Unknown input modeling and robust fault diagnosis using black box observers

Seema Manuja [a], Shankar Narasimhan [a], Sachin C. Patwardhan [b]

[a] *Department of Chemical Engineering, Indian Institute of Technology, Madras, Chennai 600 036, India*
[b] *Department of Chemical Engineering, Indian Institute of Technology, Bombay, Powai, Mumbai 400 076, India*

Keywords:
Robust fault diagnosis
Unknown input observer
Fault tolerant control
GLR method

A B S T R A C T

A key issue that needs to be addressed while performing fault diagnosis using black box models is that of *robustness* against abrupt changes in unknown inputs. A fundamental difficulty with the robust FDI design approaches available in the literature is that they require some *a priori* knowledge of the model for unmeasured disturbances or modeling uncertainty. In this work, we propose a novel approach for modeling abrupt changes in unmeasured disturbances when innovation form of state space model (i.e. black box observer) is used for fault diagnosis. A disturbance coupling matrix is developed using singular value decomposition of the extended observability matrix and further used to formulate a robust fault diagnosis scheme based on generalized likelihood ratio test. The proposed modeling approach does not require any *a priori* knowledge of how these faults affect the system dynamics. To isolate sensor and actuator biases from step jumps in unmeasured disturbances, a statistically rigorous method is developed for distinguishing between faults modeled using different number of parameters. Simulation studies on a heavy oil fractionator example show that the proposed FDI methodology based on identified models can be used to effectively distinguish between sensor biases, actuator biases and other soft faults caused by changes in unmeasured disturbance variables. The fault tolerant control scheme, which makes use of the proposed robust FDI methodology, gives significantly better control performance than conventional controllers when soft faults occur. The experimental evaluation of the proposed FDI methodology on a laboratory scale stirred tank temperature control set-up corroborates these conclusions.

1. Introduction

Fault diagnosis is primarily concerned with recognizing anomalous behavior in the functioning of process components, such as sensors and actuators, and abnormal trends in process parameters. Occurrences of such faults lead to degradation in control loop performance and can result in reduced profits. Over the last three decades, a large number of model based methods have been proposed in the literature for fault detection and identification (FDI). Based on the type of models used for representing the normal and the faulty behavior, these approaches can be broadly classified as data-driven, analytical redundancy based and knowledge based systems [1]. The problem of designing an FDI system essentially involves generation of signals (residuals) containing information on the fault followed by devising a suitable decision logic mechanism.

The initial focus of work in fault detection and identification (FDI) was on deriving approaches for isolating faults using the operating data and not much attention was paid to using this information for on-line fault accommodation and performance recovery. The upcoming research trend in the last decade features an ever increasing integration of fault diagnosis in the design and development of control laws. Various design approaches reported in the literature for achieving fault tolerance in control can be broadly classified as passive and active schemes [2,3]. The passive approach is essentially based on the principle of robust control where the controller is made insensitive to specified faults at the design stage. This approach requires representation of the fault effect just as if it were a source of modeling uncertainty. As a consequence, only a restricted class of faults can be accommodated in this approach.

A more efficient approach to achieving fault tolerance is the active scheme, which makes explicit fault correction as and when a fault is identified. Among the various approaches proposed in the literature, analytical redundancy based methods are found to be most suitable for devising an active fault tolerant scheme, since the information it provides facilitates explicit corrective measure. Recently, Prakash et al. [4,5] have developed an active fault tolerant control scheme (FTCS), which integrates a model based FDI methodology based on generalized likelihood ratio (GLR) method [6,7] with a controller, through an on-line compensation mechanism. The prime requirement of GLR method is the availability of a process model (estimator) for normal operation and fault modes, describing the process behavior for each hypothesized fault. A salient feature of their work is the use of first principles model enabling root cause analysis of various types of soft faults like sensor and actuator biases along with abrupt changes in disturbances and process parameters. However, the development of such mechanistic models is often not so feasible and practicable for complex chemical processes. On the other hand, development of an innovations form of state space models (or time series models) from input–output data are relatively easy.

