Chapter 2
Fault Detection and Diagnosis

In this chapter, we start with the formal definition and formulation of the fault detection and diagnosis problem in nonlinear systems. Then, desired attributes of a fault diagnosis system and the rationale behind each attribute are discussed. A comprehensive survey and analysis of the literature on model-based and computational intelligence (CI)-based approaches to fault diagnosis is then presented with individual emphasis on the tasks of detection, isolation and identification. A number of well-known methodologies within each approach are further demonstrated, and their respective advantages and disadvantages are highlighted. Finally, the issue of robustness in fault diagnosis is introduced and briefly discussed.

2.1 Problem Formulation

In this section, the problem of detecting, isolating, and identifying faults in a general nonlinear system is formulated. Toward this end, consider a general nonlinear dynamic system described by the following nonlinear discrete-time state-space representation:

\[
\begin{align*}
  x_{k+1} &= f(x_k, u_k) = \Gamma(x_k) w_k \\
  y_k &= h(x_k) + v_k
\end{align*}
\] (2.1)

where \(x_k \in \mathbb{R}^n\) is the system state vector, \(f : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n\), \(h : \mathbb{R}^n \to \mathbb{R}^m\) are smooth nonlinear vector-valued functions (or vector fields) on their respective domains, \(u_k \in \mathbb{R}^r\) is the control input vector, \(y_k \in \mathbb{R}^m\) is the system output vector, and \(w_k\) and \(v_k\) represent system disturbances and measurement noise, respectively. The vector fields \(f\) and \(h\) represent the dynamics and output equation of the nominal model of the system. The state-dependent function \(\Gamma(.)\) essentially represents the channel over which the external disturbances are applied to the system. In many systems this function is simply a matrix gain. It is assumed that all system states are available for measurement. It is also assumed that disturbances and measurement noise are bounded signals, that is
\[
\|w_k\| \leq D_{\text{max}}, \|v_k\| \leq N_{\text{max}} \quad \forall k \in \mathcal{N}
\] (2.2)

Under full-state measurement assumption, the output equation in (2.1) can be redefined as \( y_k = Cx_k + v_k \), where \( C \) is an \( n \times n \) identity matrix.

In this monograph, we are concerned with diagnosis of faults that occur in the components of the open-loop system. More precisely, even though the performance of the proposed FDII technique will be assessed in an operational closed-loop setting, we assume that no faults may occur in the system controller. There are two main reasons for this assumption. First, modern control systems are computer-controlled and are thus more reliable and less prone to errors. Second, faults/errors that may occur in the controller software are usually handled using an entirely different error handling and accommodation mechanisms, which are mostly developed by researchers in the computer science community.

As far as the open-loop system is concerned, the system under consideration can be decomposed into three parts including sensors, actuators, and system dynamics. Figure 2.1 shows this decomposition that is also often used in practice. As can be observed from this figure, faults may occur in any of the three components of the open-loop system as described below. Furthermore, the plant dynamics and the sensor measurements are always affected by external system disturbances (or process noises) and measurement noises, respectively. A reliable fault diagnosis system should be able to distinguish faults from system disturbances and measurement noise. More precisely, the fault diagnosis system must be robust to these uncertainties while remaining sensitive to faults. The robustness of an FDII system to various sources of uncertainties is of utmost importance, which will be further discussed in Section 2.5. In the following, we will describe the various sources of faults in the open-loop system.

(i) **Sensor faults:** Sensors are basically the output interface of a system to the external world, and convey information about a system’s behavior and its internal states. Therefore, sensor faults may cause substantial performance degradation of all decision-making systems or processes that depend on data integrity for making decisions. Such systems include, but not limited to, feedback control systems, safety control systems, quality control systems, navigation systems, surveillance and reconnaissance systems, state estimation systems, optimization systems, and
particularly health monitoring and fault diagnosis systems. For example, in a feedback control system, sensors are used either to directly measure system states or to generate state estimates for the feedback control law. Thus, the presence of faults in sensors may deteriorate state estimates and consequently result in inefficient and/or inaccurate control.

Common sensor faults/failures include: (a) bias; (b) drift; (c) performance degradation (or loss of accuracy); (d) sensor freezing; and (e) calibration error [71]. Figure 2.2 depicts the effect of the above faults on system measurements.

![Figure 2.2](image)

Moreover, the mathematical representation of the above sensor faults is as follows [71]:

\[
y_i(t) = \begin{cases} 
  x_i(t) & \forall t \geq t_0 \quad \text{No failure} \\
  x_i(t) + b_i & \text{Bias} \\
  x_i(t) + b_i(t) & \text{Drift} \\
  x_i(t) + b_i(t) & \text{Loss of accuracy} \\
  x_i(t) & \text{Sensor freezing} \\
  k_i(t)x_i & \text{Calibration error} \\
\end{cases}
\]

\[
(2.3)
\]
where $t_{Fi}$ denotes the time of fault occurrence on the $i$th sensor and $b_i$ denotes its accuracy coefficient such that $b_i \in [-\bar{b}_i, \bar{b}_i]$ where $\bar{b}_i > 0$. Furthermore, it is seen that $\bar{k}_i \in [\bar{k}_i, 1]$ where $\bar{k}_i > 0$ denotes the minimum sensor effectiveness. We can represent the above cases, except the freezing case, with the following general mathematical model:

$$y = k_mx + B$$  \hspace{1cm} (2.4)

where $K_m$ is a positive definite diagonal matrix whose elements are slowly varying within $[\bar{k}_i, 1]$ and elements of vector $B$ slowly vary within $[-\bar{b}_i, \bar{b}_i]$.

(ii) **Actuator faults**: In many electromechanical or electrochemical systems, control signals from the controller (for example, a microprocessor or a microcontroller) cannot be directly applied to the system. Actuators are needed to transform control signals to proper actuation signals such as torques and forces to drive the system. Actuators are thus the control effectors of a system. Therefore, consequences of the occurrence of anomalies in system’s actuators may vary from higher energy consumption (due to an incipient fault) to total loss of control (due to total failure of an actuator).

Actuator faults are usually dependent on the actuator type. However, common types of faults have been identified for specific types of actuators. For example, common faults in control valve actuators include stuck-open, stuck-closed, and abnormal leakage. Another common set of actuator faults especially in servomotors include (a) lock-in-place (LIP) or freezing; (b) float; (c) hard-over-failure (HOF); and (d) loss of effectiveness (LOE) [72]. Figure 2.3 depicts the effect of these faults on the actuation signal.

In the case of LIP faults, the actuator “freezes” at a certain condition and does not respond to subsequent commands. Hard-over-failure is characterized by the actuator moving to upper or lower position limit regardless of the command. The speed of

![Fig. 2.3 Common types of actuator faults [73]](image-url)
response is limited by the actuator rate limit. Float fault occurs when the actuator “floats” with zero moment and does not contribute to the control authority. Loss of effectiveness is characterized by lowering the actuator gain with respect to its nominal value. Different types of actuator faults can be mathematically represented by:

\[
\begin{align*}
    u^i_a(t) &= \begin{cases} 
    u^i_c(t) & \text{Absence of Faults/Failures} \\
    k^i(t)u^i_c(t) & 0 < \varepsilon_i \leq k^i(t) < 1, \forall t \geq t_{Fi}; \text{(LOE)} \\
    0 & \forall t \leq t_{Fi}; \text{(Float)} \\
    u^i_c(t_{Fi}) & \forall t \geq t_{Fi}; \text{(LIP)} \\
    u^i_{i_{min}} \vee u^i_{i_{max}} & \forall t \leq t_{Fi}; \text{(HOF)}
    \end{cases}
\end{align*}
\]  

(2.5)

where \(u^i_a(t)\) denotes the actuation signal (or actuator output) from the \(i\)th actuator; \(u^i_c(t)\) is the control command signal (or actuator input) to the \(i\)th actuator; \(t_{Fi}\) denotes the time of fault occurrence on the \(i\)th actuator; \(k^i(t) \in [\varepsilon_i, 1]\) is the actuator effectiveness coefficient of the \(i\)th actuator with \(\varepsilon_i > 0\) being the minimum effectiveness; and \(u^i_{i_{min}}\) and \(u^i_{i_{max}}\) are the lower and upper limits on the actuation level of the \(i\)th actuator, respectively.

