Chapter 6
Model-Based Diagnosis and Prognosis of Hybrid Dynamical Systems with Dynamically Updated Parameters

Om Prakash and A.K. Samantaray

6.1 Introduction

Fault detection and isolation (FDI) and prognosis for large complex process engineering systems are important research areas of industrial importance in order to improve the safety, reliability and availability of critical machineries/processes. In condition-based maintenance (CBM), FDI is intended for prompt detection, isolation and classification of any fault in a system and to quantify the severity of fault; whereas prognosis is intended to predict the remaining useful life (RUL) of faulty component or subsystem based on the current health status of the system and its past degradation profile or trend provided by diagnosis. Precise prediction of RUL assists the plant technicians to plan the future maintenance activities. Since diagnosis and prognosis both are concerned with the health monitoring of the industrial system, subsystems or components; it is logical to integrate them in a common framework for process supervision. There are generally two types of fault situations, namely anticipated or unanticipated types. An anticipated fault situation is generally known in advance based upon the history of system behaviour and past experience; but, the unanticipated or unexpected fault situation is generally not known in advance and that must be detected during process monitoring to maintain the safety and reliability of the system. Nowadays, many modern integrated systems or processes such as chemical plants, automobiles and airplanes use embedded system architecture where electronics and communication systems play important roles. These systems contain various dynamical components or subsystems which exhibit both continuous and discrete dynamics and are hence called hybrid systems. In a hybrid dynamical system, faulty discrete events may occur in addition to parametric faults and occurrence of these may be unknown in advance. Most of the

O. Prakash • A.K. Samantaray (✉)
Systems, Dynamics and Control Laboratory, Department of Mechanical Engineering,
Indian Institute of Technology-Kharagpur, 721302 Kharagpur, India
e-mail: omimech21@gmail.com; samantaray@mech.iitkgp.ernet.in; samantaray@lycos.com

© Springer International Publishing Switzerland 2017

W. Borutzky (ed.), *Bond Graphs for Modelling, Control and Fault Diagnosis of Engineering Systems*, DOI 10.1007/978-3-319-47434-2_6

model-based diagnosis and prognosis methods available in literature are intended for continuous systems and these approaches cannot be easily applied to hybrid dynamical systems as the supervision of such systems need tracking of continuous as well as discrete state variables.

Generally two types of fault sources may occur in a hybrid dynamical system. The first one is a parametric fault related to some component degradation and the second one is due to some unexpected transition of nominal mode of the system, i.e., discrete fault (e.g., valve stuck on or stuck off fault, controller transition command failure, etc.). Hybrid dynamics contain two types of discrete transitions: supervisory controlled discrete transition and autonomous mode discrete transition. Usually, the supervisory controlled discrete input to the plant is implemented in software and it is possible to measure the discrete input signals issued by the controller at the interface between the supervisory controller and the plant. So, we assume that the supervisory controlled discrete input signals to the plant are directly observed and known to us. Another way to determine controlled mode transitions is to use the model of the supervisory controller to predict such transitions. Even when the supervisory controlled mode change information is known to us; a discrete fault may be possible like valve stuck on or stuck off fault and pump stuck on or pump stuck off fault, etc., and such faults should also be detected and isolated. In contrast to supervised controlled discrete transition, autonomous mode discrete transitions are usually not known and may not be directly measurable. However, the conditions for autonomous mode transitions are known either in terms of measured plant output variables or in terms of state variables. So, the autonomous mode transitions can be known to diagnosis module based on the measurement of the outputs.

Most of the existing diagnosis and prognosis approaches are based on single fault hypothesis. These assume that the system or subsystem is immediately repaired once a fault is detected and isolated. However, a simple fault can lead to a sequence of other catastrophic faults and it may not be possible to repair each fault within available time. Moreover, some faults may be tolerated and the process operation may be continued in the presence of one or more known faults and RUL of such faults should be known to plant technician so that maintenance activities can be scheduled accordingly. Detection of mode transition and any subsequent fault (which may be serious) after a few known minor faults should also be possible. FDI method based on single fault hypothesis fails to predict the actual fault candidates when next fault occurs after the first fault because effects of one fault may be concealed or compensated by the effects of another fault. One solution to this problem is to use a lot of sensors to decouple fault effects. However, this is a costly approach and each process variable may not be measurable. Other approaches rely on building observers or a bank of observers (including unknown input observers) of the system and tapping measurements/inconsistencies from the observer. However, observer-based approaches cannot be easily applied to hybrid and often non-linear dynamical systems.

For better planning and scheduling of maintenance activities, a good supervision system should detect and isolate small faults and should predict the RUL of faulty/degraded components. Without isolation of correct faults and their types,

RUL estimation is not possible. Hence, it is needed to develop a method having ability to correctly detect and isolate the actual fault of unknown type and unknown degradation behaviour and at the same time, it should provide some information about the severity of the fault and predict the RUL of the faulty components keeping the system level performance constraints in the view.

In a hybrid dynamical system, various components or subsystems operate in different modes or environmental conditions. This results in varying degradation rate of the components throughout the system's life cycle. In fact, the prognosis of hybrid systems is challenging due to the fact that the same component can exhibit different degradation behaviours in different operating modes. Traditionally, RUL estimation is performed by utilizing a single degradation model which assumes that degradation rate parameters used in the degradation model are constants. Although a single degradation model may be enough for a particular degradation pattern, it does not suffice when components or subsystems have different operational profiles. Utilization of multiple degradation models which include operational modes as additional control parameter and evolve through degradation model identification is suggested in this chapter. Note that identifying appropriate multiple degradation models is a challenging task when components have dynamic degradation patterns.

This chapter precisely deals with the afore-mentioned problem faced in diagnosis and prognosis of hybrid systems. The premise of the solution proposed in this chapter is a basic assumption that there is a very rare chance of occurrence of simultaneous faults. Even apparent simultaneous faults are separated by a small time interval and we assume that time interval is large enough to carry out the necessary parametric fault or mode identification, degradation pattern identification and model updating steps. We assume that infinite mode transitions in a finite time do not occur and only partial parametric faults occur in a system. For RUL estimation, it is also assumed that mode of operation of each hybrid component is known in advance.

Different approaches for model-based diagnosis and prognosis have been developed depending on the kind of knowledge used to describe the process model. Usually, a specific methodology is applied for a specific process. Diagnosis methods may be broadly classified into two types: model-based methods and data-driven-based methods. Likewise, prognosis methods intended for RUL estimation can also be classified into three types: model-based prognosis, data-driven prognosis and experience or probability-based prognosis [1]. Every method has its own advantages and disadvantages. In the present chapter, bond graph model-based diagnosis and prognosis (MBDP) scheme is proposed.

For model-based process supervision, a precise and reliable mathematical model of the actual plant behaviour is required. A unified multi-energy domain Bond Graph (BG) [2, 3] and its extended form Hybrid Bond Graph (HBG) [4–6] are well-suited for modelling of continuous and hybrid dynamical behaviours, respectively. BG tool is also useful in the design and development of model-based FDI for both continuous and hybrid dynamical systems [6, 7]. A considerable amount of literature can be found related to model-based FDI for hybrid dynamical systems [8–11]; but, very few works are reported on integration of both diagnosis and prognosis of hybrid dynamical system in a common framework. Also, very few literatures are

available for prognosis of hybrid dynamical systems. In model-based diagnosis, the BG model is used to derive a set of consistency rules called analytical redundancy relations (ARRs). ARRs are constraints expressed in terms of measurable process variables and nominal parameters of plant [7]. These constraints remain valid until a system operates according to its normal operation model. A fault is detected by monitoring the trend of the residuals. For consistency checking, residual must be tested with predetermined fixed threshold value or an adaptive threshold value [6–8]. Diagnostic Bond Graph (DBG) method has been introduced in [12], so that residuals are directly obtained by making use of the current and few past measured sample data even when ARRs cannot be obtained in explicit symbolic form. DBG has been extended to Diagnostic Hybrid Bond Graph (DHBG) and adapted for FDI of hybrid systems [6, 8]. Few works related to the prognosis, which are based on BG approach, can be found in [13–16]. In [13, 14], it is assumed that degradation models of the faulty components are known beforehand and RUL estimation is performed by incorporating the degradation model into constrained ARRs equation. Further, it is assumed that a single component's parameter value continuously drifts with the evolution of time and all other components behave normally. In [15, 16], a BG framework is utilized for FDI but the RUL estimation is performed by using the Monte Carlo framework (particle filter technique). The existing model-based prognosis methods are intended for continuous systems and those cannot be easily applied to hybrid dynamical systems.

Components may degrade due to both internal stresses (load, torque, speed, etc.) and external stresses (wind, temperature, humidity, etc.). In order to accurately predict the RUL, it is necessary to take into account how and where the components will be used and what will be the mode of operation. According to domain knowledge about the considered system and its components, and known environmental and operational conditions, degradation model of the deteriorating components can be identified by understanding the physics of degradation in a model-based prognosis framework. A degradation model of a component may be obtained by accelerated life tests method and then that degradation model may be used to track the degradation of component once an incipient fault is detected by FDI module; whereupon only the coefficients of model need to be estimated and RUL can be predicted [14]. Thus, identification of precise dynamic degradation model and specification of a well-defined failure threshold for RUL estimation are the main challenges in model-based prognosis of hybrid system. In this regard, this chapter makes the following contributions:

- A unified sequential multiple fault diagnosis and prognosis method based on DHBG approach is developed for hybrid dynamical systems by introducing the concept of model updating after each fault identification. The proposed method is able to diagnose faults whose effect may be masked due to previously existing faults and also predicts the RUL of the faulty component if the isolated fault is of progressive type.
- RUL estimation uses the common framework, i.e., BG modelling approach, that has been used for system modelling, virtual prototyping, fault diagnosis rule

development, and parameter and system identification. Utilization of multiple degradation models for RUL estimation is suggested which include operational modes as additional control parameter and evolve through degradation model identification. Same DHBG model of the system as used in FDI module is used in a modified form to identify the degradation pattern of the components after detection of parametric fault. Models are continually evolved with time by adapting to the new information of the state of degradation of the monitored system to provide accurate RUL with bounded uncertainty value.

- Instantaneous fault sensitivity signature for both parametric and discrete faults is used for minimizing the set of suspected faults (SSF), which also provides the expected directions of corresponding parameter deviations. Accordingly, constrained parameter estimation is proposed for improving the diagnosis and prognosis tasks of hybrid system.
- The proposed approach can detect and isolate both parametric and discrete faults and can diagnose different types of fault like abrupt, incipient and progressive faults. However, the method may also detect and isolate the intermittent faults if there is a sufficient time window for parameter estimation. The proposed approach can also track the discrete mode transitions even in presence of one or more faults in a system.

6.2 Model-Based Diagnosis and Prognosis for Hybrid Systems

The performance of model-based diagnosis and prognosis (MBDP) approaches depend on the accuracy or quality of the model of the considered system. Models serve as knowledge representation of a large amount of structural, functional and behavioural information and their relationship. This knowledge representation is capitalized to create complex cause-effect reasoning leading to construction of powerful and robust automatic process supervision tools. The model development for a large complex system is a challenging task, especially for a hybrid dynamical system whose dynamical behaviour changes with the change of operating mode of the system. Assumptions taken during modelling, exclusion of minor dynamics and inclusion of major dynamics and the used modelling technique always affect the accuracy or quality of diagnosis and prognosis outcomes. If an appropriate model of a system which provides the expected behaviour of the real system in normal healthy condition is developed, then it can be utilized for the process supervision of the system. However, there is always some mismatch between the outputs of the behavioural model with the real system measurements even if there is no fault present in the real system. This generally happens because of modelling uncertainties, parameter uncertainties, unknown disturbances, measurement uncertainties, etc. For robust supervision, these factors must be taken into account. BG modelling is a good approach to deal with a multi-energy domain system

which may have continuous and discrete dynamical behaviours with different uncertainties. Also, BG modelling can be used as a common framework for system modelling, virtual prototyping, fault diagnosis rule development, parameter and system identification, and RUL estimation.

Generally, BG model-based diagnosis is broadly classified into qualitative and quantitative approaches [7]. In qualitative model-based method, model considers the cause and effect relationships, and the functional relationships between the outputs and inputs of a system are represented in terms of qualitative functions. In quantitative model-based method, dynamic behaviour of a system is generally obtained from the first principles and the functional relationships between the outputs and inputs of a system are represented in terms of mathematical equations. However, for MBDP, quantitative method is preferred as it provides the common framework for both diagnosis and prognosis module development. Further, this quantitative method may be classified as observer-based, parity relation-based, parameter estimation-based and ARRs-based [8]. These methods can be used to generate the residuals which are the primary step in the process supervision of a system. Among these methods, ARR-based methods are more popular for the development of process supervision. ARRs are constraints expressed in terms of measurable process variables and nominal parameters of plant [7]. These constraints remain valid until a system operates according to its normal operation model. A fault is detected by monitoring the trend of the residuals. Quantitative ARRs-based methods can be further classified as symbolic and numerical methods. In symbolic methods, symbolic ARRs are obtained from BG model to evaluate the residuals; whereas, in numerical methods, residuals can be numerically evaluated using the DBG/DHBG model approach.

Quantitative BG model-based fault diagnosis method consists of two main steps: generation of residual or a DBG/DHBG model and evaluation of residual. The generation of residual is a technique for constructing ARRs using the BG model of the system. In the residual evaluation step, the trends of the residuals are interpreted to check any inconsistency. The inconsistency which indicates presence of one or more faults is detected by testing whether each residual is enveloped by a prescribed adaptive threshold, which in turn is defined based on the known uncertainties in parameter estimation, operating mode information and statistical parameters of measurement noise and unknown disturbances. Once any inconsistency in the residuals is found, the fault identification module is triggered to determine the severity of the fault and its nature/type. After information of severity of fault is obtained, the system may be reconfigured or fault may be accommodated. In case of incipient fault, RUL must be estimated to assists the plant technicians for planning the future maintenance activities. In the following sections, basic concepts in model-based diagnosis and prognosis are introduced.