Recently, Patwardhan et al. [8] have shown that, it is possible to detect and identify sensor and actuator biases using linear black box observers identified from input–output data and employing the GLR method. However, their approach, as proposed, has a severe limitation since it cannot distinguish between actuator biases and abrupt changes in unmeasured disturbances. Thus, the key issue that needs to be addressed while performing fault diagnosis with such black box models is that of *robustness* against unmodeled process changes. The robust FDI methods developed so far essentially aim at detecting and identifying sensor and actuator faults while desensitizing the effect of modeling uncertainty. The two main approaches for achieving this robustness are through (i) the use of adaptive thresholds for residual evaluation, and (ii) generation of disturbance-decoupled residuals [13]. The design principle in adaptive thresholds involves modeling the uncertainty in system description as a function of process operating conditions. The relationship is derived either theoretically or empirically, and is used to vary the test thresholds accordingly while evaluating the residuals for FDI [10–12]. In the disturbance-decoupled approach, the residuals are generated in such a manner so as to minimize (or completely decouple) the influence of model uncertainty while retaining sensitivity to faults. The available design tools are unknown input observer [13,14], eigenstructure assignment [9] and frequency domain approaches [15].

A fundamental difficulty with all these design approaches is that they require some *a priori* knowledge of the model for unmeasured disturbances or modeling uncertainty. A typical uncertainty description makes use of the concept of *unknown inputs* acting upon the normal linear model of the system where the associated distribution matrix is assumed to be known. In general, for real applications such a knowledge is not available and needs to be determined either from *a priori* knowledge of the process dynamics or identified using a suitable model for uncertainty [9,16]. When the disturbances are not measured, identifying the associated distribution matrix is a nontrivial exercise. Further, a major constraint in all these robust FDI schemes is the availability of adequate design freedom for performing requisite fault isolation along with robustness.

Depending upon the diagnosis philosophy that is adopted, persistent changes in the mean of unmeasured variables can either be classified as faults or simply treated as unmeasured disturbances. In the robust FDI methods proposed in the literature, persistent changes in unmeasured disturbances are not treated as faults (as the diagnostic algorithm is made insensitive or robust to such changes). In the present work, however, we adopt an active approach and treat the mean shifts in unmeasured disturbances as faults. Even though root cause analysis of faults caused due to unmeasured disturbances/process parameters is not possible using a model identified purely from input–output data under normal operating conditions, the overall diagnostic ability of FDI scheme and the performance of FTCS is severely impaired if these changes are ignored. Thus, the main objective of this work is to develop an active online robust fault diagnosis and accommodation scheme based on linear black box observers that can detect and isolate abrupt changes in unmeasured disturbances (classified as unknown input fault) while distinguishing them from sensor and actuator faults. To achieve this objective, we first propose a new approach to construct of distribution matrix for the unknown inputs in state dynamics, which does not require any *a priori* information regarding such fault characteristics. Secondly, we have also proposed a statistically rigorous modified procedure which extends the use of likelihood ratio test statistics for distinguishing between faults modeled using different number of parameters. Finally, the proposed robust FDI method is also integrated with fault tolerant control strategies developed earlier for online fault accommodation. We test both the diagnostic and the control performance of the proposed fault accommodation method by conducting simulation studies on the benchmark shell control problem and experimental verification on a laboratory scale system of stirred tank heaters.

The paper is organized into six sections. Section 2 describes the proposed FDI method with linear identified models. Section 3 briefly explains the fault tolerant control scheme (FTCS). Section 4 contains the simulation studies carried out on the shell control problem of heavy oil fractionator example. Section 5 gives the experimental studies conducted on a coupled two-tank temperature control process. Based on the simulation and experimental results, main conclusions of the work are reported in Section 6.

2. Models for normal and fault modes

The FDI strategy in the present work uses the generalized likelihood ratio (GLR) method originally proposed by Willsky [6]. This approach requires models for normal and faulty operating modes. In this section we first describe models for normal operation and sensor/actuator biases. We then proceed to develop a model for isolating abrupt changes in unknown input (unmeasured disturbance).

