Journal of Process Control 108 (2021) 68–85

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Design of functional observers for fault detection and isolation in nonlinear systems in the presence of noises

Sunjeev Venkateswaran, M. Ziyan Sheriff, Benjamin Wilhite, Costas Kravaris *

Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX 77843, USA

ARTICLE INFO

Article history:
Received 15 June 2021
Received in revised form 16 September 2021
Accepted 2 October 2021
Available online 23 November 2021

Keywords:
Fault detection
Functional observers
Residual generators
Generalized likelihood ratio

ABSTRACT

This work deals with the problem of designing functional observers for fault diagnosis in nonlinear systems in the presence of noises. It follows up previous work of the authors on design of functional observers (residual generators) for deterministic systems from the point of view of observer error linearization. Here, we consider the effect of noises on residual generation. The effect of sensor noises on the residuals is studied analytically and the associated probability distributions are derived. Following this, a well-known statistical hypothesis testing approach, Generalized Likelihood Ratio, is used to track changes in the mean of the residual to enable robust fault detection. The approach is also extended to process noises (plant-model mismatch) numerically. Throughout the study the methods presented are tested on a non-isothermal CSTR case study. The results show that the fault diagnosis scheme is able to quickly and accurately detect faults in the presence of both sensor and process noises.

1. Introduction

Higher demand for safety in process industries has made fault detection and isolation an active research area over the past two decades. A fault is an abnormal incident resulting in a large deviation in process variables from the usual conditions. They could arise due to numerous factors such as mechanical failures, human errors, and power failures. The consequences of faults can range from off spec production that can result in loss of revenue, to explosions that result in fatalities. A case in point being the 2007 T2 laboratories explosion in Jacksonville Florida, that occurred due to cooling system failure and resulted in four fatalities [1–3] and the 2013 William Olefins plant explosion in Geismar Louisiana, that resulted in two fatalities [4,5]. Such disastrous consequences warrant the need for robust fault diagnosis schemes that quickly detect and isolate the faults in the system and enable engineers or an automated fault-tolerant controller to take corrective action.

In literature fault diagnosis methods are broadly divided into two categories, (i) model-based (ii) statistical /data-driven [6–8]. Model-based methods are based on exploiting analytical redundancies present in the system. This comprises of a digital twin using a process model which is implemented in software form on a computer [9–14]. Analytical redundancy is achieved through the known interdependence among the process variables provided by the model [9–11,13–15]. The evolution of process variables of the digital twin will follow the outputs of the real system in the absence of faults and will show a measurable deviation in the presence of faults. The essence of analytical redundancy in fault diagnosis is checking consistency of the actual system behavior against the available system model. Any inconsistency is measured in terms of residuals that deviate from zero only in the presence of a specific fault. Moreover, since accurate modeling of a real system is difficult and the effect of unknown disturbances or uncertainties could corrupt the residual signal, it is important to carefully define the residual in a way that makes it unaffected by those disturbances. The central objective in model-based fault diagnosis is to develop a residual generator for each of the possible faults, in a way that the residual is unaffected by the other faults and unknown disturbances.

One of the most widely studied approaches in the area of model-based fault detection and isolation (FDI) is the observer-based fault diagnosis approach. The first observer-based FDI method for linear systems was proposed by Beard and Jones in the early 1970s [9,16,17] which was a historic milestone in the area of fault diagnosis. In general, observer based FDI methods for linear systems can be grouped into the following four categories [9,11,15] (i) Fault Detection Filter (ii) Diagnostic Observer (iii) Parity Space Approach (iv) Frequency Domain Approach. For a review of fault diagnosis for linear systems the reader is referred to excellent surveys by Frank and Ding [10,15]. There have been efforts seeking extensions of observer-based FDI methods to nonlinear systems in the spirit of linear systems methods. For

* Corresponding author.
E-mail address: kravaris@tamu.edu (C. Kravaris).

https://doi.org/10.1016/j.jprocont.2021.10.001
0959-1524/

example, [18] used the generalized inverse approach to design functional observers for FDI in Lipschitz non-linear systems in the presence of disturbances/uncertainties. [19,20] used concepts from input–output stability and L_2 stability respectively to develop a scheme for integrated design of observer-based fault detection systems in affine nonlinear systems. Recently, [21,22] proposed an observer based FDI approach from the point of view of observer error linearization for non-linear systems. The methods were a direct nonlinear generalization of standard linear FDI methods found in [9].