6.2.1 Basic Framework of Fault Diagnosis

The global ARR, GARR($\boldsymbol{MD}$,$\boldsymbol{\theta}$, $\boldsymbol{U}$, $\boldsymbol{Y}$), of an uncertain hybrid system may be written as

$$\text{GARR}_n \pm \lambda = 0 \tag{6.1}$$

where $\boldsymbol{MD}$ represents the controlled junction mode vector, $\boldsymbol{\theta}$ represents a known parameter vector, $\boldsymbol{U}$ represents known control or input vector and $\boldsymbol{Y}$ represents sensor output vector. Online evaluation of each nominal part, GARR_n, and uncertain part, λ, using $\boldsymbol{MD}$, $\boldsymbol{\theta}$, $\boldsymbol{U}$ and $\boldsymbol{Y}$ provides residual r and adaptive threshold $\varepsilon = |\lambda|$, respectively. The usual approach to detect fault induced inconsistencies is to test whether the numerical value of GARR_n at corresponding mode remains bounded between $\pm\varepsilon$. Residual (r) and adaptive threshold (ε) may be evaluated directly from DBG [12] or DHBG model in linear fractional transformation (LFT) form [6–8]. A binary coherence vector ($\boldsymbol{C}$) is used to represent the signature for a fault; whose standard form is $\boldsymbol{C} = [c_1, c_2, \ldots, c_n]$ where c_i $(i = 1, 2, \ldots, n)$ are obtained from a decision procedure, Θ, which is used to generate the alarms, i.e., $\boldsymbol{C} = [\Theta(r_1), \Theta(r_2), \ldots, \Theta(r_n)]$. For robust FDI, each residual r_i(t) is checked against the time varying adaptive threshold ε_i(t) as follows:

$$c_i = \Theta\,(r_i) = \begin{cases} 0, & \text{if} - \varepsilon_i(t) \le r_i(t) \le \varepsilon_i(t) \\ 1, & \text{otherwise} \end{cases} \tag{6.2}$$

During online monitoring, the coherence vector ($\boldsymbol{C}$) is obtained at each and every sampled time for consistency checking. An alarm is raised if one or more than one elements of the coherence vector show nonzero value, i.e., $\boldsymbol{C} \neq [0, 0 \ldots 0]$. After detection of fault, the coherence vector is matched with the fault signature matrix (FSM) at corresponding mode for isolation of actual fault candidate [7]. A fault in a component is detectable/monitorable, if at least one of the residual is sensitive to this fault, i.e., its monitorability index represented by $M_b = 1$. A fault in a component can be isolated only when it is monitorable and its fault signature (a corresponding row in FSM) is unique; which is represented by isolability index $I_b = 1$.

6.2.1.1 FSM, GFSM, GFSSM and MCSM, MCSSM

An FSM, $\boldsymbol{S}$, represents the relation between a set of parametric faults and their assumed signatures. It is used to detect and isolate the actual faults. The elements of FSM are either 1 or 0 as determined from

$$S_{ji} = \begin{cases} 1, & \text{if } r_i \text{ is a function of } P_j \\ 0, & \text{otherwise} \end{cases} \tag{6.3}$$

where r_i is the residual of the ith column, P_j is the parameter of the jth row in FSM, and $i = (1, 2 \ldots n)$, $j = (1, 2 \ldots p)$, n is number of residuals and p is number of parameters.

Hybrid system dynamics contains both continuous and discrete modes. So, FSMs for such system are mode dependent and need to be separately derived for each mode. Global fault signature matrix (GFSM), ***GS***, is derived in a global form [6, 8], whose elements are obtained from

$$GS_{ji} = \begin{cases} f(a_1, \ldots, a_m), & \text{if } r_i \text{ is a function of } P_j \text{ depending on the values } a_1, \ldots, a_m, \\ 1, & \text{if } r_i \text{ is a function of } P_j \text{ at all modes,} \\ 0, & \text{otherwise} \end{cases} \tag{6.4}$$

where f represents a logical function of controlled junction in a bond graph model and $a_1, \ldots, a_m$, represent the controlled junction state variables, and m is the number of controlled junctions.

Global fault sensitivity signature matrix (GFSSM) is an extension of GFSM, which has capability to differentiate between increasing ($P_j \uparrow$) and decreasing ($P_j \downarrow$) parametric fault. Its elements are updated at each and every instant by using the instantaneous sign of each residual sensitivity with respect to the component parameters [9]. The elements of the GFSSM, ***GSS***, are determined from

$$GSS_{ji} = \begin{cases} -\text{sign}(\partial r_i / \partial P_j), & \text{if } r_i \text{ is a function of } P_j \text{ and } P_j \text{ is expected to increase due to fault,} \\ +\text{sign}(\partial r_i / \partial P_j), & \text{if } r_i \text{ is a function of } P_j \text{ and } P_j \text{ is expected to decrease due to fault,} \\ 0, & \text{otherwise} \end{cases} \tag{6.5}$$

This new kind of signature is termed here sensitivity signature as it can be analysed by sensitivity theory and this residual sensitivity can be derived numerically by using Sensitivity Bond Graph (SBG) approach as in [17–19], where GARRs in closed symbolic form may or may not be derivable.

Residuals are also sensitive to discrete mode fault in a hybrid system and any inconsistency in actual mode may be identified by using mode change signature matrix (MCSM). The elements of the MCSM are determined from

$$\text{MCSM}_{ki} = \begin{cases} 1, \text{ if } r_i \text{ is a function of } a_k \\ 0, \text{ otherwise} \end{cases} \tag{6.6}$$

Mode change sensitivity signature matrix (MCSSM) is an extension of MCSM with capability to differentiate between of an increasing ($a_k \uparrow$) and decreasing ($a_k \downarrow$) mode fault. Its elements are updated at each and every instant by using the instantaneous sign of each residual sensitivity with respect to the mode [9]. The elements of the MCSSM are determined from

$$\mathrm{MCSSM}_{ki} = \begin{cases} -\mathrm{sign}(\partial r_i/\partial a_k), & \text{if } r_i \text{ is a function of } a_k \text{ and } a_k \text{ is expected to change from 0 to 1,} \\ +\mathrm{sign}(\partial r_i/\partial a_k), & \text{if } r_i \text{ is a function of } a_k \text{ and } a_k \text{ is expected to change from 1 to 0,} \\ 0, & \text{otherwise} \end{cases} \tag{6.7}$$

where a_k represents the controlled junction state variables of the kth row of MCSM or MCSSM, and $a_k \in (0, 1)$, $k \in (1, 2, \ldots m)$.

6.2.1.2 Adaptive Thresholds for Robust FDI

Adaptive thresholds are used to achieve robustness in FDI by accounting for the process and measurement uncertainties and also the mode transitions so that supervision system can minimize false alarms and misdetections. An uncertainty on a particular parameter value θ_j can be introduced as

$$\theta_j = \theta_{jn}(1 + \delta_{\theta_j}) \tag{6.8}$$

$$\text{or } \theta_j = \theta_{jn} + \Delta\theta_j \tag{6.9}$$

where $\theta_j \in (I, C, R, TF, GY)$ corresponds to parameters associated with the model, and $\delta_{\theta_j} = (\Delta\theta_j/\theta_{jn})$ and $\Delta\theta_j$ are the relative and the absolute deviations of nominal parameter value θ_{jn}.

Adaptive thresholds using BG-LFT method [7, 8] can be used, in which system uncertainties in parameters are detached from their nominal parameters model and modelled as feedback loops of internal variables. For instance, when the real parameter value of a capacitance C is not accurately identified, it can be expressed as $C_n \pm \Delta C = C_n(1 \pm \delta_C)$, where C_n is represented as nominal parameter value and $\pm\Delta C = \pm\delta_C C_n$ is the uncertainty part of the parameter. If the C element is modelled in derivative causality, then its constitutive relation is given as

$$f = \frac{1}{C_n \pm \Delta C}\dot{e} = \frac{1}{C_n}\left(1 \mp \delta_{1/C}\right)\dot{e} = \frac{\dot{e}}{C_n} \mp w_{1/C} \tag{6.10}$$

where $(\mp\delta_{1/C}/C_n)\dot{e} = \mp w_{1/C}$ is the extra contribution of flow because of uncertain part of parameter and may be treated as a disturbance. Note that $\delta_{1/C}$ is the uncertainty in estimating the value of $1/C$. The C element in derivative causality and parameter uncertainty can be modelled in BG-LFT form as given in Fig. 6.1.

Similarly, for the non-linear resistive R element modelled in conductive causality for representing the flow through a non-linear valve, the constitutive relation including uncertainty is written as

$$\begin{aligned} f &= \frac{1}{R_n \pm \Delta R}\sqrt{e} = \frac{1}{R_n}\left(1 \mp \delta_{1/R}\right)\sqrt{e} \\ &= (C_d \mp \Delta C_d)\sqrt{e} = C_d(1 \mp \delta_{1/R})\sqrt{e} = C_d\sqrt{e} \mp w_{1/R} \end{aligned} \tag{6.11}$$

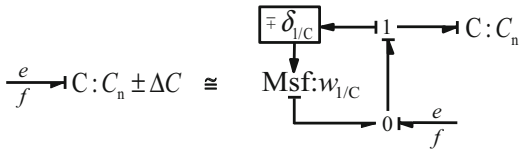

Fig. 6.1 Capacitance (C) element in derivative causality modeled in LFT form

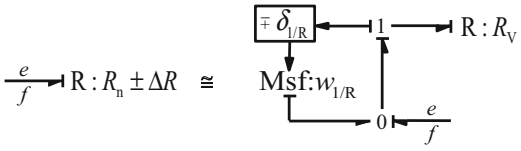

Fig. 6.2 Resistive (R) element in conductive causality modeled in LFT form

where $\delta_{1/R} = \Delta C_d/C_d$, e is the difference of pressure across the non-linear valve, C_d is the nominal value of discharge coefficient of valve, ΔC_d is the uncertainty part of the parameter C_d and $\mp w_{1/R} = \mp \delta_{1/R}\ C_d\sqrt{e}$ is the extra contribution of flow because of uncertain part of parameter and may be treated as a disturbance. The R element in conductive causality and parameter uncertainty can be modelled in BG-LFT form as shown in Fig. 6.2. Likewise, other parameters (I-element, and TF and GY two-ports) with uncertainties can be modelled [20].

6.2.2 Basic Framework of Prognosis

The term prognosis is often used in medical domain to describe the prediction of poor health of a patient by considering the actual diagnosis of one or more symptoms and their evolution compared with other similar observed cases. In industrial domain, the same reasoning of prognosis can be transposed to machines and components to answer the question about the RUL of a machine or a component once an impending failure condition is detected, isolated and identified by diagnosis module.

RUL, also called time to failure (TTF), is the time left before observing a failure of a component or subsystem given the current health status of system and its past degradation profile. Once the degradation trend of component's parameter is obtained, then that can be extrapolated with some set value of failure threshold to predict the RUL of faulty component [6].

$$\mathrm{RUL}(t, z) = t_{\mathrm{fl}} - t_0 |\ t_{\mathrm{fl}} > t_0,\ D(t) \tag{6.12}$$

where t_{fl} indicates the random variable of TTF, t_0 represents the current age of the component or system, $D(t)$ represents the past degradation profile up to the current time and z is the operating mode of the system.

Component's degradation may be captured by a continuous drift of its parameter value in its life cycle as reported in most of the existing literature on prognosis. However, in a hybrid dynamical system, various components or subsystems operate

at different modes or environmental conditions which result in varying degradation rate of the components throughout the system's life cycle. Utilization of multiple degradation models which include operational modes as additional control parameter and evolve through degradation model identification is suggested in this chapter. Models are continually evolved with time by adapting to the new information of the state of degradation of the monitored system to provide accurate RUL with bounded uncertainty value. This overcomes the drawback of other similar models, where the parameters of the model are estimated only once and then the estimated degradation pattern is kept fixed irrespective of subsequent new available observations. RUL estimation and scheduling maintenance activities of component/subsystem based on single pre-identified degraded state are not an optimal solution, especially for hybrid systems that operate under variable modes. Intelligent prognosis must adapt according to change of the state of degradation of the constantly monitored system.

The proposed approach for RUL prediction requires simultaneous monitoring of both degradation and operating modes of the system. Let $\theta_j(t, z) \in \boldsymbol{\theta}$ be the parameter associated with jth component of a system which has started to degrade detectably at time t_0, as determined (detected and isolated) by FDI module. Figure 6.3 shows the different known operating modes (z) of the system and the jth component's (θ_j) degradation behaviour evolution in response to the operating mode changes. It is shown that the change points, denoted by $t^{(i)}$, ($i = 1, 2, \ldots$)

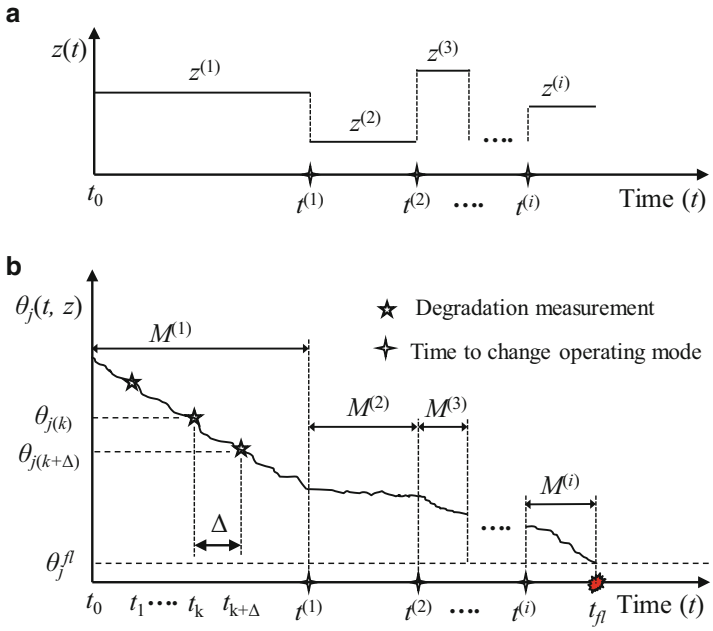

Fig. 6.3 (**a**) Different operating mode (**b**) degradation behaviour evolution at various operating mode of the component (θ_j)

of the operating mode are known and the operating mode is shifted to new mode $z^{(i)} \in z$ after $t^{(i)}$. After the shift, the state of degradation of the component changes at new mode ($z^{(i)}$) and degradation model is identified as $M^{(i)}$. Finally, the degradation curve of the component (θ_j) crosses a predefined failure threshold θ_j^{fl} at time t_{fl}. Thus the mode dependent RUL of the component (θ_j) as per defined performance (θ_j^{fl}) is estimated as

$$\text{RUL}(t, z) = t_{\text{fl}} - t_0 \tag{6.13}$$

Such a case of transition of operating condition from one mode to another is very common in hybrid dynamical systems. For instance, performance of an on–off valve which is generally used in processing plants may degrade because of fouling, which decreases the coefficient of discharge (C_d) of water flow. But the coefficient of discharge (C_d) does not continuously decrease at all times because of discrete operating nature of valve. When the valve is in open condition, fluid flow through the unit results in increased fouling such as due to sediment/lime scale deposition; but when the valve is in closed condition there is no such fouling. So, this operating condition pattern should be considered for RUL prediction.