We can represent the above cases with the following mathematical model:

\[
u^i_a(t) = \delta_i k^i u^i_c(t) + (1 - \delta_i)\tilde{u}_i
\]

(2.6)

where \(\delta_i = 1, k_i = 1\) in the absence of faults/failures, \(\delta_i = 1, 0 < k_i < 1\) in the presence of LOE, and \(\delta_i = 0\) in other types of faults (i.e., float, LIP, and HOF) with \(\tilde{u}_i\) being the position at which the actuator is locked.

(iii) Components faults: The component faults are usually represented as cases where some condition changes in the system rendering the nominal dynamic equation of the system invalid. Component faults are also dependent on the system being monitored. Some examples include power source (e.g., battery, solar arrays) failures in satellites; leak in a tank in chemical systems or propulsion systems; body damage (e.g., wing damage, control surface damage) faults in aerial vehicles; bearing faults in rotational equipments (e.g., aircraft engines); friction faults due to lubricant deterioration; and tooth breakage and crack in gears of a gearbox system (especially in helicopters). Mathematical representation or modeling of these faults is sometimes very difficult and extensive experimentation may be needed before constructing a model. Yet, in general, component faults can be represented by a change in the system’s state equation (i.e., a change in the nonlinear function \(f\) in Eq. 2.1), being either a parametric change or a structural/functional change. We will further discuss the important issue of fault modeling in Section 3.1.

Component faults may have minor to very severe consequences. For example, an unexpected failure of the gearbox in a helicopter may cause significant economic as well as fatal loss. Nonetheless, these types of faults usually occur due to wear and tear of system components. Thus, it is extremely crucial to diagnose these faults at early stages of component degradation in order to avoid catastrophic consequences. Early diagnosis of incipient component faults allows performing timely, on-demand
maintenance operations on the degraded component, which may also involve component replacement.

Now that the general sources of faults in a general nonlinear dynamic system given in Eq. 2.1 have been identified, the fault diagnosis problem can be stated as follows:

Fault diagnosis is the problem of autonomously detecting the presence, isolating the location, and identifying the type and severity of any of the three aforementioned faults in a system. Our objective in this monograph is to simultaneously achieve FDII within a unified framework. In this monograph, we will mainly focus on FDII of component faults and actuator faults, since accurate FDII of incipient faults in components and actuators of a system is vital for enhancement of the reliability and safety of the system as well as fault prognosis and consequently condition-based maintenance (CBM). In particular, the CBM technology has recently received considerable attention from various industries and Original Equipment Manufacturers (OEMs) such as Pratt & Whitney in aircraft engines, production chain of automotive industry, etc. Nevertheless, the proposed FDII approach can be easily extended to sensor and actuator faults, since they can also be represented by the fault models developed in this monograph, which are described in Section 3.1.

2.2 Desired Attributes of a Fault Diagnosis System

A fault diagnosis system should ideally meet some general requirements. Some of the most important desirable attributes of a diagnostic system are explained in the following:

- **Early detection and diagnosis**: This refers to the capability of a diagnostic system in detecting and isolating incipient faults. Early detection and isolation of faults prior to their full manifestation into a failure is of utmost importance for fault-tolerant control of safety-critical systems as well as CBM practices. While being sensitive to incipient faults, the diagnostic system should keep false alarms under healthy operational modes of the system minimized, which poses a major challenge in achieving early detection capability.

- **Isolability**: This is the capability of a diagnostic system in distinguishing the origins of a fault from other potential fault sources or to locate a faulty component among various components of a system. While being absolutely necessary for CBM, isolation capability is also crucial to obtain fault tolerance, since proper counter-measures cannot be taken without knowing the source of an anomaly in a system. Isolability of a fault does not depend only on the diagnostic system design, but also on the way the fault affects system outputs (i.e., fault observability). Moreover, various sources of uncertainties such as modeling uncertainty/errors and system disturbances pose a serious challenge to achieve a high degree of isolability. More precisely, a diagnostic system with a high degree of isolability may be too sensitive to these uncertainties.
• **Fault identifiability**: To estimate the severity, type or nature of the fault. While being useful for fault accommodation purposes, fault identifiability is a definitive requirement for fault prognosis and eventually CBM. Accurate fault identification is usually very difficult to achieve due to presence of measurement noise, system disturbances, modeling uncertainties, and last but not the least, coupling/interactions between potential fault sources in the monitored system.

• **Robustness**: Uncertainties are inevitable in practical settings. Therefore, robustness to measurements noise, system disturbances, and modeling uncertainties is one of the most highly desirable attributes of a diagnostic system intended for practical implementations. Robustness essentially augments diagnostic system reliability and effectiveness. Due to its utmost importance, the issue of robustness of the fault diagnosis system is discussed in more details in Section 2.5.

• **Novelty identifiability**: Although the well-known, industry standard failure analysis tools such as FMEA (failure mode and effects analysis) and its recent extension FMECA (failure mode, effects, and criticality analysis) provide fruitful information on potential failure modes within a system, their effects/impacts upon the system as well as the probability of failure modes against the severity of their consequences (i.e., criticality analysis), there is still a chance of novel anomalies occurring in the system. It is expected from a diagnostic system *not* to wrongly classify novel malfunctions in the system as either an a priori known type of malfunction or a healthy operational mode. While detection of novel faults is relatively easy to achieve, isolation and identification of them is extremely difficult to accomplish, since these faults cannot be modeled due to their unknown nature.

• **Multiple fault identifiability**: This refers to the ability of a diagnostic system to identify and correctly classify multiple faults that may even coexist in a system. This is a rather difficult requirement mainly due to nonlinearities and coupling/interactions that generally exist between the states and the potential fault sources of a dynamical system. Another reason is that some faults in an engineering system are extremely difficult to model because of their complexity.

• **Explanation facility**: A diagnostic system should be able to explain where a fault originated and how it propagated in the system.

• **Adaptability**: The operating conditions of the system change due to disturbances and environmental changes. Furthermore, the system components experience performance degradation over time. Hence, a fault diagnosis should intelligently adapt to these changes in order to maintain its diagnostic performance.

• **Reasonable storage and computational requirements**: Memory and computational requirements are the two fundamental characteristics of any algorithm intended for online, real-time implementation. Diagnostic algorithms, especially the ones intended for embedded on-board fault diagnosis, are by no means an exception. Therefore, while designing a fault diagnosis system, it is necessary to keep in mind that the computational and memory requirements must always meet the specifications of the application, also including power consumption specifications. Moreover, depending on the application, a reasonable compromise between these two requirements should be made.
2.3 A Review of Analytical Redundancy-Based FDI Approaches

In Chapter 1, two fundamentally distinct approaches to the general problem of fault detection and isolation, namely the hardware redundancy-based and analytical redundancy-based approaches, were discussed and compared in details. Furthermore, a general overview of some of the analytical redundancy-based methods was provided. In this section, however, we formally investigate the analytical redundancy-based approaches and explore some of the well-known FDI techniques proposed in the literature within each approach.

The investigation of various analytical redundancy-based diagnostic approaches starts essentially with classifying them into different categories according to the form of system information (or process knowledge) utilized within each approach. In view of this, most of the existing FDI methodologies can essentially be divided into model-based and CI-based approaches. In the former, the mathematical model of the system is being used as an a priori source of information on the system being monitored. However, the latter approach utilizes either quantitative historical data of the system or qualitative information on the system in the form of if-then rules.

In this section, we investigate these two fundamentally and conceptually different approaches to FDI and some of the specific FDI methods developed within each approach will also be reviewed and analyzed.