2.1. Models for normal operation and sensor/actuator biases

Consider an innovations form of state space model directly identified from the input–output data given by [17]

$$\mathbf{x}(k+1) = \mathbf{\Phi}\mathbf{x}(k) + \mathbf{\Gamma}_u\mathbf{u}(k) + \mathbf{K}\mathbf{e}(k) \quad (1)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{e}(k) \quad (2)$$

where $\mathbf{x}(k) \in R^n$ represents the states, $\mathbf{u}(k) \in R^m$ represents the manipulated inputs, $\mathbf{y}(k) \in R^r$ represents the measured outputs and $\mathbf{e}(k) \in R^r$ represents zero mean innovation sequence with covariance $\mathbf{Q}$. This model can be expressed in standard form as follows:

$$\mathbf{x}(k+1) = \mathbf{\Phi}\mathbf{x}(k) + \mathbf{\Gamma}_u\mathbf{u}(k) + \mathbf{w}(k) \quad (3)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k) \quad (4)$$

Here, $\mathbf{w}(k) \in R^n$ and $\mathbf{v}(k) \in R^r$ represent zero mean white noise processes with

$$R_1 = E(\mathbf{w}(k)\mathbf{w}(k)^{\mathrm{T}}) = \mathbf{KQK}^{\mathrm{T}} \quad (5)$$

$$R_2 = E(\mathbf{v}(k)\mathbf{v}(k)^{\mathrm{T}}) = \mathbf{Q} \quad (6)$$

$$R_{12} = E(\mathbf{w}(k)\mathbf{v}(k)^{\mathrm{T}}) = \mathbf{KQ} \quad (7)$$

Note that estimates of the Kalman gain matrix $\mathbf{K}$ and covariance matrix $\mathbf{Q}$ are generated as a part of the parameter identification exercise [17].

In order to isolate faults using GLR method, models for each type of hypothesized fault are also required. In the presence of a sensor bias, the measurement model is represented as

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k) + b_{y,i}\mathbf{e}_{y,i}\sigma(k-t) \quad (8)$$

where t is the time of occurrence of a fault and $\sigma(k-t)$ represents unit step function such that

$$\sigma(k-t) = \left\{ \begin{matrix} 0 & \text{for } k < t \\ 1 & \text{for } k \geqslant t \end{matrix} \right\}$$

Likewise, in the presence of an actuator bias, the state transition is modeled as follow:

$$\mathbf{x}(k+1) = \mathbf{\Phi}\mathbf{x}(k) + \mathbf{\Gamma}_u\mathbf{u}(k) + \mathbf{w}(k) + b_{u,i}\mathbf{\Gamma}_u\mathbf{e}_{u,i}\sigma(k-t) \tag{9}$$

In the above equations, $b_{f,i}$ represents the magnitude of the fault and $\mathbf{e}_{f,i}$ represents fault direction vector where $f \in (y,u)$. Here, $\mathbf{e}_{f,i}$ represents a unit vector of appropriate dimensions with unity at ith location and all remaining elements equal to zero.

Remark 1. The innovation form of state space model given by Eqs. (1) and (2) is equivalent to the following discrete transfer function model

$$\mathbf{y}(k) = \mathbf{G}(q)\mathbf{u}(k) + \mathbf{H}(q)\mathbf{e}(k) \tag{10}$$

$$\mathbf{G}(q) = \mathbf{C}[q\mathbf{I} - \mathbf{\Phi}]^{-1}\mathbf{\Gamma}_u; \quad \mathbf{H}(q) = \mathbf{C}[q\mathbf{I} - \mathbf{\Phi}]\mathbf{K} + \mathbf{I} \tag{11}$$

Thus, the stationary unmeasured disturbances affecting the process under normal operating conditions are modelled as a Gaussian white noise process $\{\mathbf{e}(k)\}$ passing through a filter $\mathbf{H}(q)$, i.e. as a colored noise signal.