6.2.2.1 An Integrated Framework of Diagnosis and Prognosis for Multiple Faults in Hybrid Dynamical System

Since diagnosis and prognosis both are concerned with the health monitoring of the industrial system, subsystems or components; it is reasonable to integrate them in a common framework for process supervision. Moreover, hybrid system contains both discrete and continuous dynamics; thus, discrete mode faults may occur in addition to parametric fault (abrupt or progressive type) and the occurrence of these faults are generally unknown in advance. In case of abrupt parametric fault or discrete fault, there is a step-like deviation in the corresponding component's parameter and it generally persists with the evolution of time. However in case of incipient or progressive fault, there is a slow change in the component's parameter with some dynamic degradation pattern which may be unknown beforehand. In abrupt or discrete fault, it is essential that the diagnosis scheme detects the faults quickly to avoid catastrophic consequences. In such cases, prompt fault detection and fault accommodation are the main aim of fault diagnosis. On the other hand, incipient faults are more significant in maintenance activities where it is necessary that slowly evolving faults are detected early enough to avoid more severe consequences. Once the nature of degradation pattern of incipient fault is obtained, RUL can be predicted by using the degradation model. In a large complex hybrid system, the occurrence of sequential multiple faults are much more likely, while the occurrence of simultaneous faults may be very rare and this is taken as a key assumption in the developments presented in this chapter. In this section, an integrated approach to BG-MBDP in a hybrid dynamical system for sequential multiple fault of unknown nature or type is proposed. The main goal of this section is to show how the same

DHBG model that is used for FDI of hybrid system can be further used to identify the degradation pattern of the faulty components and to estimate their RUL.

For estimation of component's degradation pattern in hybrid dynamical system, the fault magnitude of degrading parameter $\theta_j(t, z)$ is estimated with a fixed window of sample data collected at different time instants at various operating modes (z). As a result, a set of estimates of parameter values of degrading component is obtained at different time instances at various operating modes. Then degradation model, $M_{\theta_j}^{(i)}(z)$, which is best fit equation of parameter value evolution at any operating mode, is identified through curve fitting tool. Consequently, the obtained degradation models corresponding to various operating modes can be used for RUL estimation based on well-defined failure thresholds and known future operating modes of the component.

A complete flow chart of the proposed integrated MBDP for hybrid dynamical system is represented in Fig. 6.4. In a hybrid dynamical system, GARRs are used to detect any inconsistency in the nominal behaviour of the monitoring system at any mode. The general form of GARR which is obtained from DHBG model can be written as

$$\mathrm{GARR}_i = \mathrm{GARR}_i(\boldsymbol{MD}, \boldsymbol{\theta}, \boldsymbol{U}, \boldsymbol{Y}) \tag{6.14}$$

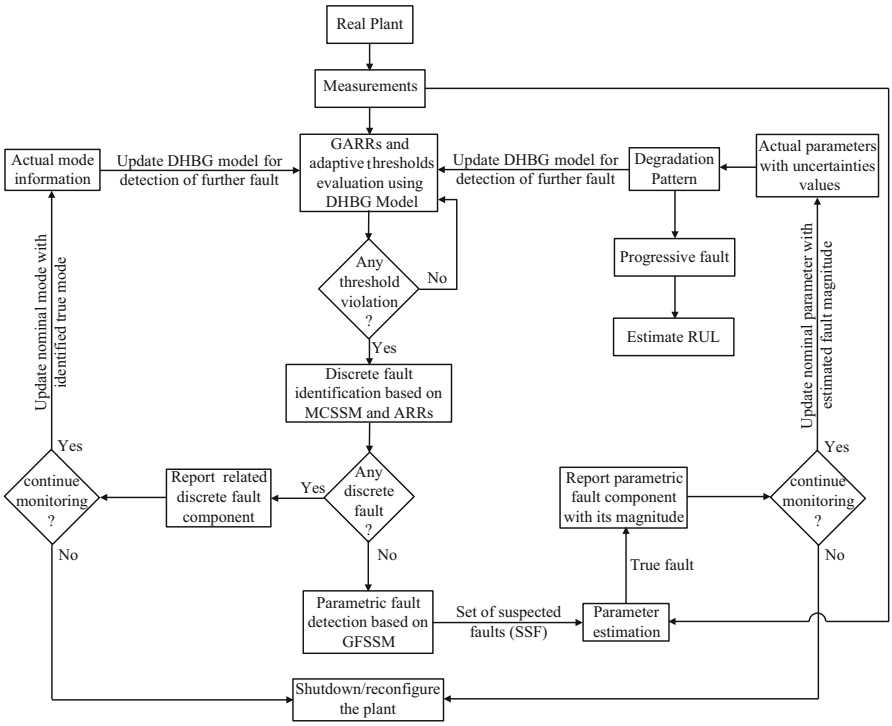

Fig. 6.4 Flow chart of proposed integrated diagnosis and prognosis method

where $\boldsymbol{MD} = [a_1, a_2, \ldots, a_k, \ldots, a_m]^T$ represents the controlled junction mode vector, $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_j, \ldots, \theta_p]^T$ represents a known parameter vector which includes p number of component's nominal parameters, **U** represents known input vector and is formed by $\boldsymbol{U} \in (\boldsymbol{Sf}, \boldsymbol{Msf}, \boldsymbol{Se}, \boldsymbol{Mse})$ where ***Sf*** or ***Msf*** represents flow or modulated flow input vector and ***Se*** or ***Mse*** represents effort or modulated effort input vector, ***Y*** denotes sensor output vector and is formed by $\boldsymbol{Y} \in (\boldsymbol{Df}, \boldsymbol{De})$ where ***Df*** and ***De*** represent flow and effort sensor output vector, respectively, and $i \in (1, 2, \ldots n)$ is the number of residuals of the system.

Online evaluation of $\text{GARR}_i = \text{GARR}_i(\boldsymbol{MD}, \boldsymbol{\theta}, \boldsymbol{U}, \boldsymbol{Y})$ using ***MD, θ***, ***U*** and ***Y*** in normal operation may be written as

$$\text{GARR}_{ni} \pm \lambda_i = 0 \tag{6.15}$$

where GARR_{ni} is nominal part and λ_i is the uncertain part of GARR_i, respectively, whose evaluations provide nominal residual r_{ni} and adaptive threshold $\varepsilon_i = |\lambda_i|$, respectively:

$$r_{ni}(t) = \text{Eval}\{\text{GARR}_{ni}(\boldsymbol{MD}, \boldsymbol{\theta}, \boldsymbol{U}, \boldsymbol{Y})\} \text{ and } \varepsilon_i(t) = \pm\text{Eval}\{\lambda_i\} \tag{6.16}$$

During normal operation of the system, $\varepsilon_i(t) \geq r_i(t) \geq -\varepsilon_i(t)$ is satisfied. The residuals (r_{ni}) are sensitive to both parametric and discrete faults. Some of the residuals which are sensitive to a particular discrete or parametric fault in a system cross either upper or lower threshold when any type of the fault occurs. If any one of the nominal parameter (say θ_j) of θ changes (more than uncertainty value) or any one of the nominal mode of ***MD*** (say a_k) is inconsistent, then only a set of particular residuals which are sensitive to the change of this parameter θ_j or a_k cross the threshold in due time. When any threshold violation occurs, we initially hypothesize that this inconsistency is due to a discrete fault. If the sensitivity signature obtained corresponding to threshold violations ($+1$ for crossing upper threshold, -1 for crossing lower threshold and 0 for lying within thresholds) has unique match in MCSSM, then the component related to a_k is declared as a faulty one. If this discrete fault (a_k) cannot be isolated because more than one components share the same signature in MCSSM, then the ARR-based mode tracking [6] is followed in which all inconsistent ARRs are evaluated with each hypothesized mode fault and the actual faulty mode, if any, is identified as the one that gives consistent residuals. Then, the current discrete fault information of a_k is fed into the DHBG model to update the GARRs and adaptive thresholds for the isolation of subsequent faults. Once discrete fault is identified, we assume that it persists indefinitely thereafter if the plant is not allowed to shut down. The ARR-based mode tracking can also be used to track the initial mode of the system if it is unknown [6].

If the inconsistency in the residuals is not due to a particular discrete fault, i.e., the initial hypothesis could not be validated, then it is hypothesized that the inconsistency is due to a parametric fault only. If the sensitivity signature obtained corresponding to threshold violations has unique match in GFSSM corresponding

to component (θ_j), then the FDI module detects the degradation of component (θ_j) (say at time t_0) and component (θ_j) is isolated as degrading component at mode $z = z^{(i)}$ (say). On the other hand, if the degrading component (θ_j) cannot be isolated because of more than one components share the same signature in GFSSM, then a set of suspected fault (SSF) is hypothesized for further course of action.

After postulation of faults candidates in SSF, a targeted parameter estimation technique is triggered for identification of true fault and its degradation pattern. This module tries to quickly estimate only a few possible parameters from which the true fault candidate (θ_j) can be isolated [7, 8, 17, 18, 21]. Now the parameter θ_j is finally declared as a degrading faulty parameter and is represented by $\theta_j(t, z^{(i)})$ at mode $z^{(i)}$. The estimated fault magnitude of degrading parameter $\theta_j(t, z^{(i)})$ at kth instant of time is represented by $\theta_j^f(k, z^{(i)})$. If the nominal part of each GARR_{ni} is evaluated again with the estimated fault magnitude $\theta_j^f(k, z^{(i)})$ (real value of degrading parameter of plant/system at that instance), and adaptive thresholds are also updated, then the corresponding evaluated residuals do not cross the corresponding residual thresholds [8]. So, the original vector $\boldsymbol{\theta}$ is updated by replacing the nominal θ_j by the estimated fault magnitude, $\theta_j^f(k, z^{(i)})$ of the real plant. Now updated $\boldsymbol{\theta}$ at kth instant is assumed to be the new nominal parameter vector that is used to update DHBG model in LFT form to predict further degradation pattern of faulty component $\theta_j(t, z^{(i)})$. Note that a large deviation in a parameter value is considered as an abrupt fault. Likewise, any discrete (mode) fault is also treated as abrupt fault.

If the detected parametric fault (θ_j) is of progressive type, then the component's parameter $\theta_j(t, z^{(i)})$ is degrading slowly according to operating mode $z^{(i)}$. Thus, the same set of residuals which are sensitive to the change of this parameter $\theta_j(t, z^{(i)})$ would cross the adaptive threshold again after some more time. Again, parameter (θ_j) is estimated and its estimated value at $(k+\Delta)$th instant of time is represented by $\theta_j^f(k + \Delta, z^{(i)})$. Since up to the current time only two data points of fault magnitude of degrading parameter are known, a linear degradation model may be assumed for the initial estimation of RUL which alerts the maintenance engineer for scheduling the maintenance activities or other tasks. This initial linear degradation model is further adapted with modified model when more parameter estimates are obtained during monitoring. For accurate estimation of degradation pattern, sufficient number of data points should be obtained at particular mode $z^{(i)}$ and then the degradation model is identified by using the curve fitting tools at corresponding mode $z^{(i)}$. If the jth operating change point occurs at time instant $t^{(j)}$ corresponding to new mode $z^{(j)}$ during data collection for estimating the parameter value in previous operating mode $z^{(i)}$, then the fresh data of a fixed window size corresponding to new mode $z^{(j)}$ are collected and fault magnitude is estimated in this new mode $z^{(j)}$. This process is repeated until the true degradation model, $M_{\theta_j}^{(i)}(\text{z})$, is obtained for parameter $\theta_j(t, z)$ at different operating modes. Then the obtained true model $M_{\theta_j}^{(i)}(\text{z})$ is further used in prognosis module to predict RUL of faulty component (θ_j) with the future known operating mode by extrapolating the model $M_{\theta_j}^{(i)}(\text{z})$. Also, the finally obtained true

degradation model $M_{\theta_j}^{(i)}(z)$ of the component (θ_j) is fed into the DHBG model, so that the subsequent faults can be detected and isolated. In the case of sequential multiple progressive faults, RUL is estimated for each progressive fault component and the RUL of the component which has least value is the significant for the maintenance engineer.

6.3 Application to a Two-Tank Benchmark System

6.3.1 Description of a Two-Tank Hybrid System

The benchmark hybrid two-tank system is adapted from [11]. Its process and instrumentation diagram and hybrid bond graph (HBG) model are shown in Figs. 6.5 and 6.6, respectively. This system consists of two tanks (T_1 and T_2) that are connected with pipes and valves (V_1 and V_2). C_{dvi} is the coefficient of discharge of non-linear valve V_i, $i = 1, 2$. The liquid level in tank T_1 is regulated by a hydraulic pump which is controlled by a PI-controller installed in this system and tries to maintain water level of 0.5 m. The flow of pump (Q_P) is proportional to the output of the PI-controller (U_{PI}). Two drainage pipes (L_1 and L_2) with coefficients of discharge C_{dL1} and C_{dL2}, respectively, showing linear behaviour are also used. This system shows the hybrid dynamics and includes both autonomous and supervisory controller transition modes. Valve V_1 is switched to on and then off state according to command input given by the supervisory controller. When water level in tank T_1 exceeds level H_{L1} then the water starts flowing from tank T_1 to tank T_2 through drainage pipe L_1 (autonomous mode, a_1). Similarly, when water level in tank T_2 exceeds level H_{L2} then the water starts flowing from tank T_2 to atmosphere through

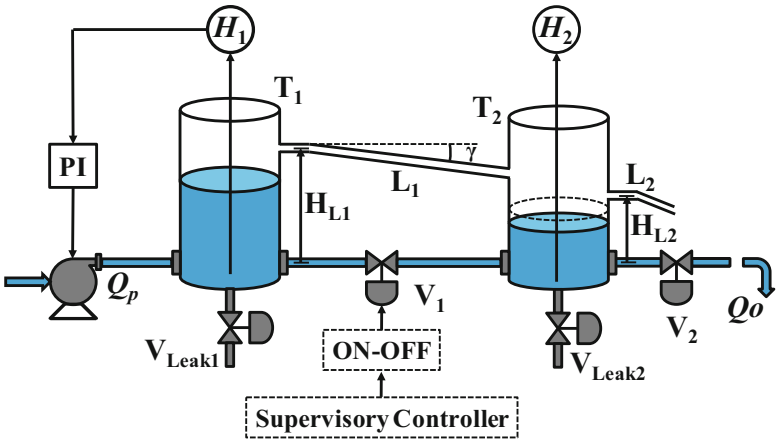

Fig. 6.5 Schematic diagram of a two-tank hybrid system

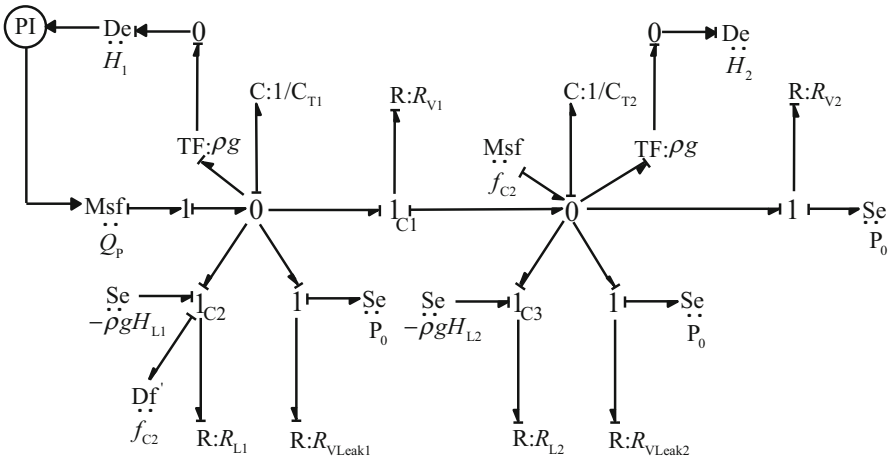

Fig. 6.6 HBG model of a two-tank hybrid system

drainage pipe L_2 (autonomous mode, a_2). Leakage fault can be introduced in tank T_i using imaginary valve V_{Leaki} having coefficient of discharge C_{dLeaki}, $i = 1, 2$.