2.3.1 Model-Based Approaches to FDI

In general, the model-based fault diagnosis approaches can be classified into two mathematically distinct categories with respect to the dynamical model and the online information/data that they use. These two categories include:

- **Discrete-event system (DES) based approaches**: These methods are pursued whenever the behavior of the system being monitored can be modeled as a finite-state machine (FSM) (or described as a discrete-event system) and the system can be observed merely as a sequence of events. Techniques under this category solve the diagnostics task by comparing the observed event sequence with the discrete-event dynamics of the model. DES-based diagnostic methods are not of interest in this monograph; however, a very good treatment of the subject can be found in Lunze and Schroder [74, 75].

- **Differential or difference equation model-based approaches**: These methods are used whenever the system being monitored can be represented by a mathematical model in the form of a differential or difference equation and the system outputs can be measured numerically. Since these systems are under consideration in this monograph, the model-based portion of the proposed hybrid fault diagnosis method falls under this category. Hence, the following sections are focused on reviewing the literature along this line of research.
Before proceeding with the literature review, it is worthwhile to mention that the above two groups of model-based diagnostic methods differ significantly in terms of their mathematical background and their associated diagnostic steps. This is due to fundamental differences between the properties of the systems that they monitor. For example, fault diagnosis approaches for continuous-variable systems are usually decomposed into two steps: residual generation and residual evaluation (see also Section 1.4), whereas in discrete-event systems, these steps cannot be defined (because the notion of a difference between events is not defined [75]). Instead, the DES-based diagnostic approaches check the consistency between the current system behavior and the DES model in a different way [74].

A quick review of literature on fault diagnosis reveals that the three tasks of FDI have not been equally investigated in the literature. This is partly due to the different levels of complexity involved in each task. In general, fault isolation and especially fault identification are more complicated than fault detection. Therefore, we need to separately review the literature corresponding to each task.

2.3.1.1 Model-Based Fault Detection

Fault detection is essentially the first step of fault diagnosis. It basically detects the presence of faults in the system. It is important to note that the detection of incipient faults (or early detection of faults) is extremely crucial for the safety of the system as well as efficient implementation of a CBM system. As was mentioned in Chapter 1, model-based fault detection is based on residual generation, where the residuals are quantities that represent the inconsistency between the actual system behavior and the mathematical model of the system.

Many residual generation methods have been proposed by various researchers in the field, some of which were reviewed in Section 1.4. Among them, nonlinear observer-based residual generation has been the most extensively studied. Observers are dynamical systems that estimate the states and consequently the outputs of a process. An observer-based residual is simply the output estimation error itself or a combination of the output estimation errors. Various nonlinear observer design techniques have been used for observer-based residual generation, since no single, universal, optimal nonlinear observer exists for all nonlinear systems. The existing nonlinear observers have to be designed usually under certain assumptions on system structure, system inputs, and/or the degree of the system nonlinearity.

In a deterministic framework, Frank et al. [44] provide a survey of the use of nonlinear observers for fault detection and isolation. More specifically, Hammouri et al. [43, 46] discuss the use of high-gain observers for fault detection of control affine nonlinear systems. Besancon and Hammouri [76] studied the observer design problem utilizing the solution of Riccati equation for Lipschitz nonlinear systems. Seliger and Frank [77] proposed nonlinear unknown input observers (UIO) as an extension of the linear UIO to a class of nonlinear systems. Ding and Frank [78] and Yang and Saif [79] proposed the use of adaptive nonlinear observers for fault
detected. Sreedhar et al. [80] designed fault detection for nonlinear systems based on sliding-mode observer.

In a stochastic setting to observer-based fault detection, Alessandri et al. [81] used extended Kalman filter (EKF) for detection of actuator faults in unmanned underwater vehicles. Caliskan and Hajiyev [82] developed an EKF-based fault detection algorithm for surface faults in aircrafts. Okatan et al. [83] developed a fault detection algorithm for magnetometers and sun sensors of the attitude determination and control subsystem of Low Earth Orbit (LEO) satellites using an approach for checking the statistical characteristics of the EKF innovation sequence. Tudoroiu et al. [84] used unscented Kalman filter (UKF) for fault detection in actuators of satellite attitude control subsystem (ACS). Finally, Li and Kadirkamanathan [85] developed a likelihood ratio approach based on particle filters [49] for fault diagnosis in nonlinear stochastic systems.

The second classical method to residual generation for fault detection is the parity space approach, which relies on analytical redundancy relations (ARR) that link a subset of selected variables of the system under consideration. The ARRs can be automatically obtained from the model equations using various elimination algorithms [58]. The ARRs can be separated into two parts. The first part depends only on known (measured) variables, while the second one, namely evaluation part, depends on the fault components. Parity residuals are generated by computing online the known part of these relations. The residual value can be interpreted by the evaluation part of the ARR [58]. Christophe et al. [58, 59] have proven that for a class of nonlinear multi-input single-output (MISO) systems a relationship exists between parity residuals and residuals generated by high-gain observers. The major drawback of the parity space approach, however, is that the residuals are computed using time derivatives of measured variables, which makes the approach very sensitive to measurement noise and system disturbances. Thus, to make it useful in a noisy environment, extra filtering and pre-processing are required. A good survey on the applications of parity space approach to nonlinear system fault detection was provided in Section 1.4.

2.3.1.2 Model-Based Fault Isolation

Once a fault is detected in a system, it should be followed by fault isolation which will distinguish (or isolate) a particular fault from others or locate the faulty component within the system. While a single residual signal is sufficient for fault detection, fault isolation requires usually a set of residuals (or a residual vector). If a residual vector can isolate all faults, it has the required fault isolability property.

Basically, there are two fundamental frameworks to create a residual set to enable fault isolation, including structured residual set and directional residual set. Almost all model-based fault-isolation methodologies can be classified into either of these two frameworks. In the following, the overall concept of each framework is individually reviewed and some of the well-known model-based fault isolation techniques within each framework are discussed.
(A) Structured residual set: One approach to fulfill the fault isolation task is to design a set of structured residuals, where each residual is designed to be sensitive to a subset of faults, while remaining insensitive to the remaining ones. The design procedure consists of two steps: the first step is to specify the sensitivity and insensitivity relationships between residuals and faults according to the assigned isolation task and the second is to design a set of residual generators according to the desired sensitivity and insensitivity relationships [24].

The structured residuals can be designed in two conceptually different ways, namely dedicated residual set and generalized residual set. These two schemes are shown in Fig. 2.4 for an example of isolating three different faults \( \{ f_1, f_2, f_3 \} \).

\[ r_i(t) > T_i \implies f_i(t) \neq 0; \quad i \in \{ 1, 2, \ldots, L \} \quad (2.7) \]

where \( L \) is the total number of faults \( (f_i) \) to be isolated and \( T_i(i = 1, 2, \ldots, L) \) are thresholds corresponding to residuals \( r_i(i = 1, 2, \ldots, L) \). The dedicated residual set is very simple and all faults can be detected simultaneously; however, there is normally no design freedom left to achieve other desirable attributes of a fault diagnosis system such as robustness to various sources of uncertainties (i.e., measurement noise, system disturbances, and modeling errors). As will be seen in Chapter 3 and demonstrated in Chapter 5, some of the characteristics of the series-parallel FDI scheme proposed in this monograph are similar to that of the dedicated scheme. In particular, a portion of the fault isolation decision logic of the series-parallel scheme is analogous to that of the dedicated scheme. Furthermore, both methods are equally sensitive (or non-robust) to measurement noise.

Various fault isolation techniques have been developed in the literature under the dedicated scheme. Clark [86], in his pioneering work, designed a dedicated observer scheme (DOS) for sensor fault detection, which was actually (and surprisingly) the original inspiration for the concept of dedicated residual set (or dedicated...

---

**Fig. 2.4** Two schemes of structured residual set: (a) dedicated scheme and (b) generalized scheme for an example of isolating three faults [24]
scheme). In DOS, an observer reconstructs all of the system outputs except one (i.e., \( y_j(t), \ j = 1, \ldots, m, \ j \neq i \)) using all of the system inputs and the only left-out output, namely \( y_i \). Then, the difference between the estimate and the measurement indicates the possibility of a fault in the \( i \)th sensor. If this technique is applied for all \( m \) outputs of the system, namely \( y_i, \ i = 1, \ldots, m \), then a bank of \( m \) dedicated observers are needed to monitor \( m \) sensors of the system. Chen and Saif [87] recently extended Clark’s DOS to actuator fault isolation. Their scheme is able to detect and isolate multiple actuator faults using a bank of \( r \) observers, where \( r \) is the total number of actuators in the system under consideration.