2.2. *Modeling of unknown input fault*

Using the above identified model, sensor and actuator biases can be easily identified. However, if an abrupt change occurs in unmeasured disturbances/ parameter, then it gets misidentified as an actuator bias [8] as no model is hypothesized for such a change. In this section, we develop a lumped model for disturbance/parametric faults, which captures the unqualified bias effect on state dynamics. Note that all faults caused by step changes in unmeasured disturbances or process parameters (referred to as *unknown inputs*) are lumped and modeled as a single artificial fault and no attempt is made to isolate the true cause of this fault.

We propose to model the step changes in the unknown inputs occurring at time t using an unknown additive vector to the state evolution equation as follows:

$$\mathbf{x}(k+1) = \mathbf{\Phi}\mathbf{x}(k) + \mathbf{\Gamma}_u\mathbf{u}(k) + \mathbf{w}(k) + \sigma(k-t_f)\mathbf{f} \tag{12}$$

where $\mathbf{f} \in R^n$ is an unknown vector which represents the direction and magnitude of the fault. In the absence of any *a priori* knowledge as to which states are affected by the fault, we have assumed that all the states are affected. It is also assumed that the fault direction and magnitude defined by $\mathbf{f}$ is time invariant, which may be valid in the neighborhood of the current operating point for faults that occur as step changes. Recently, Patton and Chen [9,16] have described the above approach as one possible way to model the unknown inputs. This approach is similar to the one used in model predictive control to handle disturbance and parameter changes without explicit knowledge of the source of fault.

One possibility is to view the elements of the vector $\mathbf{f}$ as additional states and estimate them from input–output data. Consider the deterministic part of the model corresponding to Eq. (12) together with additional equation $\mathbf{f}(k+1) = \mathbf{f}(k)$. Using the following equivalent augmented model, we have

$$\mathbf{x}_a(k+1) = \mathbf{\Phi}_a\mathbf{x}_a(k) + \mathbf{\Gamma}_{u,a}\mathbf{u}(k) \tag{13}$$

$$\mathbf{y}(k) = \mathbf{C}_a\mathbf{x}_a(k) \tag{14}$$

where

$$\mathbf{x}_a(k) = \begin{bmatrix} \mathbf{x}(k) \\ \mathbf{f}(k) \end{bmatrix}; \quad \mathbf{\Phi}_a = \begin{bmatrix} \mathbf{\Phi} & I \\ [0] & I \end{bmatrix}$$

$$\mathbf{\Gamma}_{u,a} = \begin{bmatrix} \mathbf{\Gamma}_u \\ [0] \end{bmatrix}; \quad \mathbf{C}_a = [\mathbf{C} \quad [0]]$$

The observability matrix of the above augmented system is given by

$$\mathbf{w}_a = [\mathbf{w}_x \quad \mathbf{w}_f] = \begin{bmatrix} \mathbf{C} & [0] \\ \mathbf{C\Phi} & \mathbf{C} \\ \mathbf{C\Phi}^2 & \mathbf{C\Phi} \\ \vdots & \vdots \\ \mathbf{C\Phi}^{n-1} & \mathbf{C\Phi}^{n-2} \end{bmatrix} \tag{15}$$

where $\mathbf{w}_x$ represents observability gramian for the normal behavior model given by Eqs. (3) and (4). Assuming that the normal model is observable i.e. Rank $(\mathbf{w}_x) = n$, it can be shown that the rank of matrix $\mathbf{w}_a$ cannot exceed $n + r$ where r represents the dimension of the measurement vector ($\mathbf{y}$). Thus, from the measured data we can estimate at most $n + r$ elements of the augmented state vector $\mathbf{x}_a(k)$. If the number of states is less than or equal to the number of measurements, then all elements of the unknown fault vector $\mathbf{f}$ can be estimated [9,16]. However, this requirement is unrealistic in the case of chemical processes where the number of measured variables (r) is typically far less than the number of state variables (n). This implies that in order to estimate the unknown fault vector we have to parameterize it using at most r unknown parameters.