Two level sensors H_1 and H_2 and one flow sensor Q_P are installed in the system for measuring the water levels in tanks T_1 and T_2 and water input flow by pump, respectively. The atmospheric pressure is assumed to be the reference pressure. The small angle γ, volume of water in the drainage pipe and inertia effect of water flow are neglected in this system; with the former two being considered as part of the uncertainties in tank capacities. This way of simplified model building by neglecting minor dynamics improves the speed of diagnosis without unduly complicating the process of development of the supervision system.

Pump saturation characteristic (Q_P) and PI-controller output law (Φ_{PI}) are, respectively, given as

$$Q_P = \begin{cases} U_{PI}, & 0 \le U_{PI} \le f_{max} \\ 0, & U_{PI} \le 0 \\ f_{max}, & U_{PI} \ge f_{max} \end{cases} = \Phi_P(U_{PI}) \tag{6.17}$$

$$\begin{aligned} U_{PI} &= K_P(S_{pt} - \rho \cdot g \cdot H_1(t)) + K_I \int (S_{pt} - \rho \cdot g \cdot H_1(t))dt \\ &= \Phi_{PI}(H_1(t)) \end{aligned} \tag{6.18}$$

where f_{max} is the maximum flow from the pump, S_{pt} is a pressure (or level) set point and K_P and K_I are the proportional and integral gains, respectively. In this work, we do not consider the actuator, controller and sensor faults; they may be diagnosed with additional hardware/sensor redundancies and hence are not relevant for this study.

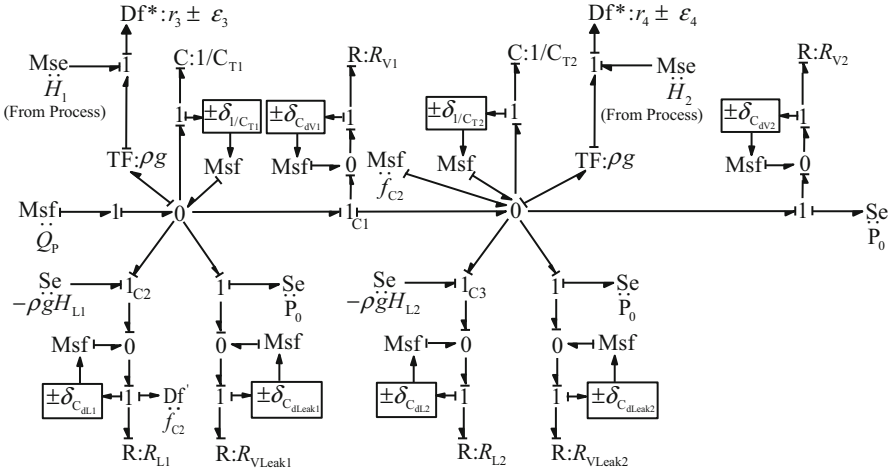

Fig. 6.7 DHBG-LFT model of a two-tank hybrid system

6.3.2 DHBG Model and GARRs Generation

Diagnostic hybrid bond graph (DHBG) model is obtained from HBG model of a system, in which all controlled junctions and storage elements are assigned with appropriate causalities so that all active BG elements remain active at all operating modes. This permits consistent causal description and generation of GARRs for a hybrid dynamical system from its DHBG model. The DHBG of the considered two-tank hybrid system in LFT form is shown in Fig. 6.7. Two virtual flow sensors (Df^*) are used to derive two constraints GARR_3 and GARR_4.

The ARRs for the actuators and the controllers (which will not be used in this study) are simply obtained from comparisons of input and output relationships as

$$\text{ARR}_1 : Q_\text{P} - \Phi_\text{P}(U_\text{PI}) = 0 \tag{6.19}$$

$$\text{ARR}_2 : U_\text{PI} - \Phi_\text{PI}(H_1(t)) = 0 \tag{6.20}$$

$$\begin{aligned}\text{GARR}_3 : {} & Q_\text{P} - C_\text{T1} \cdot \frac{d}{dt}(\rho \cdot g \cdot H_1(t)) - a_\text{v1} \cdot C_\text{dv1} \cdot \sqrt{|\rho \cdot g \cdot (H_1(t) - H_2(t))|} \\ & \cdot \text{sign}(\rho \cdot g \cdot (H_1(t) - H_2(t))) - a_1 \cdot C_\text{dL1} \cdot \rho \cdot g \cdot (H_1(t) - H_\text{L1}) \\ & - C_\text{dLeak1} \cdot \sqrt{|\rho \cdot g \cdot H_1(t)|} \pm \lambda_3 = 0\end{aligned} \tag{6.21}$$

$$\begin{aligned}\text{GARR}_4 : {} & a_\text{v1} \cdot C_\text{dv1} \cdot \sqrt{|\rho \cdot g \cdot (H_1(t) - H_2(t))|} \cdot \text{sign}(\rho \cdot g \cdot (H_1(t) - H_2(t)) \\ & + a_1 \cdot C_\text{dL1} \cdot \rho \cdot g \cdot (H_1(t) - H_\text{L1}) - a_2 \cdot C_\text{dL2} \cdot \rho \cdot g \cdot (H_2(t) - H_\text{L2}) \\ & - C_\text{T2} \cdot \frac{d}{dt}(\rho \cdot g \cdot H_2(t)) - C_\text{dv2} \cdot \sqrt{|\rho \cdot g \cdot H_2(t)|} - C_\text{dLeak2} \cdot \\ & \sqrt{|\rho \cdot g \cdot H_2(t)|} \pm \lambda_4 = 0\end{aligned} \tag{6.22}$$

where $C_{T1} = A_1/g$, $C_{T2} = A_2/g$, $a_1 = \begin{cases} 0, & H_1(t) \le H_{L1} \\ 1, & H_1(t) > H_{L1} \end{cases}$ and $a_2 = \begin{cases} 0, & H_2(t) \le H_{L2} \\ 1, & H_2(t) > H_{L2} \end{cases}$

$GARR_3$ and $GARR_4$, as presented in (6.21) and (6.22), respectively, contain the uncertain parts λ_3 and λ_4. The effects of uncertainties in various parameters on a GARR are un-correlated with possibility of cancelling out each other. Hence, absolute values of the individual effects are considered for adaptive threshold evaluation as follows:

$$\begin{aligned}\lambda_3 = {} & |\delta_{C1} \cdot C_{T1} \cdot \frac{d}{dt}(\rho \cdot g \cdot H_1(t))| + |a_{v1} \cdot \delta_{C_{dv1}} \cdot C_{dv1} \cdot \sqrt{|\rho \cdot g \cdot (H_1(t) - H_2(t))|}| \\ & + |a_1 \cdot \delta_{C_{dL1}} \cdot C_{dL1} \cdot \rho \cdot g \cdot (H_1(t) - H_{L1})| + |\delta_{C_{dLeak1}} \cdot C_{dLeak1} \cdot \sqrt{\rho \cdot g \cdot H_1(t)}|\end{aligned} \tag{6.23}$$

$$\begin{aligned}\lambda_4 = {} & |\delta_{C2} \cdot C_{T2} \cdot \frac{d}{dt}(\rho \cdot g \cdot H_2(t))| + |a_{v1} \cdot \delta_{C_{dv1}} \cdot C_{dv1} \cdot \sqrt{|\rho \cdot g \cdot (H_1(t) - H_2(t))|}| \\ & + |a_1 \cdot \delta_{C_{dL1}} \cdot C_{dL1} \cdot \rho \cdot g \cdot (H_1(t) - H_{L1})| + |a_2 \cdot \delta_{C_{dL2}} \cdot C_{dL2} \cdot \rho \cdot g \cdot (H_2(t) - H_{L2})| \\ & + |\delta_{C_{dv2}} \cdot C_{dv2} \cdot \sqrt{\rho \cdot g \cdot H_2(t)}| + |\delta_{C_{dLeak2}} \cdot C_{dLeak2} \cdot \sqrt{\rho \cdot g \cdot H_2(t)}|\end{aligned} \tag{6.24}$$

Using (6.4) and (6.6) on (6.21) and (6.22), the GFSM and MCSM for two-tank hybrid system are obtained as shown in Tables 6.1 and 6.2, respectively; whereas using (6.5) and (6.7) on (6.21) and (6.22), the GFSSM and MCSSM are found as shown in Tables 6.3 and 6.4, respectively. The parameters related to leakage fault in tank T_1 and T_2, i.e., C_{dLeak1} and C_{dLeak2}, respectively, have only increasing possibility (i.e., leakage), while for other parameters (i.e., C_{dv1}, C_{dv2}, C_{dL1}, C_{dL2}), both increasing (i.e., leakage) and decreasing faults (i.e., blockage) are possible. Likewise, discrete stuck on and off faults for a_{v1} of valve V_1 are also possible.

Table 6.1 GFSM (***GS***) for the two-tank hybrid system

Parameter	$GARR_3$ (r_3)	$GARR_4$ (r_4)	M_b	I_b Single fault
C_{dv1}	a_{v1}	a_{v1}	1	$\bar{a}_{v1} = (1 - a_1)a_{v1}$
C_{dv2}	0	1	1	0
C_{dL1}	a_1	a_1	a_1	$\bar{a}_1 = (1 - a_{v1})a_1$
C_{dL2}	0	a_2	a_2	0
C_{dLeak1}	1	0	1	1
C_{dLeak2}	0	1	1	0

Table 6.2 MCSM for the two-tank hybrid system

Parameter	$GARR_3$ (r_3)	$GARR_4$ (r_4)	M_b	I_b Single fault
a_{v1}	1	1	1	1

Table 6.3 GFSSM (***GSS***) for the two-tank hybrid system

Parameter	GARR_3 (r_3)	GARR_4 (r_4)
$C_{\text{dv1}} \uparrow$	$a_{\text{v1}}\text{sign}(H_1(t) - H_2(t))$	$-a_{\text{v1}}\text{sign}(H_1(t) - H_2(t))$
$C_{\text{dv1}} \downarrow$	$-a_{\text{v1}}\text{sign}(H_1(t) - H_2(t))$	$a_{\text{v1}}\text{sign}(H_1(t) - H_2(t))$
$C_{\text{dv2}} \uparrow$	0	$+1$
$C_{\text{dv2}} \downarrow$	0	-1
$C_{\text{dL1}} \uparrow$	a_1	$-a_1$
$C_{\text{dL1}} \downarrow$	$-a_1$	a_1
$C_{\text{dL2}} \uparrow$	0	a_2
$C_{\text{dL2}} \downarrow$	0	$-a_2$
$C_{\text{dLeak1}} \uparrow$	$+1$	0
$C_{\text{dLeak2}} \downarrow$	0	$+1$

Table 6.4 MCSSM for the two-tank hybrid system

Parameter	GARR_3 (r_3)	GARR_4 (r_4)
$a_{\text{v1}} \uparrow$	$\text{sign}(H_1(t) - H_2(t))$	$-\text{sign}(H_1(t) - H_2(t))$
$a_{\text{v1}} \downarrow$	$-\text{sign}(H_1(t) - H_2(t))$	$\text{sign}(H_1(t) - H_2(t))$

These possibilities, called technological specifications which are derived from deep knowledge of the system, are considered in Tables 6.3 and 6.4, respectively.

Note that the sensitivity signatures presented in GFSSM and MCSSM [9] are simplified expressions derived from GARRs. These can be numerically obtained from SBG model if GARRs cannot be symbolically derived. For example, if we consider the sensitivity signature element $GSS_{1,1}^{C_{\text{dv1}}\uparrow}$ due to the fault $C_{\text{dv1}} \uparrow$, then the corresponding element in the GFSSM is calculated as

$$\begin{aligned} GSS_{1,1}^{C_{\text{dv1}}\uparrow} &= -\text{sign}(\partial \text{GARR}_3 / \partial C_{\text{dv1}}) \\ &= -\text{sign}\big(- a_{\text{v1}} C_{\text{dv1}} \sqrt{|\rho g(H_1(t) - H_2(t))|}\ \big) \\ &= a_{\text{v1}}\text{sign}(H_1(t) - H_2(t)) \end{aligned} \tag{6.25}$$

6.3.3 Simulation Study and Results

In this section, an integrated MBDP method for sequential multiple faults of unknown nature/type in a hybrid dynamical system is tested through simulation. Also, model-based process supervision scheme using the most recent existing approach [9] which considers single fault hypothesis and the newly proposed approach which considers sequential multiple faults hypotheses, along with comparison between these two approaches are presented. This section also shows that the methods which are based on single fault hypothesis fail to predict the actual fault candidates in case of sequential multiple faults and without correct isolation of fault

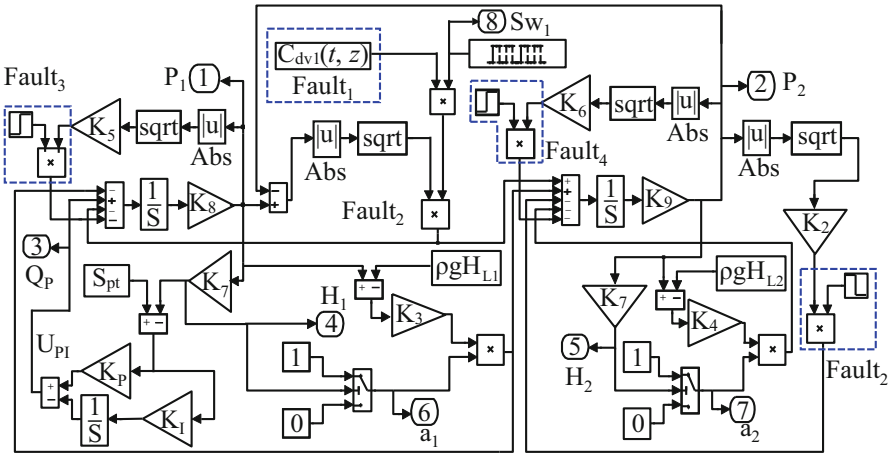

Fig. 6.8 Simulink model of two-tank system with provisions to introduce faults

candidates, prognosis is irrelevant. It is also shown that the newly proposed method gives improved fault isolation capabilities in both single and multiple fault scenarios and also improves the RUL estimation if the identified fault is of progressive type.