Another very important group of fault isolation methods that essentially fall under the dedicated scheme are the multiple-model (MM) approaches. Over the past few decades, the use of multiple models has become very popular and widely applied across various domains of research including state estimation, control, target tracking, and fault diagnosis of stochastic systems. In the literature, there are mainly two types of MM algorithms, namely non-interacting MM and interacting MM (IMM). Non-interacting MM approach was originally proposed by Magill [87] for optimal adaptive estimation of sampled linear stochastic processes. As mentioned therein, the MM estimator is composed of a set of elemental estimators and a corresponding set of weighting coefficients. However, the model-based elemental filters independently operate in parallel at all times without any interaction between them. Such an approach is not suitable for fault diagnosis problem since it assumes that there are no mutual interactions among the multiple models, whereas in general, the system structure or parameters do indeed change as a system component (as well as a sensor or an actuator) fails. Nonetheless, the MM approach has also been developed for fault diagnosis in different engineering applications but mainly for the purpose of detection rather than isolation. For example, see Laparo et al. [88] on leak detection in heat exchanger systems and Manke and Maybeck [89] on sensor/actuator failure detection in the Vista F-16 fighter aircraft. Furthermore, Alessandri et al. [81] used a bank of non-interacting extended Kalman filters (EKF) for isolation of faults in actuators of unmanned underwater vehicles.

The interacting multiple-model (IMM) approach, initially proposed by Blom and Bar-Shalom [90] for state estimation of stochastic systems, presented a notable advance to MM-based estimation (also see the book by Bar-Shalom et al. [91] for more details on IMM and its application to tracking and navigation). The IMM approach uses modal probabilities to weight the inputs and outputs of a bank of parallel filters at each instant of time. Furthermore, the IMM approach overcomes the weakness of the non-interacting MM approach by explicitly modeling the abrupt changes of the system by “switching” from one model to another in a probabilistic manner. This approach is one of the most cost-effective adaptive estimation techniques for systems involving structural as well as parametric changes [67].

Faults/failures usually create structural and parametric changes in the system. Since the IMM approach explicitly models and effectively handles the structural and/or parametric changes in the system, it presents a very promising and effective candidate approach for fault detection and isolation. Mehra et al. [92] and Zhang and Xiao [67] independently and almost simultaneously proposed IMM
approach for fault detection and diagnosis for the first time. The IMM-based nonlinear fault diagnosis assumes that the system being monitored can be modeled, at any time, sufficiently accurately by the following jump Markov hybrid nonlinear system [67]:

\[
x(k + 1) = f(k, m(k + 1), x(k), u(k)) + T(k, m(k + 1)), w(k, m(k + 1))
\]
\[
z(k) = g(k, m(k), x(k), u(k)) + v(k, m(k))
\]

with \( x_0 \sim N(\hat{x}_0, P_0) \); where the mode of the system at time \( k \) is selected by a discrete process \( m(k) \) that is modeled as a discrete-time, \( L \)-state, first-order Markov chain with transition probabilities \( \pi_{ij}(k) \) given by:

\[
\pi_{ij}(k) = P\{m_j(k+1)|m_i(k)\}, \forall m_i, m_j \in S
\]

where \( \pi_{ij}(k) \) is the probability of the transition from mode \( i \) at time-step \( k \) to mode \( j \) at time-step \( k+1 \) and

\[
0 \leq \pi_{ij}(k) \leq 1, i = 1, \ldots, N; j = 1, \ldots, N;
\]
\[
\sum_j \pi_{ij}(k) = 1, i = 1, \ldots, N
\]

where \( S = \{m_1, m_2, \ldots, m_L\} \) is the set of all possible modes of the system including healthy and various faulty modes and \( L \) is the total number of modes in \( S \).

In IMM-based fault diagnosis, one mathematical model has to be designed per mode in the set \( S \). This is the so-called model set design step of the IMM approach. This is the initial and a key step in IMM approach because the model set has to be designed such that it represents as many system modes as possible. Therefore, the design of a proper model set requires a priori knowledge of the potential system faults/failures.

Once a model set is designed, a model-based recursive filter has to be designed based on each model in the IMM model set in order to estimate system states. Various stochastic filtering techniques can be used for this purpose. The filter that has been commonly used for nonlinear systems is the extended Kalman filter (EKF) (see, for instance, Zhang and Xiao [67] and Tuduroiu and Khorasani [93]). More recently, Tuduroiu et al. [84] developed an interactive bank of unscented Kalman filters for fault detection and isolation in the reaction wheel actuators of the satellite attitude control subsystem.

Each filter in the IMM bank recursively calculates a model-conditional estimate of the system states and then these estimates are combined to obtain an overall estimate, also called mixed estimate, of system states. The mixed estimates are calculated using the so-called model (or mode) probabilities. It should be noted that the model probabilities are different from transition probabilities introduced above. The transition probabilities comprise a matrix that is a parameter of the IMM algorithm and is usually set to a fixed value; however, the model probabilities comprise a vector \( (\mu_i, i = 1, 2, \ldots, N \text{ in Fig. 2.5}) \) that is essentially part of the state vector of the
IMM algorithm and is recursively updated at each time-step of the algorithm operation. The model probabilities at each instant of time represent the probability of each mode currently in effect. Therefore, the largest model probability indicates clearly the mode in effect at that instant, hence fault is isolated. Furthermore, the value of the largest model probability provides a quantitative measure of the confidence level of IMM-based diagnoser in its decision, which is almost an exclusive property of IMM-based fault diagnosis. This can definitely be considered as an advantage of the IMM method, since the confidence information can be very effectively used for information fusion in fault diagnosis systems comprising of more than one diagnoser (or decision-maker). Figure 2.5 depicts the block diagram representation of the IMM-based fault diagnosis algorithm.

(A-2) Generalized scheme of the structured residual set: The generalized scheme for designing the structured residual consists of making each residual sensitive to all but one faults [24], i.e.,

$$
\begin{align*}
    r_1(t) &= R(f_2(t), \ldots, f_L(t)) \\
    \cdots \\
    r_i(t) &= R(f_{i-1}(t), f_{i+1}(t), \ldots, f_L(t)) \\
    \cdots \\
    r_L(t) &= R(f_1(t), \ldots, f_{L-1}(t))
\end{align*}
$$

(2.11)
The above set of residuals is defined as generalized residual set. If a bank of observers is used for generation of all residuals in the generalized residual set (i.e., a bank of observer-based residual generators), the structure is known as the generalized observer scheme (GOS) (see survey paper of Frank [94] and Lunze and Schroder [75] for application of GOS to sensor and actuator fault diagnosis of discrete-event systems). The isolation in generalized scheme can be performed by using the following logic [24]:

\[
\begin{align*}
    r_i(t) & \leq T_i \\
    r_j(t) & > T_i \quad \forall j \in \{1, \ldots, i - 1, i + 1, \ldots, L\}
\end{align*}
\]  

(2.12)

for \( i = 1, 2, \ldots, L \).

The GOS-based FDI, depicted in Fig. 2.6 for both sensor and actuator fault detection and isolation, is more robust than DOS with respect to parameter uncertainties and measurement noise. This is mainly due to the fact that in GOS, more than one output \( y_i \) is fed into the observers [94], as can also be seen in Fig. 2.6.