Now, in the absence of noise and knowledge of $\mathbf{x}(k-n+1)$, we can find an $\mathbf{f}$ that exactly satisfies the following equation

$$\mathbf{w}_f\mathbf{f} = \mathbf{Y} - \mathbf{w}_x\mathbf{x}(k-n+1) \tag{16}$$

$$\mathbf{Y} = [\mathbf{y}(k-n+1)^{\mathrm{T}} \quad \mathbf{y}(k-n+2)^{\mathrm{T}} \quad \dots \quad \mathbf{y}(k)^{\mathrm{T}}]^{\mathrm{T}} \tag{17}$$

Let us denote RHS of Eq. (16) by the vector $\mathbf{Y}_x = \mathbf{Y} - \mathbf{w}_x\mathbf{x}(k-n+1)$. We propose to base diagnosis of fault $\mathbf{f}$ on best fit of model residuals, which in the absence of noise and exact knowledge of $\mathbf{x}(k-n+1)$, implies best fit of vector $\mathbf{Y}_x$. Any lower dimensional approximation we make to $\mathbf{f}$ will result in a corresponding lack of fit to $\mathbf{Y}_x$. Thus, we choose the lower dimensional approximation of $\mathbf{f}$, which causes the least error in fit to observed $\mathbf{Y}_x$. Otherwise, we may not correctly classify it as fault $\mathbf{f}$ and may identify some other fault that better fits $\mathbf{Y}_x$. Our choice of model for abrupt changes in the unmeasured disturbances is motivated by the following result.

Theorem 1. *Let $\mathbf{f}_{\text{approx}} = \mathbf{\Gamma}_d\mathbf{d}$ where $\mathbf{\Gamma}_d$ is a $n \times r$ matrix of full column rank, $\mathbf{d} \in R^r$ and vector ε_d denotes the approximation error such that*

$$\mathbf{w}_f\mathbf{f} = \mathbf{Y}_x = \mathbf{w}_f\mathbf{f}_{\text{approx}} + \varepsilon_d \tag{18}$$

Also, let the singular value decomposition of matrix $\mathbf{w}_f$ be defined as follows:

$$\mathbf{w}_f = \mathbf{U\Sigma}V^{\mathrm{T}} \tag{19}$$

$$\mathbf{\Sigma} = [\sigma_1\mathbf{e}^{(1)}\sigma_2\mathbf{e}^{(2)} \quad \dots \quad \sigma_n\mathbf{e}^{(n)}]$$

$$\mathbf{V} = [\mathbf{v}^{(1)} \quad \mathbf{v}^{(2)} \quad \dots \quad \mathbf{v}^{(n)}]$$

where $\mathbf{U}$ and $\mathbf{V}$ are square unitary matrices of dimension $2nr \times 2nr$ and $n \times n$, respectively, vector $\mathbf{e}^{(i)} \in R^{2rn}$ represents unit vector such that ith element of the vector is unity and rest all elements are zero, vector $\mathbf{v}^{(i)} \in R^n$ represents ith column of matrix $\mathbf{V}$ and $\{\sigma_i : i = 1, 2, \dots, n\}$ represent the singular values of $\mathbf{w}_f$ such that $\sigma_1 \geqslant \sigma_2 \geqslant \dots \geqslant \sigma_n$. Then, the choice matrix $\mathbf{\Gamma}_d$ that yields the smallest upper bound on the relative approximation error, $\|\varepsilon_d\|_2/\|\mathbf{f}\|_2$, for an arbitrary vector $\mathbf{f}$ corresponds to

$$\mathbf{\Gamma}_d = [\mathbf{v}^{(1)} \quad \dots \quad \mathbf{v}^{(r)}] \tag{20}$$

Proof. Refer to Appendix A. □

Thus, in absence of any knowledge of fault vector $\mathbf{f}$ a fault model for unknown inputs is chosen as

$$\mathbf{x}(k+1) = \mathbf{\Phi}\mathbf{x}(k) + \mathbf{\Gamma}_u\mathbf{u}(k) + \mathbf{w}(k) + \mathbf{\Gamma}_d\mathbf{d}\sigma(k-t) \tag{21}$$

where the unknown input fault coupling matrix $\mathbf{\Gamma}_d$ is given by Eq. (20). It may be noted that $\mathbf{\Gamma}_d$ is completely determined from the system matrices $(\mathbf{C}, \mathbf{\Phi})$ with no *a priori* knowledge about specific instances of fault.