The HBG model of the two-tank hybrid system (shown in Fig. 6.6) is converted into MATLAB-SIMULINK model (with provisions of faults introduction in model as shown in Fig. 6.8) to generate the process data through simulation for the validation of the proposed method. In Fig. 6.8, $C_{\text{dv1}}(t, z)$ block represents the progressive fault function block for valve V_1 and $K_2 = C_{\text{dv2}}, K_3 = C_{\text{dL1}}, K_4 = C_{\text{dL2}}$, $K_5 = C_{\text{dLeak1}}$, $K_6 = C_{\text{dLeak2}}$, $K_7 = 1/\rho g$, $K_8 = g/A_1$, $K_9 = g/A_2$ are the corresponding gains blocks. Similarly, the DHBG-LFT model (shown in Fig. 6.7) can be converted into MATLAB-SIMULINK model for residuals and adaptive thresholds evaluation [11]. Here, the residuals and thresholds evaluation are done in a simple MATLAB program that evaluates, respectively, the nominal GARR_3 and GARR_4 in (6.21) and (6.22), and the uncertain parts λ_3 and λ_4 in (6.23) and (6.23). These, correspondingly, provide the residuals r_3 and r_4, and adaptive thresholds $\varepsilon_3 = \pm\lambda_3$ and $\varepsilon_4 = \pm\lambda_4$. Sensor's measurements data, parameter's nominal values and parameter's uncertainty values are the inputs to the program [7, 11]. The overall threshold may be evaluated to account for the sensor's noise as, $\varepsilon_i = \pm(\lambda_i + k_i)$, $i = 3, 4$, where k_i is the static threshold which is chosen based on the sensor characteristics or with the model of additive sensor biases presented in [8]. In the simulations, we have assumed $k_i = 0$.

The two-tank hybrid system is simulated for duration of 1800 s using a fixed step size of 0.02 s by initializing all state variables to zero. The nominal value of system's parameters used in the model are given in Table 6.5. Two types of faults are introduced in the simulation. The first one is a progressive fault in valve V_1 (see Table 6.6). The parameter C_{dv1} of on–off valve V_1 (nominal value = $1.593 \times 10^{-2}\,\text{kg}^{1/2}\,\text{m}^{1/2}$ in on state) is considered to drift slowly with time according to

Table 6.5 Nominal parameters of the hybrid two-tank

Symbol	Description	Nominal value
K_P	Proportional gain of controller	1 ms
K_I	Integral gain of controller	5×10^{-2} m
S_{pt}	Set point of the PI-controller	0.5 m
f_{max}	Maximum outflow from pump	1 kg/s
A_i	Cross-sectional area of tank $T_i (i = 1, 2)$	1.472×10^{-2} m^2
C_{dvi}	Discharge coefficient of valve V_i including connected pipe $(i = 1, 2)$	1.593×10^{-2} kg$^{1/2}$ m$^{1/2}$
C_{dLi}	Discharge coefficient of drainage pipe L_i $(i = 1, 2)$	1×10^{-3} ms
C_{dLeaki}	Discharge coefficient of $V_{Leaki} (i = 1, 2)$	0 kg$^{1/2}$ m$^{1/2}$
H_{L1}	Height of the drainage pipe L_1 of tank T_1 from datum	0.58 m
H_{L2}	Height of the drainage pipe L_2 of tank T_2 from datum	0.40 m
P_0	Atmospheric pressure	0 N/m^2
ρ	Density of water	1000 kg/m^3
g	Acceleration due to gravity	9.81 m/s^2

Table 6.6 Simulated faults in the model

Parameter	Description	Degradation nature	Start time, t_{fi} (s)	End time
C_{dv1}	Valve blockage	Progressive type as per (6.26)	225	1800
C_{dv2}	Valve blockage	Abrupt type as per (6.27)	1475	1800

operating mode ($z = z^{(i)} = a_{v1}$) as shown in Fig. 6.9b. Second one is an abrupt fault in valve V_2. In the simulation, we have introduced the fault in parameters C_{dv1} and C_{dv2} at a time instants $t_{f1} = 225$ s and $t_{f2} = 1475$ s, respectively, as defined through the functions given in (6.26) and (6.27), respectively.

$$C_{dv1}(t, z) = \begin{cases} C_{dv1n}(z) \cdot a_{v1}, & t < t_{f1} \\ C_{dv1n}(z) \left(e^{-r(z) \cdot t_{on}}\right) \cdot a_{v1}, & t \geq t_{f1} \end{cases} \tag{6.26}$$

$$C_{dv2}(t, z) = \begin{cases} C_{dv2n}(z), & t < t_{f2} \\ 0.9 C_{dv2n}(z), & t \geq t_{f2} \end{cases} \tag{6.27}$$

where $C_{dvin}(z)$ is the nominal parameter value of valve $V_i (i = 1, 2)$ at corresponding operating mode (z), $r(z) = 1.0 \times 10^{-4}$ s^{-1} at $z = z^{(1)} = a_{v1} = 1$, with each $z^{(1)}$ for 80 s and $r(z) = 0$ s^{-1} at $z = z^{(2)} = a_{v1} = 0$, with each $z^{(2)}$ for 30 s, $t_{on} = \int_{t_{f1}}^{t} a_{v1} \cdot dt$, and t_{f1} and t_{f2} are the time instances when the first and second faults start, respectively.

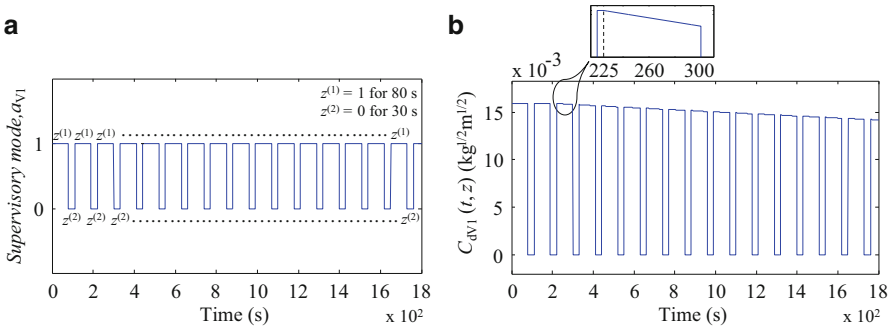

Fig. 6.9 (**a**) Operating mode (**b**) introduced degradation of on–off valve V_1

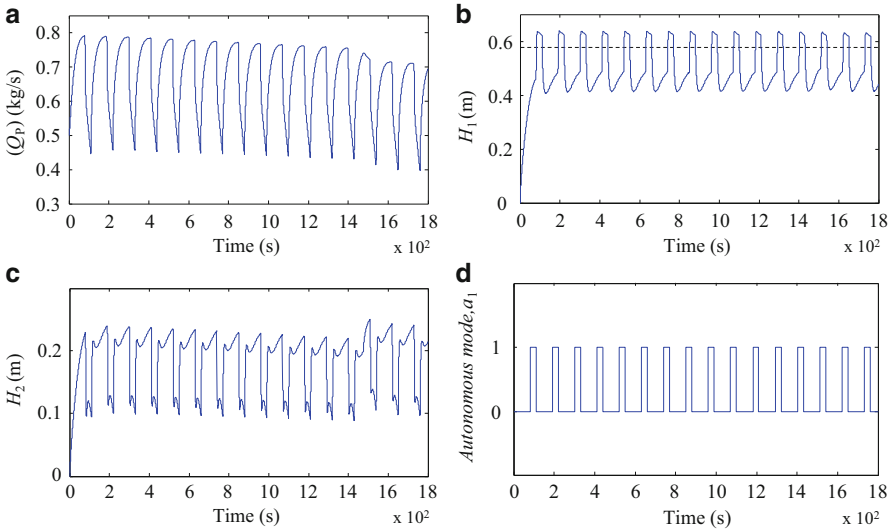

Fig. 6.10 Time responses of measurements (**a**) Pump flow (**b**) Water level in tank T_1 (**c**) Water level in tank T_2 and (**d**) Autonomous mode of drainage pipe L_1

Note that parameter C_{dv1} is considered to drift slowly only in on state of the valve, while there is no drift of parameter C_{dv1} during off state as the flow is zero in that state. The drifting of parameter C_{dv1} just after $t_{fl} = 225$ s is clearly shown in enlarged view of a portion of Fig. 6.9b. The failure threshold (C^{fl}_{dv1}) is considered as half of the nominal value in on state.

The measurements (i.e., Q_P, H_1 and H_2) from the simulated model at a sampling rate of 0.02 s are fed to the residuals and thresholds evaluation program. The time responses of measured input (Q_P), outputs (H_1, H_2) and obtained autonomous mode (a_1) are shown in Fig. 6.10. The transition of mode a_1 from 0 to 1 subjected to given conditions (when $H_1(t)$ exceeds $H_{L1} = 0.58$ m) corresponding to drainage pipe L_1 is clearly noticeable. However, no autonomous mode change is found for drainage pipe L_2; hence $a_2 = 0$ throughout the simulation period.

Table 6.7 GFSSM (***GSS***) for the two-tank hybrid system

Parameter	GARR$_3$ (r_3)	GARR$_4$ (r_4)	M_b	I_b Single fault
$C_{dv1} \uparrow$	a_{v1}	$-a_{v1}$	a_{v1}	$\bar{a}_{v1} = (1 - a_1)a_{v1}$
$C_{dv1} \downarrow$	$-a_{v1}$	a_{v1}	a_{v1}	$\bar{a}_{v1} = (1 - a_1)a_{v1}$
$C_{dv2} \uparrow$	0	$+1$	1	0
$C_{dv2} \downarrow$	0	-1	1	$\bar{a}_2 = (1 - a_2)$
$C_{dL1} \uparrow$	a_1	$-a_1$	a_1	$\bar{a}_1 = (1 - a_{v1})a_1$
$C_{dL1} \downarrow$	$-a_1$	a_1	a_1	$\bar{a}_1 = (1 - a_{v1})a_1$
$C_{dL2} \uparrow$	0	a_2	a_2	0
$C_{dL2} \downarrow$	0	$-a_2$	a_2	0
$C_{dLeak1} \uparrow$	$+1$	0	1	1
$C_{dLeak2} \uparrow$	0	$+1$	1	0

Table 6.8 MCSSM for the two-tank hybrid system

Parameter	GARR$_3$ (r_3)	GARR$_4$ (r_4)	M_b	I_b Single fault
$a_{v1} \uparrow$	$+1$	-1	1	1
$a_{v1} \downarrow$	-1	$+1$	1	1

6.3.3.1 Implementation of Integrated Diagnosis and Prognosis Approach

In practice, the elements of GFSSM and MCSSM have to be updated at each and every instant by using the instantaneous measurement data. It is observed from Fig. 6.10 that the measurement H_1 is always greater than H_2, so the obtained GFSSM and MCSSM during the duration of observation (0–1800 s) for the considered two-tank hybrid system can be simplified to the forms given in Tables 6.7 and 6.8, respectively. The response of residuals (r_3 and r_4) and adaptive thresholds obtained from DHBG-LFT model using previously existing method [9] and the newly proposed method with dynamically updated parameter and updated adaptive thresholds are shown in Figs. 6.11 and 6.12, respectively (solid lines indicate residuals and dashed/dotted lines indicate adaptive thresholds). The response of residuals (r_3 and r_4) and adaptive thresholds during normal operation (up to $t = 225$ s) and during identification of degradation behaviour of valves V_1 and V_2 using proposed method (after $t = 225$ s) is clearly shown in Fig. 6.12. Note that the first two columns of the coherence vector related to actuator and controller faults, which are irrelevant in this study, have been dropped from the analysis.

A discussion on fault isolation capabilities using existing FDI methods considering single fault hypothesis is presented here. From the simulated faults listed in Table 6.6, both valves V_1 and V_2 are faulty between 1475 and 1800 s as the fault in valve V_1 is not repaired. The observed coherence vector ($\boldsymbol{C}$) just after 225 s (after initiation of blockage fault in V_1) is obtained from Fig. 6.11a, b as $\boldsymbol{C} = [1\ 1]$ (if not considering residual sensitivity signature) or $\boldsymbol{C} = [-1\ +\ 1]$ (if considering the residual sensitivity signature). According to coherence vector $\boldsymbol{C} = [1\ 1]$ or $\boldsymbol{C} = [-1 + 1]$, the set of suspected faults (SSF) is C_{dv1} and a_{v1} using GFSM/GFSSM

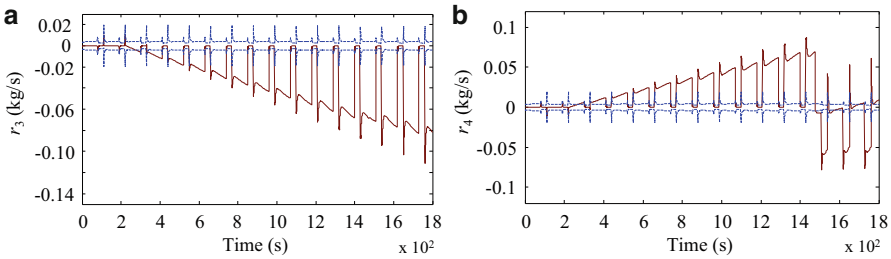

Fig. 6.11 Response of residuals (**a**) r_3 and (**b**) r_4 using previously existing method [9]

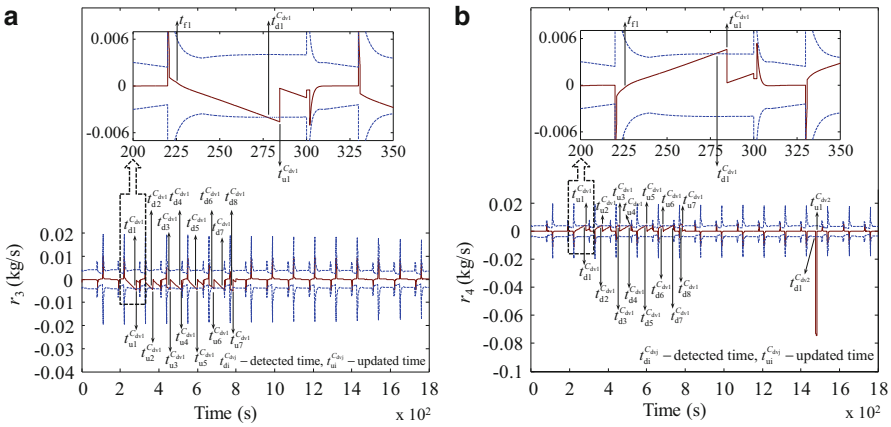

Fig. 6.12 Response of residuals (**a**) r_3 and (**b**) r_4 using proposed method with dynamically updated parameter and adaptive threshold

and MCSM/MCSSM at mode $a_1 = 0$; and if mode $a_1 = 1$, SSF is C_{dv1}, C_{dL1} and a_{v1} (see Tables 6.1, 6.2 and Tables 6.7, 6.8). In both modes, i.e., at mode $a_1 = 1$ and $a_1 = 0$, fault is non-isolatable as signature of C_{dv1}, C_{dL1} and a_{v1} are same. As the dynamics of valve V_1 is mode dependent, the residuals lie within the thresholds even after fault for $a_{\mathrm{v0}} = 0$ (See Fig. 6.11) and under such situation, fault cannot be detected and isolated until the system moves into a different mode ($a_{\mathrm{v1}} = 1$).