\( \)
vector, also called detection filters (see Beard [95] and Jones [96]). A directional residual set is a vector that lies in a fixed and fault-specific direction (or subspace) in the residual space, in response to that particular fault [24]. In mathematical notation, we want to have:

$$r(t|f_i(t)) = \beta_i(t)\vec{l}_i; \quad i = 1, 2, \ldots, L$$

(2.13)

where the constant vector $\vec{l}_i$ is the signature direction of the $i$th fault in the residual space and $\beta_i$ is a scalar that depends on the fault size and dynamics [24]. A fault is then isolated by determining the fault signature direction that is the closest to the generated residual vector. Therefore, in order to isolate faults reliably (i.e., to reduce incorrect isolation rate) there must be a one-to-one correspondence between fault signatures and potential fault sources (i.e., each fault signature must be uniquely associated with one fault).

Although directional residual set is simpler to implement (not necessarily to design, which is more problem dependent) than the structured residual set and it also provides more reliable fault isolation capability under ideal conditions, it is really difficult to make it robust against various sources of uncertainties, especially modeling errors and system disturbances.

A number of fault-isolation methods have been proposed in the literature within the model-based directional residual generation framework. Fault detection filters proposed by Beard [95] and Jones [96] (also known as Beard–Jones fault detection filter) is one of the pioneering methods that has actually inspired the directional residual concept. Being originally designed for FDI of linear systems, fault detection filter is a Luenberger observer-based method, where the observer gain is chosen so that the direction of the residual vector in the output residual space can be used to identify the failed component. Note that in the Beard–Jones detection filter design, faults are viewed as inputs and the residuals are viewed as outputs.

The Beard–Jones detection filter, developed following the directional residual set concept, has also inspired the celebrated geometric approach to fault isolation which indeed falls under the dedicated residual set category. Massoumnia [97] first proposed a geometric formulation of the Beard–Jones fault detection filter problem for linear systems using the concept of unobservability subspaces, which is a subspace in the residual space that can be made “unobservable” via “output-reduction” and “output-injection” leading to a quotient (observable) subsystem unaffected by all faults except one. This approach is known in the literature as the geometric approach to FDI. Later, Massoumnia et al. [98] proved that the basic necessary and sufficient condition for the fault-isolation problem to be solvable is the existence of an unobservability subspace. The unobservability subspace can be determined by means of a simple recursive algorithm. Massoumnia et al. [98] also showed that the geometric approach to FDI is the dual version of the problem of non-interacting control by means of dynamic feedback.

De Persis and Isidori [99] extended Massoumnia’s geometric approach to nonlinear systems by proposing a differential-geometric approach that gives the necessary and sufficient conditions for solving the problem of nonlinear FDI. Detailed
description of their approach requires many background mathematical definitions
and concepts, which is out of the scope of this monograph, but the interested reader
can refer to the work of De Persis Isidori [99] for further details and information.
The background mathematical concepts and definitions can also be found in Isidori’s
book [100]. To put it in a nutshell, in a nonlinear geometric approach an unobserv-
ability distribution is computed by means of suitable algorithms, which results in a
coordinate transformation in the state and the output space of the system that induces
and “observable” quotient subsystem unaffected by all faults but one. Then, a fault
detection filter (i.e., a nonlinear observer) is designed for the quotient subsystem.

In mathematical notations, in the nonlinear geometric approach it is assumed that
the nonlinear system can be described by the following model:

\[
\dot{x} = f(x) + \sum_{i=1}^{r} g_i(x)u_i + \sum_{i=1}^{L} l_i(x)m_i + \Gamma(x)w
\]

\[
y = h(x)
\]  

(2.14)

where \(u_i, i = 1, \ldots, r; \ m_i, i = 1, \ldots, L\) and \(w\) denote the input channels for
control purposes, the fault/malfunction signals whose occurrence has to be detected
and isolated and system disturbance signals, respectively. The objective is then to
find for each fault signal \(m_i, i = 1, \ldots, L\), a quotient subsystem that is affected by
the fault signal \(m_i\) and decoupled from other faults \(m_j, j = 1, \ldots, L; j \neq i\) (and,
if possible, decoupled from disturbance \(w\) to also achieve robustness with respect
to disturbance). The algorithm that verifies the existence of such a quotient sub-
system is a constructive algorithm that provides the state coordinate transformation
\(z = \Phi(x)\) required for fault isolation. Once the coordinate transformation function
\(\Phi(.)\) is found and applied to the system’s state-space equations, the new state-space
representation of the system in terms of the transformed coordinate \(z\) is obtained.
Then, a nonlinear observer is designed for estimating \(z\), with the residual defined as
\(r = z - \hat{z}\), where \(\hat{z}\) is the estimate from the observer.

The main power of the nonlinear geometric approach is in providing necessary
and sufficient conditions for the solution of fault-isolation problem, supported by
unprecedented rigorous mathematical proofs. Furthermore, under full-state mea-
surement conditions (i.e., the function \(h(.)\) in Eq. 2.14 is simply a unity matrix),
finding the coordinate transformation \(\Phi(.)\) is fairly simple (though, this is not always
the case under partial-state measurement). However, it also possesses some draw-
backs. Its major drawback is lack of robustness to modeling errors. Since the trans-
formation \(\Phi(.)\) is obtained based on system’s nominal equations, any discrepancies
between the actual system and its nominal model (due to unmodeled dynamics,
parameter uncertainties, parameter variations, etc.) may render the analytical results
invalid. To a lesser extent, the measurement noises will also affect the performance
of the geometric approach. As far as robustness to system disturbances are con-
cerned, sometimes little design freedom is left to decouple residuals from distur-
bances.
2.3.1.3 Model-Based Fault Identification

Despite its undeniable importance, model-based fault identification has received less attention from the research community as compared to model-based FDI. This is especially true for nonlinear systems. Nonetheless, possibly the first formal effort to estimate the severity of faults is in the seminal works of Isermann [26, 60]. In his work it is assumed that faults are reflected in the physical parameters of the system; hence, faults can be identified through online estimation of system parameters. However, the parameter estimation approach of Isermann was developed for linear systems due to availability of very well-known linear parameter estimation methods. More recently, Tan and Edwards [101] applied the concept of “equivalent output estimation error injection” – proposed by Edwards [45] – to reconstruct faults for linear systems using sliding-mode observers. Once again, however, their approach was developed for linear systems.

Chen and Saif [87] recently extended the approach proposed by Tan and Edwards [101] to actuator fault identification in a class of nonlinear systems. More specifically, they modified the approach proposed by Tan and Edwards [101] in two ways. First, instead of linear systems they consider a specific class of uncertain nonlinear systems. Second, instead of reconstructing only faults, they reconstruct the inputs and the faults at the same time. They estimate actuator faults using equivalent control method in the sliding-mode observer design. Nevertheless, their approach also has two limitations. First, it has been developed specifically for actuator faults and its application to identification of component faults in a nonlinear system has not been discussed. Second, it is applicable only to a specific class of nonlinear systems rather than a general one.

One may also use multiple-model (MM) approach for fault identification, where multiple models in the model bank correspond to different levels of fault severity. However, this will introduce an inevitable quantization error in fault estimation. This quantization error can be reduced as more models are used in the bank. But the use of more models will increase the computational requirements of the algorithm. In order to very precisely identify faults, ideally infinite number of models (or quantization levels) should coexist in the model bank, which makes the approach computationally unfeasible and thus impractical. The IMM approach, to a lesser extent has a similar problem, though Zhang and Xiao [67] suggest that in the IMM approach, the magnitude (size) of a fault can be determined by the probabilistically weighted sum of the fault magnitudes of the corresponding partial fault models. However, this idea has not been well elaborated and, as was mentioned previously, the fine-tuning of the IMM approach is not easy to accomplish especially if precise fault identification is required. Zhang and Jiang [172] have also developed a two-stage adaptive Kalman filter (or a dual Kalman filter) for simultaneous (or joint) state and fault parameter estimation, which is applicable to identification of only actuator (not component) faults.