3. Robust FDI using black box observer

The normal behavior model given by Eqs. (3)–(7), can be used to obtain the optimal state estimates from the Kalman filter (referred to as *normal* Kalman filter in the rest of the text) for use in FDI [8,18]

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{\Phi}\hat{\mathbf{x}}(k|k-1) + \mathbf{\Gamma}_u\mathbf{u}(k) + \mathbf{K}\gamma(k) \tag{22}$$

$$\gamma(k) = \mathbf{y}(k) - \mathbf{C}\hat{x}(k|k-1) \tag{23}$$

where $\gamma(k)$ is a vector of innovations (or residuals) and $\hat{\mathbf{x}}(k|k-1)$ denote the state estimates predicted at time k using all measurements made up to time $(k-1)$. It can be shown that, under a fault free situation, the residuals (or innovations) $\gamma(k)$ form a Gaussian zero mean white noise sequence. Therefore, a simple statistical test referred to as the fault detection test (FDT) is applied at each time instant to estimate the time of occurrence of a fault. If FDT is rejected at some time t, the occurrence of a fault is further confirmed by applying a cumulative statistical test known as the fault confirmation test (FCT) on the innovation sequence in the time interval $[t, t+N]$ (see [4,8] for details). If this test rejects the null hypothesis (no fault hypothesis) then the occurrence of the fault at time t is confirmed.

3.1. Review of conventional GLR method for fault isolation

After a fault is detected using FDT and later confirmed using FCT, we make use of the GLR method for identifying the fault and estimating its magnitude. It can be noted from the fault model Eqs. (8), (9), and (21), that while sensor and actuator biases can be described in terms of one unknown parameter (representing the fault magnitude), unknown input faults are characterized by r unknown parameters (representing the fault magnitude and direction). We develop a modified GLR method in order to identify faults characterized using different number of unknown parameters. To motivate our development, we review the conventional GLR method that can be used to identify sensor and actuator biases.

Fault identification in the GLR method is carried out using the innovation sequence $\{\gamma(t)\ldots\gamma(t+N)\}$ generated by the *normal* Kalman filter over the time window $[t, t+N]$. It can be noted from Eqs. (8) and (9) that the effect of sensor and actuator biases is described by additive changes to the normal process behavior. Thus, under different fault hypothesis, the innovation sequence generated by the *normal* Kalman filter can be shown to be a Gaussian random process with different unknown means $\boldsymbol{\mu}_{f,i}(k;t)$, but with the same known covariance matrix $\mathbf{V}$. The hypothesis for the presence or absence of a fault in the observed data can be written as

$$H_0 : \boldsymbol{\mu}(k;t) = 0$$

$$H_{1f,i} : \boldsymbol{\mu}_{f,i}(k;t) = b_{f,i}\mathbf{G}_f(k;t)\mathbf{e}_{f,i}$$

$$k \in [t, t+N] \quad \text{and} \quad f \in y, u$$

The signature matrix $\mathbf{G}_f(k;t)$ for different types of faults can be computed recursively using the normal Kalman filter equations and the fault models. For actuator and sensor biases, the recurrence relations for computing the signature matrices are as follows [8]:

- *Sensor bias:*

$$\mathbf{G}_y(k;t) = \mathbf{I} - \mathbf{C\Phi J}_y(k-1;t) - \mathbf{CKG}_y(k-1;t) \tag{24}$$

$$\mathbf{J}_y(k;t) = \mathbf{\Phi J}_y(k-1;t) + \mathbf{KG}_y(k-1;t) \tag{25}$$

- *Actuator bias:*

$$\mathbf{G}_u(k;t) = -\mathbf{C\Phi J}_u(k-1;t) + \mathbf{C\Gamma}_u - \mathbf{CKG}_u(k-1;t) \tag{26}$$

$$\mathbf{J}_u(k;t) = \mathbf{\Phi J}_u(k-1;t) - \mathbf{\Gamma}_u + \mathbf{KG}_u(k-1;t) \tag{27}$$