When the next fault in valve V_2 is introduced at 1475 s, the coherence vector is observed as $\boldsymbol{C} = [1\ 0]$ (if not considering residual sensitivity signature) and $\boldsymbol{C} = [-1\ 0]$ (if considering the residual sensitivity signature) Fig. 6.11a, b. According to coherence vector $\boldsymbol{C} = [1\ 0]$, possible SSF may be C_{dLeak1} at mode $a_{\mathrm{v1}} = 1$ and this lead to a wrong fault isolation. On the other hand, $\boldsymbol{C} = [-1\ 0]$ does not have any match in GFSSM and MCSSM and the fault is not isolated. While the fault is detected through both approaches (GFSM/MCSM and GFSSM/MCSSM) in this case of sequential multiple faults, the misdiagnosis is natural because the residuals are not diagonal or structured and thus, not suitable for multiple fault diagnosis [7]. Here, one fault effect hides the other fault effect that results in wrong fault signatures.

To identify the actual fault candidates in such cases, it is needed to estimate all the parameters simultaneously by using the non-linear parameter estimation techniques [7, 8, 17, 18, 21] and tracking the real mode using ARR-based mode identification technique [6]. However, this is a very tough and computationally complex task for a large complex system; especially considering the fact that prompt detection and true faults isolation are the main objectives in diagnosis task. Under the assumptions stated beforehand for sequential multiple faults, we need to estimate only a few postulated fault candidates or possible conflicts [17]. The parameter estimation can be done with least squares response matching or minimization of residuals (see [7, 8, 17, 18, 21] for details). Here, we will use least squares response matching approach for parameter estimation. This approach gives best performance when there are few parameters to estimate (as detected from GFSSM) and there are constraints on parameter values (expressed in terms of penalty functions added to the objective function to be minimized). For example, to estimate value of C_{dv1}, the parameter value is constrained between specified minimum (in this case 0) and maximum values. For the considered fault scenario (Table 6.6), GFSSM indicates a possible decrease in value of C_{dv1} and thus its value may be constrained between 0 and the earlier known value (nominal value). This way, the search zone is reduced. In addition, parameter estimation requires initial guess values of parameters. For discharge coefficient of valve V_1, it may be assigned as αC_{dv1}, where $0 < \alpha < 1$. In fact, value of α can be approximately obtained from the rate of change of residuals. The reduced search zone and closer guess values improve the convergence of the optimization process. The estimation procedure requires transient data after the fault. Therefore, a small delay amounting to a chosen window length is present between fault detection and its isolation. We assume that no more faults occur within this chosen window length after detection of a fault event.

A discussion on fault isolation capabilities using the new proposed integrated MBDP method, considering sequential multiple faults hypotheses using GFSSM and MCSSM, is presented here. When the simulated fault (progressive type) for valve V_1 is introduced at 225 s as shown in Fig. 6.9b, the coherence vector (for a short time just after 225 s) is obtained from Fig. 6.12a, b as $\boldsymbol{C} = [-1+1]$. According to coherence vector $\boldsymbol{C} = [-1+1]$, the possible SSF can be $C_{\mathrm{dv1}} \downarrow$ and $a_{\mathrm{v1}} \downarrow$ using GFSSM and MCSSM at mode $a_1 = 0$; otherwise, the possible SSF can be $C_{\mathrm{dv1}} \downarrow$, $C_{\mathrm{dL1}} \downarrow$ and $a_{\mathrm{v1}} \downarrow$ if mode $a_1 = 1$ (see Tables 6.7 and 6.8). In both cases, i.e., at mode $a_1 = 0$ and $a_1 = 1$, fault is non-isolatable as signature of $C_{\mathrm{dv1}} \downarrow$, $C_{\mathrm{dL1}} \downarrow$ and $a_{\mathrm{v1}} \downarrow$ are same. Since we assume the discrete mode fault occurs first, the ARR-based mode identification algorithm [6] is triggered. Whether the discrete mode fault $a_{\mathrm{v1}} \downarrow$ (valve V_1 stuck off) occurs or not is checked by evaluating all sensitive ARRs at current mode information and it is found that mode $a_{\mathrm{v1}} = 1$ is consistent. This indicates that the inconsistency is due to parametric faults and elements of SSF are refined as $C_{\mathrm{dv1}} \downarrow$ only at $a_1 = 0$, otherwise elements of SSF are $C_{\mathrm{dv1}} \downarrow$ and $C_{\mathrm{dL1}} \downarrow$ at $a_1 = 1$. Then the parameter estimation technique is triggered and $C_{\mathrm{dv1}} \downarrow$ is isolated as a true fault in both modes, i.e., $a_1 = 0$ and $a_1 = 1$. This blockage fault in valve $\mathrm{V}_1 (C_{\mathrm{dv1}} \downarrow)$ is detected at time instant $t_{\mathrm{d1}}^{C_{\mathrm{dv1}}} = 277.68$ s at its on state ($z = z^{(1)} = a_{\mathrm{v1}} = 1$ and $a_1 = 0$). There is some time delay between fault detection

and its occurrence because of slow evolution of fault and the uncertainties included in residual threshold evaluation.

Since the direction of parameter variation of C_{dv1} is known by GFSSM (decrease in value of C_{dv1}), this information is used in constrained parameter estimation technique for degradation model identification. This makes degradation model identification fast, which is one of the main requirements in CBM. For the estimation of fault magnitude of C_{dv1}, a small time window of 7 s transient data after the fault has been selected. Now the fault magnitude of degrading parameter C_{dv1} is estimated after detecting the fault at time instant $t_{\mathrm{d1}}^{C_{\mathrm{dv1}}} = 277.68\,\mathrm{s}$. The estimated parameter value is known at time instant $t_{\mathrm{u1}}^{C_{\mathrm{dv1}}} = 277.68\,\mathrm{s} + 7\,\mathrm{s} = 284.68\,\mathrm{s}$ and is represented by $C_{\mathrm{dv1}}^{f}(t_{\mathrm{u1}}^{C_{\mathrm{dv1}}}, z^{(1)})$, whose value is $C_{\mathrm{dv1}}^{f}(t_{\mathrm{u1}}^{C_{\mathrm{dv1}}}, z^{(1)}) = 1.584 \times 10^{-2}\,\mathrm{kg}^{1/2}\,\mathrm{m}^{1/2}$. Now, the DHBG model is updated by replacing the nominal value $C_{\mathrm{dv1}} = 1.593 \times 10^{-2}\,\mathrm{kg}^{1/2}\,\mathrm{m}^{1/2}$ by the estimated fault magnitude $C_{\mathrm{dv1}}^{f}(t_{\mathrm{u1}}^{C_{\mathrm{dv1}}}, z^{(1)}) = 1.584 \times 10^{-2}\,\mathrm{kg}^{1/2}\,\mathrm{m}^{1/2}$ of the real plant. Now the residuals and adaptive thresholds evaluation obtained through updated nominal part of GARRs and updated uncertainties parts obtained from the updated DHBG model in LFT form force the residuals to lie within the updated adaptive thresholds (see Fig. 6.12 where this updated time instant is denoted as $t_{\mathrm{u1}}^{C_{\mathrm{dv1}}}$). Since, the parameter C_{dv1} is degrading progressively at mode $z = z^{(1)} = a_{\mathrm{v1}} = 1$, again a set of residuals (r_3 and r_4) which are sensitive to the change of this parameter, C_{dv1}, cross the adaptive threshold with the evolution of time (marked as $t_{\mathrm{d2}}^{C_{\mathrm{dv1}}} = 364.02\,\mathrm{s}$ in Fig. 6.12) and again the new fault magnitude of degrading parameter C_{dv1} is estimated and known at new time instant $t_{\mathrm{u2}}^{C_{\mathrm{dv1}}} = 364.02\,\mathrm{s} + 7\,\mathrm{s} = 371.02\,\mathrm{s}$ and represented by $C_{\mathrm{dv1}}^{f}(t_{\mathrm{u2}}^{C_{\mathrm{dv1}}}, z^{(1)})$, whose value is $C_{\mathrm{dv1}}^{f}(t_{\mathrm{u2}}^{C_{\mathrm{dv1}}}, z^{(1)}) = 1.575 \times 10^{-2}\,\mathrm{kg}^{1/2}\,\mathrm{m}^{1/2}$. Since only two data points are known up to the current time a linear degradation model is assumed for the initial estimation of RUL to alert the maintenance engineer. The obtained data points at different time instances ($t_{\mathrm{u1}}^{C_{\mathrm{dv1}}}$ and $t_{\mathrm{u2}}^{C_{\mathrm{dv1}}}$) corresponding to mode $z = z^{(1)} = a_{\mathrm{v1}} = 1$ are used to find the linear degradation mathematical model for the parameter C_{dv1} using the curve fitting tool in Matlab-Simulink and is shown in Fig. 6.13a. For curve fitting, t_{on} is used as a time reference in abscissa. Initial estimated value of RUL using the linear degradation model at on state of valve V_1 and no degradation at off state based on defined failure threshold ($C_{\mathrm{dv1}}^{\mathrm{fl}} = 0.5C_{\mathrm{dv1}}$) and known future operating modes of valve V_1 is found as 6822.19 s (see Fig. 6.14a). This initial degradation model is further adapted with modified model when more new information of data points are obtained during monitoring. In Fig. 6.14, $t_{\mathrm{u1}}^{C_{\mathrm{dv1}}}$ is considered as a zero time reference for RUL estimation.

Note that at different time instances, the valve V_1 is switched to off state as per the command input given by the controller. In this duration, there is no flow through the valve V_1. Although valve V_1 is isolated as a faulty element, residuals r_3 and r_4 become zero (see Fig. 6.12) at the corresponding off state ($z = z^{(2)} = a_{\mathrm{v1}} = 0$) which provides $C_{\mathrm{dv1}}(t, z^{(2)}) = 0\,\mathrm{kg}^{1/2}\,\mathrm{m}^{1/2}$.

For accurate estimation of degradation pattern, sufficient data points should be used. As the monitoring is continued, newer information of degradation data is obtained. Estimated magnitude of degrading parameter C_{dv1} at the same mode

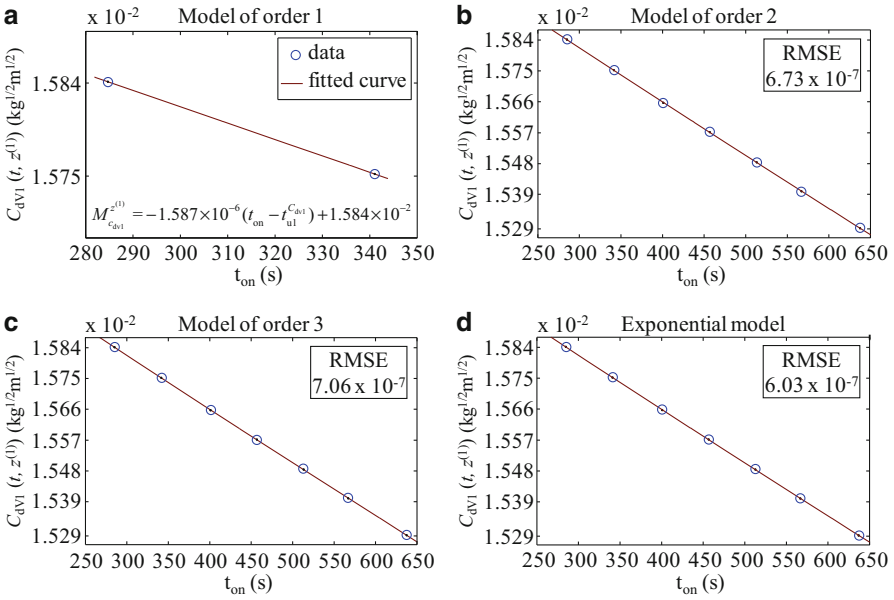

Fig. 6.13 (**a**) Linear degradation model at mode $a_{v1} = 1$ as two data points are known (**b**)–(**d**) Identification of degradation model at mode $a_{v1} = 1$ with new information of data points

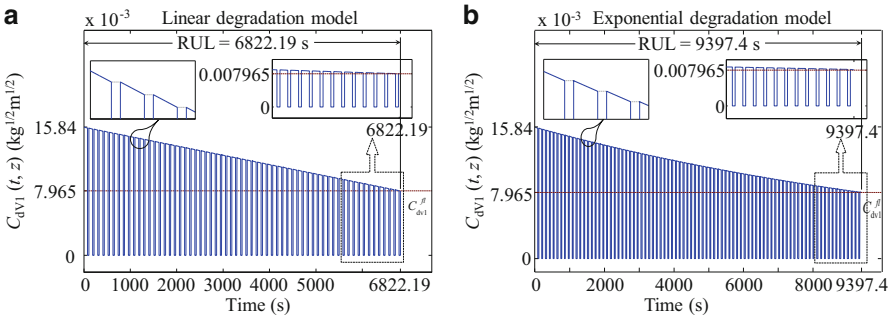

Fig. 6.14 Estimated RUL of parameter C_{dv1} (**a**) Initially assumed linear degradation model when only two data points are known (**b**) Finally confirmed exponential degradation model with the new information of degradation state at on state of the valve V_1 and no degradation at off state

$z = z^{(1)} = a_{v1} = 1$ at different time instances $t^{C_{dv1}}_{ui} (i = 3, 4, \ldots, 7)$ are shown in Table 6.9. The new obtained data points at different time instances with previous known data points are used to refine the degradation behaviour of parameter C_{dv1}. Various degradation models of second order, third order polynomial fit and exponential fit, etc., are tried and shown in Fig. 6.13b–d along with root mean square error (RMSE) values. Exponential fit model is found as a best degradation model

Table 6.9 Estimated data points at different time instances

Parameter	Detected time (s) ($t_{\mathrm{di}}^{C_{\mathrm{dvj}}}$)	Updated time (s) ($t_{\mathrm{ui}}^{C_{\mathrm{dvj}}}$)	t_{on} (s)	Estimated magnitude $C_{\mathrm{dvj}}^{f}(t_{\mathrm{ui}}^{C_{\mathrm{dvj}}}, z)$ ($\mathrm{kg}^{1/2}\,\mathrm{m}^{1/2}$)
$C_{\mathrm{dv1}} \downarrow$	277.68	284.68	284.68	1.584×10^{-2}
	364.02	371.02	341.02	1.575×10^{-2}
	453.56	460.56	400.56	1.566×10^{-2}
	509.50	516.50	456.50	1.557×10^{-2}
	595.78	602.78	512.78	1.548×10^{-2}
	679.76	686.76	566.76	1.539×10^{-2}
	780.04	787.04	637.04	1.529×10^{-2}
$C_{\mathrm{dv2}} \downarrow$	1475.02	1482.02	–	1.434×10^{-2}

at on state of the valve $\mathrm{V}_1(a_{\mathrm{v1}} = 1)$ according to its goodness of fit (RMSE $= 6.03 \times 10^{-7}$) which is the actual degradation pattern as per simulated fault. The obtained exponential model having coefficient a and b with 95 % confidence bound is represented as

$$M_{C_{\mathrm{dv1}}}^{z(1)} = a \cdot \exp(b(t_{\mathrm{on}} - t_{\mathrm{u1}}^{C_{\mathrm{dv1}}})) \tag{6.28}$$

where; $a = 1.583 \times 10^{-2}$ and $b = -1.001 \times 10^{-4}$

Using the above degradation model RUL is estimated as 9397.4 s (see Fig. 6.14b) according to set failure threshold ($C_{\mathrm{dv1}}^{\mathrm{fl}} = 0.5C_{\mathrm{dv1}}$). Note that in the simulation, we have inserted a fast rate of progressive degradation, but in real situation the degradation rate may be very slow.