In practical processes, random disturbances resulting in measurement noises and/or plant-model mismatch are common, which necessitates formulating the system in a probabilistic setting [7,9]. In contrast to deterministic systems (the focal point of model-based methods), the future state of stochastic systems is not completely determined by the past and present states [7]. The measurements are considered to be a statistical time series and when the process is operating normally the observations (or residuals generated from the observations) correspond to a given probability distribution, and in the event of a fault/mishap (out of control) the underlying distributions change and can be characterized by a change in mean and/or variance. This enables the utilization of statistical methods, where samples of the output are taken sequentially and decisions (whether a change in mean/variance has occurred or not) can be made based on observations up to the current time.

Statistical methods have been implemented both as a purely data-driven scheme, where no knowledge of a quantitative model is assumed, and through hybrid frameworks as an add-on to existing model-based fault diagnosis schemes [7,23–26]. Purely data driven approaches in literature have made use of a variety of statistical techniques such as Principal Component Analysis (PCA)/Partial Least Squares (PLS) [7,27–30], Generalized Likelihood Ratio (GLR) [23,27,31–34], neural networks [7,23,27,35–39], and autoencoders [40,41] to name a few. A recent study [27] showed that an interval PCA-based GLR algorithm for fault detection and classification was able to outperform a number of neural network-based approaches. This demonstrates that the GLR technique when appropriately designed and integrated can even outperform certain sophisticated neural network-based algorithms, thus encouraging its implementation and use in our current work.

Schemes that integrate model-based and data-driven methods, however, make use of quantitative models and model-based fault diagnosis techniques augmented with statistical classification techniques to tackle any noises and uncertainties prevalent in the system [7,9]. In general, such schemes use the disturbance decoupled model-based fault diagnosis schemes for residual generation, where the residual is now a statistical time series with a known probability distribution (calculated analytically or empirically) when in control. Statistical tools are then used downstream to analyze the residual for changes in mean/ variance to detect faults (out of control). In linear systems, prior research on such integrated schemes has focused on integrating Kalman filters with maximum likelihood estimation [7,9,26,42,43], parity space methods with temporal and spatial whitening of the residuals [42], and Markov models with Monte Carlo estimation [44]. For nonlinear systems however, methods integrating model-based methods with statistical tools have been limited.

The goal of this work is to design an integrated model-based and statistical fault diagnosis scheme for nonlinear processes with random disturbances (sensor noises and plant model mismatch) that cannot be decoupled by deterministic methods in [9, 21,22]. The model-based part focuses on building observers for residual generation from the point of view of observer error linearization [21,22,45]. The advantages of model-based schemes stem from the fact that it is closely connected to physical variables and easily lends itself to fault isolation. The presence of noises, however, corrupt the residuals generated by these schemes, in turn leading to false alarms. In the first part of this work, the effect of sensor noises on the residuals generated by model-based schemes will be studied analytically. Following this, GLR is applied on the residuals to distinguish between normal and faulty operating conditions on the basis of exceeding a GLR-based threshold that are functions of the residual data. Finally, the approach will be extended to process noises in an entirely numerical manner.

In the next section, a review of disturbance decoupled detection of a single fault in the absence of noises will be presented. Then, the problem of fault detection in the presence of Gaussian sensor noises will be tackled and the probability distributions of the residuals will be derived. Following this, the hypothesis testing method, GLR, will be reviewed and an algorithm using GLR for detecting faults in the presence of sensor noises will be presented. Finally, the approach will be extended to process noises, where a numerical approach will be used to obtain the fault free distribution before using GLR to evaluate the residual.

2. Disturbance decoupled fault detection in the absence of noises

Consider a nonlinear process described by:

$$\begin{aligned} x(k+1) &= F(x(k), W(k), f(k)) \\ y(k) &= H(x(k)) + G(x(k))\,W(k) + E(x(k))\,f(k) \end{aligned} \tag{2.1}$$

where $x(k)\epsilon\mathbb{R}^n$ denotes the vector of states, $y(k)\epsilon\mathbb{R}^p$ denotes the vector of measured outputs. $f(k)\epsilon\mathbb{R}$ and $W(k)\epsilon\mathbb{R}^m$ are the fault and the disturbances/uncertainties respectively (system inputs) and $E(x), F(x), G(x), H(x), J(x), K(x)$ are smooth functions.

We will study the problem of disturbance-decoupled fault detection on the basis of calculating a quantity $\hat{z}$ called the residual, whose evolution differs depending on the presence (i.e. when $f(k)\neq 0$) or absence of faults (i.e. when $f(k)=0$), and is unaffected by the disturbances W.