The GLR method identifies the fault corresponding to the most probable hypothesis. For this purpose, the likelihood ratio test statistics, $T_{f,i}$, corresponding to each fault hypothesis can be constructed. The test statistics are given by

$$T_{f,i} = d_{f,i}^2/c_{f,i} \tag{28}$$

$$d_{f,i} = \mathbf{e}_{f,i}^{\mathrm{T}} \sum_{k=t}^{t+N} \mathbf{G}_f^{\mathrm{T}}(k;t)\mathbf{V}^{-1}\gamma(k) \tag{29}$$

$$c_{f,i} = \mathbf{e}_{f,i}^{\mathrm{T}} \sum_{k=t}^{t+N} \mathbf{G}_f^{\mathrm{T}}(k;t)\mathbf{V}^{-1}\mathbf{G}_f(k;t)\mathbf{e}_{f,i} \tag{30}$$

The fault corresponding to the maximum of the above test statistics is identified as the fault that has occurred. The maximum likelihood estimate of the fault magnitude can also be computed as follows:

$$\hat{b}_{f,i} = \frac{d_{f,i}}{c_{f,i}}$$

Instead of comparing the test statistics, fault identification can equivalently carried out by computing the upper quantile probabilities of the test statistics under the null hypothesis. It can be shown that under the null hypothesis, the GLR test statistics corresponding to sensor or actuator biases follow a χ^2 distribution with one degree of freedom [7]. The upper quantile probabilities $\alpha_{f,i}$ are defined as

$$\alpha_{f,i} = \Pr\{\chi_1^2 > T_{f,i}\}$$

where χ_1^2 denotes a χ^2 random variable with one degree of freedom. The fault which corresponds to the least value of $\alpha_{f,i}$ can be identified, which corresponds to the fault with the maximum test statistic. The advantage of using the upper quantile probabilities instead of the test statistics directly is that it allows the GLR method to be conveniently extended for identifying faults characterized by different number of parameters. However, numerical computation of the upper quantile probabilities may pose problems which need to be overcome. We describe our proposed modified GLR strategy for this purpose in the following subsection.

3.2. Modified GLR method for identifying unknown input faults

We first derive the test statistic for unknown input faults. Eq. (9) shows that the effect of unknown input faults is also modeled as an additive effect on the normal process behavior. Therefore, the innovation sequence in the time interval $[t, t+N]$ after the occurrence of an unknown input fault can be shown to be a Gaussian random process with expected value $\boldsymbol{\mu}_d(k,t) = \mathbf{G}_d(k,t)\mathbf{d}$. The recurrence relations for the signature matrix $\mathbf{G}_d(k,t)$ can be derived using the fault model given by Eq. (21) together with Eqs. (22) and (23). These are derived in Appendix B and are given by

$$\mathbf{G}_d(k;t) = -\mathbf{C\Phi J}_d(k-1;t) + \mathbf{C\Gamma}_d - \mathbf{CKG}_d(k-1;t) \tag{31}$$

$$\mathbf{J}_d(k;t) = \mathbf{\Phi J}_d(k-1;t) - \mathbf{\Gamma}_d + \mathbf{KG}_d(k-1;t) \tag{32}$$

$$k \in [t, t+N] \tag{33}$$

The GLR test statistic T_d for unknown input faults can now be obtained as

$$T_d = \mathbf{d}_d^{\mathrm{T}}\mathbf{C}_d^{-1}\mathbf{d}_d \tag{34}$$

$$\hat{\mathbf{d}} = [\mathbf{C}_d]^{-1}\mathbf{d}_d \tag{35}$$

$$\mathbf{d}_d = \sum_{k=t}^{t+N} \mathbf{G}_d^{\mathrm{T}}(k;t)\mathbf{V}^{-1}\gamma(k) \tag{36}$$

$$\mathbf{C}_d = \sum_{k=t}^{t+N} \mathbf{G}_d^{\mathrm{T}}(k;t)\mathbf{V}^{-1}\mathbf{G}_d(k;t) \tag{37}$$