When the abrupt fault in valve V_2 (blockage) at 1475 s is inserted, the coherence vector ($\boldsymbol{C}$) just after 1475 s, i.e., after blockage fault initiation in V_2, is obtained from Fig. 6.12a, b as $\boldsymbol{C} = [0 - 1]$. This gives the possible SSF as $C_{\mathrm{dv2}} \downarrow$ (i.e., blockage fault in V_2) using GFSSM since $a_2 = 0$ at all times for the considered system. If the mode $a_2 = 1$ for some other system configuration, then the SSF would be obtained from GFSSM as $C_{\mathrm{dv2}} \downarrow$ and $C_{\mathrm{dL2}} \downarrow$; and parameter estimation technique can be used for the suspected parameters C_{dv2} and C_{dL2} to isolate the actual fault ($C_{\mathrm{dv2}} \downarrow$) and its magnitude. In the current configuration, the fault $C_{\mathrm{dv2}} \downarrow$ is directly isolatable at $a_2 = 0$. Still we need to estimate the fault magnitude for updating the DHBG model for diagnosis of subsequent faults. In this case, $C_{\mathrm{dv2}} \downarrow$ is found as actual fault parameter with estimated fault magnitude of $C_{\mathrm{dv2}}^{f} \approx 0.9C_{\mathrm{dv2}}$ and this estimated parameter value is then onwards considered as the new nominal parameter value. Now the residuals and adaptive thresholds evaluation obtained through updated nominal part of GARRs and updated uncertainties parts obtained from the updated DHBG model in LFT form (i.e., C_{dv2} is replaced by C_{dv2}^{f}) force the residuals to lie within the updated adaptive thresholds (see Fig. 6.12b after time $t_{\mathrm{u1}}^{C_{\mathrm{dv2}}} = 1482.02$ s). Now the residuals (r_3 and r_4) remain inside the adaptive threshold until the next fault occurs. This way, subsequent parametric faults and their degradation pattern can be isolated and RUL can be estimated for progressive faults.

6.4 Equivalent Hybrid Electrical System

A great range of equivalent systems can be derived from a given system by using the equivalences of various BG elements and also transformer (*TF*) and gyrator (*GY*) equivalences of BG theory. A linear hydraulic system can be modeled by its analogous linear electrical system because of the similarities in their dynamics or governing differential equations. In the present work, an equivalent electrical system is modelled for the considered two-tank system and is further scaled down for easy and low cost experimental implementation. Note that while the equivalent electrical model is linear, the hybrid nature (mode dependent changes) are retained in it.

6.4.1 Circuit Layout

An equivalent electrical circuit of the two-tank hybrid system is shown in Fig. 6.15, whose HBG and DHBG models are similar to the two-tank hybrid system as shown in Figs. 6.6 and 6.7, respectively. In the electrical domain, power variables current and voltage, respectively, are equivalent to the flow rate and pressure in hydraulic domain. In the circuit, two electrical capacitors C_1 and C_2 are considered instead of two tanks T_1 and T_2, respectively. Resistors R_1, R_2 corresponding to valves V_1, V_2, and R_{d1}, R_{d2} corresponding to drainage pipes L_1, L_2 are considered, respectively. Diodes D_1 and D_2 are considered as switches to permit the current in equivalent drainage pipe resistors R_{d1} and R_{d2} corresponding to set threshold voltages V_{set1} and V_{set2}, respectively. I_{d1} is considered as a drainage current source flowing through resistor R_{d1} and it charges the capacitor C_2 only if the corresponding set condition is reached. A modulated current source I_{in} is used in place of modulated pump

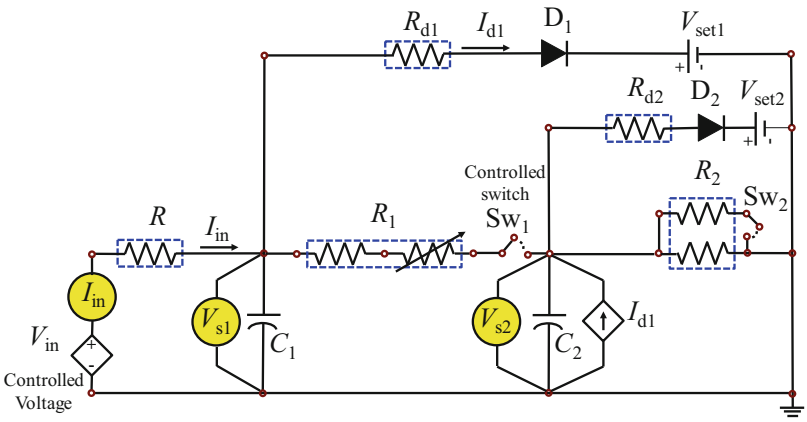

Fig. 6.15 Schematic diagram of equivalent electrical circuit of two-tank

flow Q_P. The modulated current source I_{in} is achieved by controlling the input voltage V_{in} which is the PI-controller's output and by using a known resistor R. The PI-controller, whose output is the modulated voltage V_{in}, is constructed by using various operational amplifiers (op-amps), resistors and capacitor combinations. The objective of PI-controller is to maintain a constant set voltage ($V_{set} = 5\,V$) across the capacitor C_1 which is equivalent to level or pressure set point in tank T_1. A microcontroller (Arduino Uno) is used to open and close the relay switch Sw_1, which sets resistor R_1 at on and off state as per the command given by the controller. Switch Sw_2 is used to introduce the abrupt fault in resistors R_2 by using parallel connection of resistors where due to parallel connection, the resistance is low when the switch is on (nominal state) and high when it is off (simulated faulty state). Also, a series connection of a fixed resistor and a variable resistor is used to introduce the progressive fault in equivalent resistor R_1 at a particular instance according to its operating mode (see Table 6.13 for different time instances of introduced faults and their nature). A controlled servomotor is used to vary the resistance value of resistor R_1 at a particular rate at on state of resistor R_1. For controlling the servomotor, another microcontroller board (Arduino Uno) is used. The flow sensor (Q_P) and level sensors (H_1, H_2) in the hydraulic system are replaced by current sensor (I_{in}) and voltage sensors (V_{s1}, V_{s2}), respectively, in the electrical domain.

The PI-controller law as used for the modulated voltage source, V_{in}, and the modulated current source I_{in} are given as

$$V_{in}(t) = K_P(V_{set} - V_{s1}(t)) + K_I \int (V_{set} - V_{s1}(t))dt \tag{6.29}$$

$$I_{in}(t) = \frac{V_{in}(t) - V_{s1}(t)}{R} \tag{6.30}$$

The values of pump input flow Q_P and the pressures P_1, P_2 of linear hydraulic two-tank system at steady state were measured by doing a simulation in Matlab-Simulink and obtained as follow:

$$\begin{aligned} Q_P &= 0.8\,\text{kg/s} \\ P_1 &= \rho \cdot g \cdot H_1 = 1000 \times 9.81 \times 0.5 = 4905\,\text{N/m}^2 \\ P_2 &= \rho \cdot g \cdot H_2 = 1000 \times 9.81 \times 0.25 = 2452.5\,\text{N/m}^2 \end{aligned} \tag{6.31}$$

A hydraulic-electrical analogy technique is used to determine the nominal parameters of the equivalent electrical circuit. For example, linear resistances R_1 and R_2 corresponding to the two valves V_1 and V_2 are obtained as

$$\begin{aligned} R_1 &= \tfrac{P_1 - P_2}{Q_P} = \tfrac{4905 - 2452.5}{0.8} = 3065.625\,\Omega \\ R_2 &= \tfrac{P_2}{Q_P} = \tfrac{2452.5}{0.8} = 3065.625\,\Omega \end{aligned} \tag{6.32}$$

In the similar manner, the rounded off values of all parameters of the equivalent electrical system are obtained as follows: $R = R_1 = R_2 = 3066\,\Omega$, $R_{d1} = R_{d2} = 1000\,\Omega$, $C_1 = C_2 = 1500\,\mu\text{F}$.

The op-amps, capacitances, resistances and other electronic/electrical components purchased from the market come with uncertainties. Thus, the rounding-off errors may be considered as uncertainties in corresponding parameters.

6.4.2 Model Scaling

For conducting the experiment, the voltage and capacitance values were found to be out of range. For instance, as per (6.31), we need to operate the circuit around 5 kV electrical voltage range. Thus, a dual time and amplitude scaling was performed. For linear systems, the response $x(t)$ of the original system may be scaled as

$$x_s(t_s) = \alpha x(\beta t) \tag{6.33}$$

where α and β are the constants for amplitude and time scaling, respectively. We have considered $\alpha = 1/981$ which means the input current is reduced or equivalently, the voltage set point for PI-controller is set at $4905/981 = 5$ V. The hydraulic system has a slow response. We reduced the response time by $\beta = 10$ times. For that, the time constant ($\tau = RC$) was adjusted by retaining the value of R and reducing the value of C by ten times. The PI-controller's gains are adjusted to match with time constant of the modified system. The final scaled parameter values given in Table 6.10 are considered for developing an experimental setup.

Table 6.10 Nominal parameters of the scaled equivalent electrical circuit

Symbol	Description	Nominal value
K_P	Proportional gain of controller	0.3066
K_I	Integral gain of controller	1
V_{set}	Set point of the PI-controller	5 V
$I_{in(max)}$	Maximum input electrical current	5 mA
C_i	Electrical capacitance value of ith capacitor ($i = 1, 2$)	150 μF
R_i	Electrical resistance corresponding to valve $R_i (i = 1, 2)$	3066 Ω
R_{di}	Electrical resistance corresponding to drainage pipe $R_{di} (i = 1, 2)$	1000 Ω
V_{set1}	Switch threshold voltage for R_{d1}	5.1 V
V_{set2}	Switch threshold voltage for R_{d2}	2.8 V

6.5 Experimental Study

An experimental setup of equivalent electrical hybrid system is shown in Fig. 6.16a. The circuit in the breadboard is the equivalent electrical hybrid system of the simulated two-tank hybrid system in reduced scale whose dynamics is nearly similar to the two-tank hybrid system dynamics. The schematic diagram of equivalent electrical circuit of two-tank hybrid system is already shown in Fig. 6.15. The various components of the circuit including PI-controller and the microcontroller are marked in Fig. 6.16b.

The experimental data were collected at a fixed sampling rate of 0.002 s from the current sensor I_{in} and voltage sensors V_{s1} and V_{s2} using a data-acquisition card (NI-USB6211). For computation of residuals, thresholds, mode identification and parameter estimation, LabVIEW-Matlab interface was used. The experimentally collected data were fed into the DHBG-LFT model of the electrical hybrid system for evaluation of residuals and adaptive thresholds. The GFSSM and MCSSM of the equivalent electrical system presented in Tables 6.11 and 6.12, respectively, are nearly same as that of the hydraulic two-tank system; only the corresponding parameter's nomenclature is changed. The coefficient of discharge parameter $C_{\text{d}i} \downarrow$ is considered equivalent to $1/R_i \downarrow$, where R_i is the linear resistance in electrical domain, i.e., the signature of $C_{\text{d}i} \downarrow$ is replaced by $R_i \uparrow$. Similarly, the signature of $C_{\text{d}i} \uparrow$ is replaced by $R_i \downarrow$.

As in the simulation, two types of faults were introduced in the experimental model. First one is a progressive fault in resistor R_1 as per Eq. (6.34) in its on state. The second one is an abrupt fault in resistor R_1 as per Eq. (6.35) (see Table 6.13). The failure threshold (R_1^{fl}) of resistor R_1 is considered as twice of its nominal value at on state for RUL estimation.

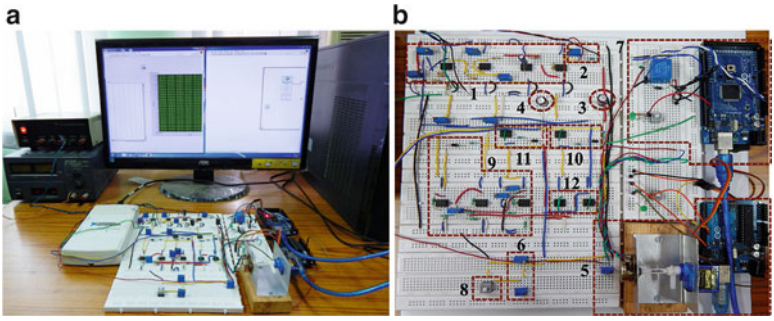

Fig. 6.16 (**a**) An experimental setup of electrical hybrid system (**b**) Enlarge view of breadboard showing different components within *dotted boxes*: (*1*) PI-controller, (*2*) R, (*3*) C_1, (*4*) C_2, (*5*) Variable resistor (R_1) with microcontroller, (*6*) R_2, (*7*) Relay switch (Sw_1) with microcontroller, (*8*) Sw_2, (*9*) I_{d1}, (*10*) D_1, R_{d1} and V_{set1}, (*11*) D_2, R_{d2} and V_{set2}, (*12*) Buffers

Table 6.11 GFSSM (***GSS***) for the electrical hybrid system

Parameter	GARR_3 (r_3)	GARR_4 (r_4)	M_b	I_b Single fault
$R_1 \downarrow$	a_{R1}	$-a_{\text{R1}}$	a_{R1}	$\bar{a}_{\text{R1}} = (1-a_1)a_{\text{R1}}$
$R_1 \uparrow$	$-a_{\text{R1}}$	a_{R1}	a_{R1}	$\bar{a}_{\text{R1}} = (1-a_1)a_{\text{R1}}$
$R_2 \downarrow$	0	+1	1	0
$R_2 \uparrow$	0	−1	1	$\bar{a}_2 = (1-a_2)$
$R_{\text{d1}} \downarrow$	a_1	$-a_1$	a_1	$\bar{a}_1 = (1-a_{\text{R1}})a_1$
$R_{\text{d1}} \uparrow$	$-a_1$	a_1	a_1	$\bar{a}_1 = (1-a_{\text{R1}})a_1$
$R_{\text{d2}} \downarrow$	0	a_2	a_2	0
$R_{\text{d2}} \uparrow$	0	$-a_2$	a_2	0
$R_{\text{Leak1}} \downarrow$	+1	0	1	1
$R_{\text{Leak2}} \downarrow$	0	+1	1	0

Table 6.12 MCSSM for the electrical hybrid system

Parameter	GARR_3 (r_3)	GARR_4 (r_4)	M_b	I_b Single fault
$a_{\text{R1}} \uparrow$	+1	-1	1	1
$a_{\text{R1}} \downarrow$	-1	+1	1	1

Table 6.13 Introduced faults in the experimental model

Parameter	Description	Degradation nature	Start time, t_{fi}, (s)
R_1	Resistance increase	Progressive type as per (6.34)	323.5
R_2	Resistance increase	Abrupt type as per (6.35)	1249.7

$$R_1(t,z) = \begin{cases} R_{1\text{n}}(z) \cdot (1/a_{\text{R1}}), & t < t_{\text{f1}} \\ (R_{1\text{n}}(z) + k(z)t_{\text{on}}) \cdot (1/a_{\text{R1}}), & t \geq t_{\text{f1}} \end{cases} \tag{6.34}$$

$$R_2(t,z) = \begin{cases} R_{2\text{n}}(z), & t < t_{\text{f2}} \\ 1.2R_{2\text{n}}(z), & t \geq t_{\text{f2}} \end{cases} \tag{6.35}$$

where $R_{\text{in}}(z)$ is the nominal parameter value of resistor $R_i (i = 1, 2)$ at corresponding operating mode (z), $k(z) = 2.35\,\Omega/\text{s}$ if $z = z^{(1)} = a_{\text{R1}} = 1$, each $z^{(1)}$ is for 8 s and $k(z) = 0\,\Omega/\text{s}$ if $z = z^{(2)} = a_{\text{R1}} = 0$, each $z^{(2)}$ is for 3 s, $t_{\text{on}} = \int_{t_{\text{f1}}}^{t} a_{\text{R1}} \cdot dt$, and t_{f1} and t_{f2} are the time instances when the first and second faults start, respectively.