2.3.2 Computational Intelligence-Based Approaches to FDI

The model-based approaches to fault diagnosis rely on the analytic mathematical model of the process being monitored. This implies that the accuracy of the
model has direct impact on diagnostic system performance and reliability. More precisely, the more accurate the model, the more reliable will be the model-based fault diagnosis scheme. However, for complex and uncertain systems, the derivation of high-fidelity mathematical models from physical principles can become very complicated, time consuming, and even sometimes unfeasible (for instance, some systems cannot be represented accurately enough by a lumped parameter system). Moreover, even with the possibility of deriving a mathematical model using first principles, obtaining accurate model parameter values may become a very tedious job or even practically impossible due to proprietary issues regularly imposed by OEMs and/or system integrators. Last but not the least, some systems exhibit uncertain behaviors such as higher order dynamics and high-frequency oscillations, collectively called unmodeled dynamics, which cannot be precisely modeled.

Mathematical methods in computational intelligence and learning theory – neural networks, fuzzy-logic, neuro-fuzzy systems, and genetic algorithms – represent a promising way of circumventing the above-mentioned modeling precision problems in model-based fault diagnosis. Indeed, during the past decade, computational intelligence (CI)-based fault diagnosis methods have been extensively developed and successfully applied to various engineering systems. A number of survey papers and books in the literature review the use of CI techniques in fault diagnosis. Among the pioneers is the survey paper of Patton et al. [27] that outlines some of the residual generation methods based upon artificial intelligence techniques, which integrate both quantitative and qualitative knowledge of the system in fault diagnosis. More recently, Palade et al. [102] published a book consisting of a set of papers reviewing the main CI techniques and their applications to fault diagnosis. They have also discussed the main advantages and disadvantages of each methodology. It is also shown that hybrids of CI-based diagnostic techniques are often used in practice to utilize their individual advantages and overcome their individual disadvantages. Two other recent books, namely Korbics et al. [63] and Vashtsevanos et al. [3] also review intelligent fault diagnosis methods.

Using CI-based techniques enables one to exploit both quantitative (numerical) and qualitative (symbolic) information about the system being monitored. Qualitative information is normally expressed in the form of Boolean or fuzzy if-then rules. For systems represented by Boolean rules, causal reasoning and fault tree analysis methods have been historically used particularly in aerospace and nuclear industries (see Zampino [103] and the pioneering work of Crosetti [104], respectively). On the other hand, fuzzy-logic is the right tool for fault diagnosis whenever the system behavior is described by a set of fuzzy if-then relations derived either by an expert or using qualitative physics. More details regarding the use of fuzzy models for fault diagnosis can be found in Dexter [105] and Mendonca et al. [106]. The key advantage of the qualitative CI-based approaches is that they can provide valuable information for the system operators to identify the root cause of anomalies (i.e., the series of events/anomalies that ended up a failure).

Though seemingly attractive, qualitative CI-based fault diagnosis methods also suffer from a major drawback. In many engineering applications, deriving Boolean and/or fuzzy if-then rules is by no means straightforward and requires extensive
expert knowledge of the system. Instead, the knowledge that describes the system behavior is contained in large quantitative datasets stored in databases. Neural networks are ideal mathematical tools for such situations due to their universal nonlinear function approximation property (Cybenko theorem; see Cybenko [107]) and their ability to learn and reproduce system behavior from quantitative system datasets (i.e., historical system input–output data). Neural networks do indeed provide an excellent framework for identification of nonlinear systems (see the seminal work of Narendra and Parthasarathy [65]).

All these properties make neural networks a promising tool for applications as diverse as feature extraction, pattern recognition, clustering, classification, information integration, and as mentioned above in system identification, which all can effectively be applied for fault diagnosis and health monitoring. As a result, neural networks have been extensively applied to fault diagnosis. In the following, we will review three of the most commonly used neural network (NN)-based approaches to fault diagnosis.

(I) Neural network-based pattern recognition approach to fault diagnosis:
In pattern recognition approaches, neural networks are used mainly for feature classification. In other words, the neural network is only used as a fault classifier. For example, in Li et al. [108], the bearing vibration frequency features and time-domain characteristics are applied to a neural network to build an automatic motor bearing fault diagnosis machine. In these applications, neural networks are merely used to examine the possibility of a fault or abnormal features in system measurements and give a fault classification signal to declare the health state of the system. This approach of using only system output measurements produces valid fault diagnosis results mainly for static systems or steady-state processes. However, this is not usually the case for fault diagnosis of dynamic systems (especially nonlinear ones), where a change in system inputs can also affect certain features of system outputs. Therefore, the NN-based pattern recognition approach to fault diagnosis of nonlinear dynamic systems can generate incorrect fault information while only the system inputs have been changed. This problem has been resolved by the following second approach to NN-based fault diagnosis.

(II) Neural network-based residual generation decision-making scheme:
This NN-based diagnostic scheme was initially proposed by Patton et al. [109]. In this scheme, depicted in Fig. 2.7, neural networks are utilized at two stages: residual generation and decision-making (for fault isolation). At residual generation stage, neural networks are used as prediction models. An important feature of a neural network-based prediction model is that it will automatically “learn” the nonlinear system dynamics during the training process made over several training cycles, with training data coming from historical input–output data of the system. Neural network-based prediction models have potential advantages over traditional prediction and estimation methods, including powerful nonlinear mapping properties, noise tolerance, self-learning and self-adapting, and parallel processing capabilities.

Various NN-based nonlinear system identification architectures can be used as prediction model at residual generation stage. Three widely used architectures
include nonlinear autoregressive exogenous (NARX) model neural networks, recurrent neural networks, and dynamic neural networks, as shown in Fig. 2.7. These architectures differ in terms of the way dynamics have been introduced into the network architecture. In the following, we will briefly review the literature on NN-based identification of nonlinear dynamic systems.

A large body of literature has been dedicated to the identification of nonlinear dynamic systems using neural networks. These efforts are justified by the following four important features of neural networks, namely (i) their nonlinear characteristics that make them suitable for dealing with nonlinear systems, (ii) their parallel and pipeline processing characteristics that allow them to perform computations more efficiently, (iii) their self-learning and self-adapting characteristics that are ideal for adapting to different environmental conditions, and (iv) their tolerance to noise.

One may classify NN-based nonlinear dynamic system identification schemes into four main categories. The first category utilizes tapped delay lines (TDL) along with a static neural network in its structure. The TDLs are used to introduce dynamics into the network by generating delayed inputs and outputs of the system that are then fed to a static network as the regressor vector. The network then performs a static nonlinear map on this regressor vector so that the desired output is obtained. This model is called nonlinear autoregressive exogenous (NARX) model. For further details, refer to Narendra and Parthasarathy [65].
The second category is recurrent neural networks. In this approach, a dynamic input–output representation is constructed using a recurrent structure. This method has been investigated in Funahashi and Nakamura [110] and Ku and Lee [111]. More specifically, Funahashi and Nakamura [110] proved that the proposed recurrent neural network is capable of identifying any nonlinear dynamic system provided that the initial states of the network are chosen appropriately with respect to the initial conditions of the system.

The third category is embedded dynamic neural networks. The embedded dynamic neural networks are constructed by utilizing dynamic neurons whose model is different from that of the static neurons. In the former, one or more dynamic elements are utilized to obtain a specific dynamical input–output map. Several dynamic neuron structures have been reported in the literature. Atiya and Parlos [112] introduced a spatio-temporal neuron in which the conventional weight multiplication operation was replaced by a linear filtering (an all zero filter) operation. Gamma neuron model was developed by Principe and Motter [113] for identification of nonlinear systems. The structure of the Gamma model is similar to the TDL structure, but instead of using simple shift elements in the line, a first-order linear filter is utilized to generate a dynamic input–output map. Yazdizadeh and Khorasani [114] introduced an embedded dynamic neural network in which adaptive linear filters are augmented before the NN’s hidden-layer activation functions in order to generate a dynamic input–output map. In this network, learning takes place by adapting both the embedded linear filter parameters and the neural network weights. The well-known time delay neural network (TDNN) was first introduced by Waibel et al. [115] for phoneme recognition. In TDNN, each weight is associated with a delay. The adaptive version of TDNN was introduced by Yazdizadeh [116] for identifying two classes of nonlinear dynamic systems denoted as “the first” and “the fourth” class of nonlinear systems by Narendra and Parthasarathy [65].