In a similar way as done in simulation (Sect. 6.3.3.1), sequential multiple faults are detected and isolated by the new proposed method using the experimental data of the electrical hybrid system and RUL is predicted for progressive faults. The evaluated residuals and adaptive thresholds with the real measurement data collected from the experimental setup are shown in Fig. 6.17. Note that due to discrete derivatives, residuals show spikes that appear during mode transitions. The observed coherence vector just after 323.5 s is $\boldsymbol{C} = [-1\ +1]$ which gives $R_1 \uparrow$ and $a_{\text{R1}} \downarrow$ as SSF elements at $a_1 = 0$; otherwise $R_1 \uparrow$, $R_{\text{d1}} \uparrow$ and $a_{\text{R1}} \downarrow$ are

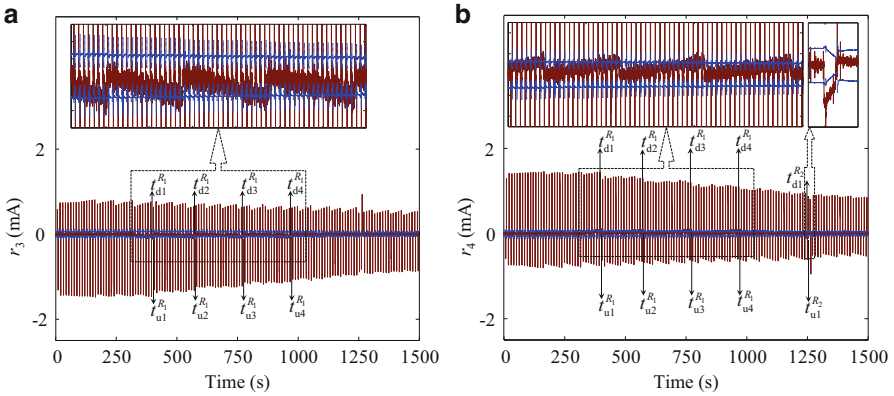

Fig. 6.17 Response of residuals (**a**) r_3 and (**b**) r_4 using proposed method with dynamically updated parameter and adaptive threshold

Table 6.14 Estimated data points at different time instances

Parameter	Detected time (s) $(t_{ui}^{R_j})$	Updated time (s) $(t_{ui}^{R_j})$	t_{on} (s)	Estimated magnitude $R_j^f(t_{ui}^{R_j}, z)$ (Ω)
$R_1 \uparrow$	399.768	402.768	402.768	3326.53
	565.572	568.572	523.572	3605.01
	749.380	752.380	656.380	3917.57
	938.692	941.692	794.692	4250.02
$R_2 \uparrow$	1249.702	1252.702	–	3740.5

SSF elements at $a_1 = 1$. Consistency in mode $a_{R1} = 1$ is found by evaluating the all sensitive ARRs at current mode information which indicates the observed inconsistency is due to parametric fault. Upon parameter estimation and successive updation of DHBG model at various time instances, it is confirmed that the fault $R_1 \uparrow$ is of progressive nature (see Fig. 6.17 where residuals r_3 and r_4 are updated at various time instances during degradation pattern identification of R_1). The estimated magnitude of degrading resistance R_1 at the mode $z = z^{(1)} = a_{R1} = 1$ at different time instances $t_{ui}^{R_1}(i = 1, 2, 3, 4)$ is shown in Table 6.14.

Initially, the obtained data points at time instances $(t_{u1}^{R_1})$ and $(t_{u2}^{R_1})$ corresponding to mode $z = z^{(1)} = a_{R1} = 1$ are used to find the linear degradation model for the resistor R_1 as presented in Fig. 6.18a. Initial estimated value of RUL using the linear degradation model at on state of resistor R_1 and infinite resistance value at off state is found as 1672.92 s (see Fig. 6.19a). This initial degradation model is further updated with modified model when newer information of parameter estimates are obtained during monitoring.

As the monitoring is continued, the new information of degradation data of resistor $R_1(t, z)$ is obtained via parameter estimation at time instances $t_{ui}^{R_1}, (i = 3, 4)$ at the same mode $z = z^{(1)} = a_{R1} = 1$, which are presented in Table 6.14. Various

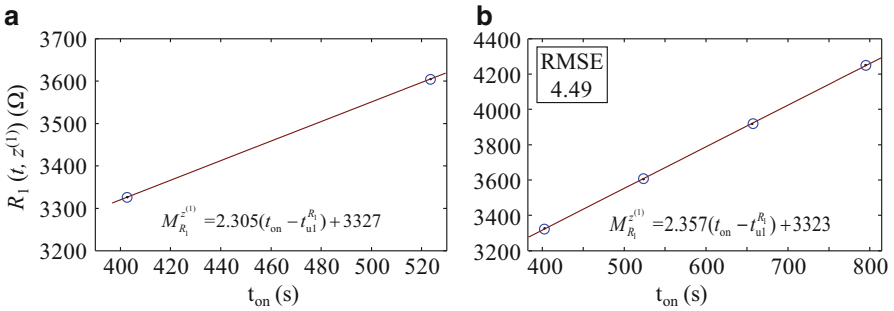

Fig. 6.18 (**a**) Linear degradation model at mode $a_{R1} = 1$ as two data points are known, (**b**) finally identified degradation model at mode $a_{R1} = 1$ with new information of data points

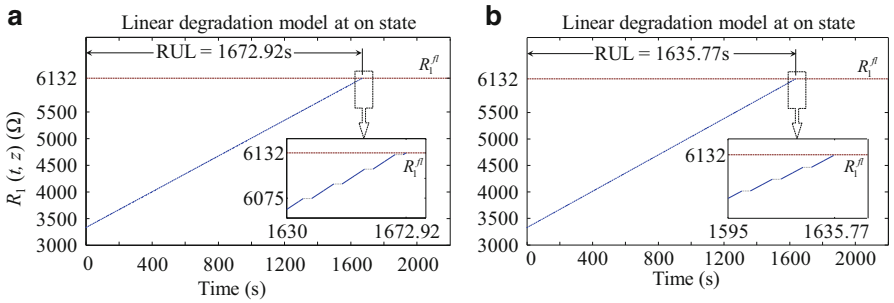

Fig. 6.19 Estimated RUL of parameter R_1 (**a**) With initially assumed linear degradation model when only two data points are known (**b**) with finally confirmed linear degradation model with the new information of degradation state at on state of resistor R_1, and considering infinite resistance in off state

degradation models are tried and linear degradation model was found as a best fit model for on state of the resistor R_1 ($a_{R1} = 1$) according to goodness of fit (RMSE = 4.49). Indeed, this is the actual degradation introduced in the experiment. The final linear degradation model having coefficient P_1 and P_2 with 95 % confidence bound is given as

$$M_{R_1}^{z^{(1)}} = P_1(t_{on} - t_{u1}^{R_1}) + P_2 \tag{6.36}$$

where $P_1 = 2.357$ and $P_2 = 3323$.

Using the degradation model (6.36), the RUL is estimated as 1635.77 s (see Fig. 6.19b) according to set failure threshold ($R_1^{fl} = 2R_1$).

When the abrupt fault in resistor R_2 at 1249.7 s is inserted, the coherence vector ($\boldsymbol{C}$) just after 1249.7 s, i.e., after fault initiation in resistor R_2, is obtained from Fig. 6.17a, b as $\boldsymbol{C} = [0{-}1]$. This gives the possible SSF as $R_2 \uparrow$ (i.e., increasing fault in resistor R_2) using GFSSM since $a_2 = 0$. Even though the fault $R_2 \uparrow$ is directly isolable at $a_2 = 0$, we still need to estimate the fault magnitude for updating the

DHBG model for diagnosis of subsequent faults. In this case, $R_2 \uparrow$ is found as actual fault parameter with estimated fault magnitude of $R_2^f \approx 1.2R_2$ at $(t_{\text{ul}}^{R_2}) = 1252.702\,\text{s}$ and this estimated parameter value is then onwards considered as the new nominal parameter value.

6.6 Conclusions

This chapter proposes a strategy for fault diagnosis and prognosis of hybrid dynamical system by using bond graph modelling as a common framework for system modelling, virtual prototyping, fault diagnosis rule development, parameter and system identification, and RUL estimation. The proposed approach detects and isolates sequential multiple faults of different types, i.e., discrete mode faults, abrupt and progressive parametric faults; and also predicts the RUL if the detected fault is due to progressive parameter drift. The developed method is first applied to a simulated benchmark problem and then it is experimentally validated on a scaled equivalent electrical circuit model.

The response time (time constant) of the system and the time taken for degradation model are the two critical parameters that govern the applicability of this diagnosis and prognosis scheme. For processes or systems with slow response time, the time spent for parameter estimation is not significant. However, for fast systems like the equivalent electrical system considered in this study, the parameter estimation needs to be faster. This is achieved by narrowing down the number of suspected faults and using sensitivity signature for information regarding fault direction. Global fault sensitivity signatures of the hybrid dynamical system are used to identify the possible fault directions from the residual responses. The information of operating mode, possible fault candidates and directions of corresponding parameter deviations are used in a constrained least square error minimization-based estimation of fault parameters. It was found that richer information gleaned from sensitivity signatures allows quicker and reliable identification of faults and degradation model for RUL estimation.

For RUL estimation, use of multiple degradation models which include operational modes as additional control parameter and evolve through degradation model identification is considered in this study for hybrid dynamical system. Models are continually evolved with time by adapting to the new information of the state of degradation of the monitored system to provide accurate RUL with bounded uncertainty value. This overcomes the drawbacks of various existing methods.

References

1. Vachtsevanos, G., Lewis, F. L., Roemer, M., Hess, A., & Wu, B. (2006). *Intelligent fault diagnosis and prognosis for engineering systems*. Hoboken: Wiley.
2. Mukherjee, A., Karmakar, R., & Samantaray, A. K. (2006). *Bond graph in modelling, simulation and fault identification*. New Delhi: I. K. International Pvt. Ltd.

3. Borutzky, W. (2010). *Bond graph methodology: Development and analysis of multidisciplinary dynamic system models*. London: Springer.
4. Mosterman, P. J., & Biswas, G. (1995). Behavior generation using model switching: A hybrid bond graph modelling technique. *Transactions of The Society for Computer Simulation, 27*(1), 177–182.
5. Roychoudhury, I., Daigle, M. J., Biswas, G., & Koutsoukos, X. (2010). Efficient simulation of hybrid systems: A hybrid bond graph approach. *Simulation: Transactions of the Society for Modeling and Simulation International, 87*(6), 467–498.
6. Wang, D., Yu, M., Low, C. B., & Arogeti, S. (2013). *Model-based health monitoring of hybrid systems*. New York: Springer.
7. Samantaray, A.K., & Bouamama, B. O. (2008). *Model-based process supervision: A bond graph approach*. London: Springer.
8. Borutzky, W. (2015). *Bond graph model-based fault diagnosis of hybrid systems*. Switzerland: Springer.
9. Levy, R., Arogeti, S., Wang, D., & Fivel, O. (2015). Improved diagnosis of hybrid systems using instantaneous sensitivity matrices. *Mechanism and Machine Theory, 91*, 240–257.
10. Narasimhan, S., & Biswas, G. (2007). Model-based diagnosis of hybrid systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 37*(3), 348–361.
11. Ghoshal, S. K., Samanta, S., & Samantaray, A. K. (2012). Robust fault detection and isolation of hybrid systems with uncertain parameters. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 226*(8), 1013–1028.
12. Samantaray, A. K., Medjaher, K., Bouamama, B. O., Staroswiecki, M., & Dauphin-Tanguy, G. (2006). Diagnostic bond graphs for online fault detection and isolation. *Simulation Modelling Practice and Theory, 14*, 237–262.
13. Medjaher, K., & Zerhouni, N. (2009). Residual-based failure prognostic in dynamic systems. In *The 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes Held in Barcelona, Spain* (pp. 716–721).
14. Medjaher, K., & Zerhouni, N. (2013). Hybrid prognostic method applied to mechatronic systems. *International Journal of Advanced Manufacturing Technology, 69*(1–4), 823–834.
15. Ming, Y., Wang, D., & Luo, M. (2015). An integrated approach to prognosis of hybrid systems with unknown mode changes. *IEEE Transactions on Industrial Electronics, 62*(1), 503–515.
16. Jha, M. S., Dauphin-Tanguy, G., & Bouamama, B. O. (2016). Particle filter based hybrid prognostics for health monitoring of uncertain systems in bond graph framework. *Mechanical Systems and Signal Processing, 75*, 301–329.
17. Samantaray, A. K., & Ghoshal, S. K. (2007). Sensitivity bond graph approach to multiple fault isolation through parameter estimation. *Proceedings of the Institution of Mechanical Engineers. Part-I: The Journal of Systems and Control Engineering, 221*(4), 577–587.
18. Gawthrop, P. J. (2000). Sensitivity bond graphs. *Journal of the Franklin Institute, 337*(7), 907–922.
19. Borutzky, W. (2011). Incremental bond graphs. In *Bond graph modelling of engineering systems* (pp. 135–176). New York: Springer.
20. Merzouki, R., Samantaray, A. K., Pathak, P. M., & Bouamama, B. O. (2012). *Intelligent mechatronic systems: Modeling, control and diagnosis*. New York: Springer Science and Business Media.
21. Low, C. B., Wang, D., Arogeti, S., & Luo, M. (2009). Fault parameter estimation for hybrid systems using hybrid bond graph. In *Proceedings of the 3rd IEEE Multi-Conference on Systems and Control, Saint Petersburg, Russia*, July 8–10 (pp. 133–134).