The fourth category of dynamic neural identifiers, which is proposed by Abdollahi et al. [117], consists of a feed-forward static neural network architecture cascaded/followed by a fixed stable linear filter. During the training/learning process, neural network weights comprise the only adaptive parameters of the proposed dynamic neural identifier and the parameters of the stable linear filter remain unchanged.

In the second stage of the NN-based fault diagnosis scheme, namely decision-making stage, a neural network-based classifier is used to partition the residual vector to patterns corresponding to different healthy and faulty modes of the system. The NN-based classifier is trained to recognize complex features in residuals and then generates fault detection and isolation information. The training can take place in both supervised and unsupervised modes; however, supervised classifiers are generally more accurate. Nonetheless, they have a major disadvantage of requiring data from all possible fault situations for classifier training. A supervised NN-based classifier trained using only fault-free situations cannot be expected to perform well for faulty situations.

(III) Neural network-based multiple-model residual generation and classification: This NN-based fault diagnosis scheme, originally proposed by Patton et al.
Fig. 2.8 A generic neural network-based multiple-model fault detection and isolation scheme

[27], follows the idea of multiple-model based FDI scheme described in Section 2.3.1.2, where the mathematical models have been replaced by parallel NN-based dynamic identifiers. This scheme, depicted in Fig. 2.8, also consists of two stages: **NN-based multiple-model residual generation** and **isolation decision-making**. In the former stage, each fault model in the residual generation block is a dynamic neural network that identifies a class of system behavior. The dynamic neural identifiers that were discussed in the previous NN-based fault diagnosis scheme are equivalently applicable in here. The major difference is that, as opposed to residual generation decision-making scheme, the NN-based multiple-model scheme requires data from all healthy and faulty situations at residual generation stage in order to be able to learn all classes of system behavior. This can be considered as one of the main drawbacks of the NN-based multiple-model approach.

In the isolation decision-making stage, the main task is to classify the generated residuals into a number of distinguishable patterns corresponding to different healthy as well as faulty situations. Thus, another neural network is used for this purpose based on the classification capability of neural networks. Once again, various NN-based classifier architectures and algorithms can be utilized at this stage. These include multi-layer perceptron (MLP) network, radial basis function (RBF) network, support vector machines (SVM), probabilistic neural networks (PNN), and fuzzy neural networks for supervised classification; and competitive neural networks (e.g., Kohonen network, self-organizing maps (SOM)), and adaptive resonance theory (ART) networks (i.e., ART-II, fuzzy-ART) for unsupervised classification.
The above-mentioned CI-based diagnostic methods use either qualitative or quantitative information about a system in order to achieve fault diagnosis. Both methodologies have been successfully applied to fault diagnosis of various engineering systems; however, integrating both quantitative and qualitative information can greatly enhance the diagnostic system performance and robustness. Such diagnostic systems are collectively called integrated computational intelligence-based fault diagnosis systems. There are basically two main ideas within the integrated CI-based framework. One is to generate residuals using NN-based methods and then allocate the decision-making (or isolation decision-making) process to a fuzzy-logic inference engine. This approach allows system operators to describe the system behavior or the fault-symptom relationship with simple if-then rules.

The second integrated CI-based diagnostic concept, depicted in Fig. 2.9 from Chen and Patton [24], revolves around using neural networks for two main purposes: (i) residual generation using quantitative historical input–output data of the system and (ii) learning (or determining) the parameters of the fuzzy model of the system (i.e., the fuzzy if-then rules that qualitatively describe the system behavior) from quantitative data of the system. This integration of quantitative and qualitative knowledge of the system is accomplished through a neuro-fuzzy system (or a fuzzy neural network) that makes it feasible to combine the learning ability of neural networks with the explicit knowledge representation of fuzzy-logic. According to Patton et al. [27], a potential way of implementing a neuro-fuzzy system is to use B-Spline neural networks. For further information on neuro-fuzzy modeling, refer to Brown and Harris [118].

Moreover, a trained neural network can be used to evaluate the reliability of information provided by either quantitative or qualitative methods and decide which has to be accordingly weighted in the information fusion, as depicted in Fig. 2.9.

2.4 Methodology Developed in This Monograph: Hybrid Approach to FDII

The approach proposed in this monograph is essentially a hybrid approach to fault diagnosis. More precisely, the proposed fault diagnosis methodology simultaneously exploits both the a priori mathematical model information of the system and the nonlinear approximation and adaptation capability of neural networks.
More specifically, mathematical model of the system is used as a basis for fault modeling and isolation, and the capability of neural networks in adaptive nonlinear function approximation is used as a basis for online fault severity identification.

Only a few fault diagnosis methodologies exist in the literature, which simultaneously take advantage of mathematical model of a system and the exclusive capabilities of CI techniques, especially neural networks, in a hybrid framework. For example, Alessandri [69] proposed a hybrid approach to fault detection in nonlinear systems. In his work, fault detection and isolation is accomplished by means of a bank of estimators, which provide estimates of parameters that describe actuator, plant, and sensor faults. These estimators, also called finite-memory filters, perform according to a receding-horizon strategy and are designed using nominal mathematical model of the system and the models of the failures. The problem of designing such estimators for general nonlinear systems is solved by searching for optimal estimation functions. These functions are approximated by feedforward neural networks and the problem is reduced to finding the optimal neural weights, hence the name finite-memory neural filters. The learning process of the neural filters is split into two phases: an offline initialization phase using any possible “a priori” knowledge on the statistics of the random variables affecting the system states and an online training phase for online optimization of neural weights.

In another example of hybrid approach to diagnostics, Xiaodong et al. [119] presented a robust fault detection and isolation scheme for abrupt and incipient faults in nonlinear uncertain dynamic systems. The diagnostic architecture proposed therein consists of a bank of \( N+1 \) nonlinear adaptive estimators, where \( N \) is the number of potential faults that may affect the nonlinear system. One of the nonlinear adaptive estimators is the fault detection and approximation estimator (FDAE) used to detect faults. The remaining ones are fault-isolation estimators (FIEs) that are used for isolation purposes only after a fault has been detected. Under normal operating conditions (without faults), the FDAE is the only estimator monitoring the system. Once a fault is detected, the bank of FIEs is activated and the FDAE adopts the mode of approximating the fault function. The nominal mathematical model of the system is explicitly used for designing both FDAE and FIEs. Furthermore, a key component of FDAE is an online approximator, which, in presence of a fault, provides the adaptive structure for online approximation of the unknown nonlinear fault function. This is where the extreme capability of neural networks in adaptively representing nonlinear multivariable functions is employed to implement the online approximator of FDAE.

Very recently, Talebi and Khorasani [120] presented a hybrid intelligent fault detection and isolation scheme for a general nonlinear system using a neural network-based observer. The proposed NN-based observer employs nominal mathematical model of the system in conjunction with two recurrent neural networks, which are used to identify general unknown actuator and sensor faults. The distinct advantage of their method is that, unlike many previous methods in the literature, it does not rely on the availability of full-state measurements.

The above works, however, either have not addressed the important problem of fault severity estimation (or fault identification) or have addressed it in a way that is not of use to fault prognosis and consequently condition-based maintenance (CBM).
More precisely, the approach proposed by Alessandri [69] is only a fault detection and isolation method, leaving fault identification problem unsolved. On the other hand, the approaches proposed by Xiaodong et al. [119] and Talebi and Khorasani [120] estimate/identify the fault function that represents the overall impact of faults on system states. Though estimating this overall impact is often sufficient for fault accommodation (and thus achieving fault-tolerant control) and is also useful for identifying actuator faults (especially, static actuators or actuators with negligible dynamics), it is not appropriate for fault prognosis and CBM of system components. The reason is that it is either impossible or extremely difficult to obtain fault trend information for a specific system component from the aforementioned fault function estimate.

The hybrid fault diagnosis approach presented in this monograph, however, is able to detect, isolate, and identify the severity of faults in components of a general nonlinear system within a unified, integrated framework. This is achieved through the use of a bank of parameterized fault models and a corresponding bank of adaptive neural parameter estimators (NPEs) to estimate fault parameter (FP) vector and thus fault severities. The nominal mathematical model of the system is used in both PFM bank and NPE design, and neural networks are used in NPE design; hence being a hybrid approach to fault diagnosis.

Finally, in order to achieve FDII under partial-state measurement, a separate nonlinear observer is designed to continuously estimate system states from inputs and measurements even in presence of faults in system components. We call such an observer a fault-tolerant observer (FTO) or a fault-tolerant state estimator (FTSE). To the best of our knowledge, the FTO terminology proposed in this monograph appears for the first time in the literature. A similar concept of fault tolerance in state estimation has been investigated in the literature under the unknown input observer (UIO) terminology. However, the UIOs have been developed and employed in the literature as a means of making fault diagnosis algorithms robust with respect to unknown uncertainties such as modeling errors and external disturbances. In other words, instead of faults, the modeling errors and external disturbances are modeled as unknown inputs and the UIOs are designed in order to decouple the state estimates from these uncertainties.

The FTSE method proposed in this monograph is a Kalman filter structure preserving neural state estimator (NSE). It is a hybrid approach to nonlinear filtering, since it utilizes both mathematical model of the system and the adaptive nonlinear function approximation capability of neural networks. Chapter 4 discusses, in more details, the proposed NSE and its integration with the proposed FDII method in order to achieve fault diagnosis under partial-state measurement.

### 2.5 Robustness of FDI to Uncertainties

Model-based fault diagnosis (FD) approaches rely on the key assumption that a perfectly accurate and complete mathematical model of the system under supervision is available. However, such assumption is usually not valid in practice since it
2.5 Robustness of FDI to Uncertainties

is difficult to obtain the necessary modeling accuracy required for construction of reliable analytical redundancy-based FD architectures. Unavoidable modeling and environmental uncertainties that arise due to modeling errors, parameter variations, time variations, unknown external disturbances, and measurement noise deteriorate the performance of the FD schemes by causing false alarms. This performance deterioration can happen to an extent that makes the model-based FD scheme totally useless. This necessitates the development of FD algorithms, which have the ability to reliably detect, isolate, as well as identify faults and failures in presence of various sources of uncertainties. Such algorithms are referred to as robust fault diagnosis algorithms.

To overcome the difficulties introduced by modeling and environmental uncertainties, a model-based FDII has to be made robust, i.e., insensitive to uncertainty [24]. However, merely reducing the sensitivity to uncertainties may not solve the problem because such a sensitivity reduction may be undesirably accompanied by a reduction of sensitivity to faults. Therefore, a more meaningful formulation of the robust FDII problem is to increase the robustness to various sources of uncertainty without losing sensitivity to faults. In other words, an FDII scheme that is designed to provide satisfactory sensitivity to faults, associated with the necessary robustness with respect to modeling and environmental uncertainties, is called a robust FDI scheme.

The importance of robustness in model-based FDI has been widely recognized by both academia and industry. More specifically, robust FDI for linear systems has been extensively investigated by many researchers during the last two to three decades. As a result, a number of methods have been proposed to tackle the linear robust FDI problem [24] such as the UIO method [121], eigen-structure assignment [122], and optimally robust parity relation methods [123].

Traditionally, the robust FDI problem for nonlinear dynamic systems has been approached in two steps. The model is first linearized around an operating point, and then robust linear FDI techniques are applied to generate residuals that are insensitive to uncertainties but responsive to faults. This method only works well when the linearization does not cause a large mismatch between linear and nonlinear models and the system operates close to the specified operating point. As another alternative to robust nonlinear FDI, one might think of just simply increasing the threshold levels of the residuals generated by the nonlinear FDI scheme and thus reducing the number of false alarms. However, the increase in the threshold levels will at the same time decrease the fault sensitivity of the FDI scheme.

This imposes a tradeoff between reducing the number of false alarms and the number of missed alarms (i.e., missing to detect the presence of an actually occurred fault). A reliable solution to such a trade-off problem is not trivial in practice especially due to the nonlinear behavior of the system dynamics and the presence of different sources of unknown uncertainties. Therefore, there is a high demand for development of techniques that make the nonlinear FDII problem robust to modeling and environmental uncertainties to remarkably reduce the number of false alarms when the nonlinear system is under healthy mode of operation, while reliably diagnosing faults or failures.
However, the problem of robust FDII for nonlinear systems has not been investigated as extensively as its linear counterpart. In particular, very few works have been reported in the literature on robust fault isolation and severity identification – rather than just detection – for nonlinear systems. Some examples of robust fault detection and isolation (but not identification) techniques for uncertain nonlinear systems can be found in the works of Xiaodong et al. [119], Talebi and Khorasani [120], Chen and Saif [87], and Wu and Saif [124].

In this monograph, we address the robustness of FDII with respect to external disturbances and particularly measurement noise. Robustness of FDII to measurement noise is of utmost importance especially in applications with low SNR. Robustness in the analytical redundancy-based framework to FDII is, in general, achieved by either making the residual generation process or the residual evaluation process insensitive to uncertainties. In this monograph, we adopt the former approach by reconfiguring the architecture of the proposed FDII scheme from series-parallel into parallel. The robustness of the parallel FDII scheme will be further explained in Chapter 3 and demonstrated in Chapter 5.

### 2.6 Conclusions

In this chapter, the problem of fault detection, isolation, and identification (FDII) in nonlinear systems was defined and formulated. Potential sources of faults in an open-loop system were also introduced including actuator, sensor, and component faults, and some common types of faults were identified. Simple mathematical models of common types of sensor and actuator faults were also presented.

Based upon the formal definition of the FDII problem, various analytical redundancy-based fault diagnosis approaches and methodologies in the literature were reviewed. Based on the a priori source of information on the system being used for diagnostic purposes, these approaches were divided into two categories, namely model-based and computational intelligence (CI)-based. While the model-based approaches exploit the mathematical model of the system for FDII design, the CI-based approaches use quantitative data or qualitative information (i.e., if-then rules) or a combination of both.

The literature on model-based approaches to fault diagnosis was reviewed separately for the three tasks of detection, isolation and identification. The reason for this individual investigation was the different levels of complexity associated with each task and the varying number of contributions within each domain. Specifically, model-based fault-isolation methods were very comprehensively reviewed and analyzed in terms of the concepts behind each method. Furthermore, some examples of FDI techniques developed based on those concepts were also mentioned and analyzed. Even though being less investigated and researched in the literature, some recent efforts in model-based fault identification or severity estimation were also reviewed.
For the CI-based diagnostic approaches, the literature survey was not separated based on the specific task in the FDII problem, since such a distinction can hardly be made within the CI-based fault diagnosis domain. Instead, some extensively used concepts and schemes to achieve fault diagnosis without having a mathematical model of the system were introduced. More precisely, methods that use quantitative data of the system for residual generation based on the learning capability of neural networks were extensively reviewed. Diagnostic methodologies that use qualitative information of the system (mainly in the form of *if-then* rules) were also explored, which are mostly based on fuzzy-logic theory. Furthermore, a general scheme for integrating both quantitative data and qualitative information of the system for fault diagnosis purposes was proposed. It was shown that this integrated scheme consists of various techniques from CI domain such as neural networks, fuzzy systems, and neuro-fuzzy systems.

Eventually, the proposed approach to fault diagnosis in this monograph was reviewed, which is essentially a *hybrid* approach to FDII. It is called *hybrid* in the sense that both a priori mathematical model information of the system and the adaptive nonlinear function approximation capability of neural networks are simultaneously used to accomplish FDII. It was mentioned that the *hybrid* approach to fault diagnosis is relatively new to the research community and actually few works have been reported in the literature following this approach (which were also reviewed in this chapter). However, it is certainly a very promising approach and sounds to be the inevitable choice of future in the fault diagnosis domain.
Fault Diagnosis of Nonlinear Systems Using a Hybrid Approach
Sobhani-Tehrani, E.; Khorasani, K.
2009, XX, 268 p. 155 illus., Softcover
ISBN: 978-0-387-92906-4