More specifically, this work will study the design of a **linear** functional observer, called the residual generator, of the form

$$\begin{aligned} \hat{\xi}(k+1) &= A\hat{\xi}(k) + By(k) \\ \hat{z}(k) &= C\hat{\xi}(k) + Dy(k) \end{aligned} \tag{2.2}$$

where $\hat{\xi}(k)\,\epsilon\mathbb{R}^{\nu}$ is the observer state and $\hat{z}(k)\,\epsilon\mathbb{R}$ is the scalar residual. It is desired that the response of the residual $\hat{z}$ in the series connection of (2.1) followed by (2.2)

$$\begin{aligned} \begin{bmatrix} x(k+1) \\ \hat{\xi}(k+1) \end{bmatrix} &= \begin{bmatrix} F(x(k), W(k), f(k)) \\ A\hat{\xi}(k) + B[H(x(k)) + G(x(k))\,W(k) + E(x(k))\,f(k)] \end{bmatrix} \\ \hat{z}(k) &= \left[C\hat{\xi}(k) + D[H(x(k)) + G(x(k))\,W(k) + E(x(k))\,f(k)]\right] \end{aligned} \tag{2.3}$$

has the following properties:

(a) $\hat{z}(k)$ asymptotically approaches zero when f is identically zero
(b) $\hat{z}(k)$ is unaffected by the disturbances W
(c) $\hat{z}(k)$ is affected by the fault f.

In other words, for any initial conditions $\begin{bmatrix} x(0) \\ \hat{\xi}(0) \end{bmatrix}$ and any disturbances W(k),

$\lim_{k\to\infty} \hat{z}(k) = 0$ iff $k = 0$
$\lim_{k\to\infty} \hat{z}(k) \neq 0$ iff $k \neq 0$

The responsiveness of $\hat{z}$ to faults and insensitivity to disturbances ensures fault detection while precluding the possibility of false alarms.

For (a) to hold in the absence of disturbances, there must exist a differentiable map T (x) from $\mathbb{R}^n$ to $\mathbb{R}^\nu$ such that:

$$T(F_* (x)) = AT (x) + BH (x) \tag{2.4}$$

$$0 = CT (x) + DH (x) \tag{2.5}$$

where $F_* (x) = F(x, 0, 0)$. If conditions (2.4) and (2.5) are satisfied, the functional observer's error dynamics, with error defined as $e (k) = \hat{\xi} (k) - T(x(k))$, is as follows:

$$\begin{aligned} e(k+1) &= Ae(k) + B [G (x(k)) W (k) + E (x(k)) f(k)] \\ &\quad - [T (F (x (k) , W (k) , f(k))) - T (F_*(x(k)))] \\ \hat{z} (k) &= Ce(k) + D [G (x(k)) W (k) + E (x(k)) f(k)] \end{aligned} \tag{2.6}$$

It should be noted here that the zero-input ($W = 0, f = 0$) dynamics of (2.6) becomes $e(k+1) = Ae(k)$ i.e. is exactly linear and moreover, if the matrix A has eigenvalues in the unit disc, the zero-input response is

$$e(k) = A^k e(0) \to 0$$

$$\hat{z} (k) = CA^k e(0) \to 0$$

which means that the residual $\hat{z} (k)$ asymptotically approaches zero. It is possible to derive the following necessary and sufficient condition that Eqs. (2.4) and (2.5) can hold if and only if there exist $\beta_0, \beta_1, \ldots, \beta_\nu \in R^p$ such that [21,45]:

$$\beta_0 HF_*^\nu(x) + \beta_1 HF_*^{\nu-1}(x) + \cdots + \beta_{\nu-1} HF_*(x) + \beta_\nu H (x) = 0 \tag{2.7}$$

where in the equation above we have used the notation $F^j (x) = \underbrace{F \circ F \ldots F \circ F(x)}_{j \text{ times}}$ and $HF (x) = (H \circ F)(x)$ with A, B, C, D given by

$$\begin{aligned} A &= \begin{bmatrix} 0 & 0 & \cdots & 0 & -\alpha_\nu \\ 1 & 0 & \cdots & 0 & -\alpha_{\nu-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & -\alpha_2 \\ 0 & \cdots & 0 & 1 & -\alpha_1 \end{bmatrix}, \\ B &= \begin{bmatrix} \beta_\nu - \alpha_\nu \beta_0 \\ \beta_{\nu-1} - \alpha_{\nu-1}\beta_0 \\ \beta_{\nu-2} - \alpha_{\nu-2}\beta_0 \\ \vdots \\ \beta_1 - \alpha_1 \beta_0 \end{bmatrix}, C = [0\ 0\ \cdots\ 0\ 1], D = \beta_0 \end{aligned} \tag{2.8}$$

and

$$T (x) = \begin{bmatrix} -\beta_0 HF_*{}^{\nu-1}(x) - \cdots - \beta_{\nu-2} HF_*(x) - \beta_{\nu-1} H (x) \\ \vdots \\ -\beta_0 HF_*(x) - \beta_1 H (x) \\ -\beta_0 H (x) \end{bmatrix} \tag{2.9}$$

The second requirement for the functional observer is that the residual $\hat{z}$ must remain completely unaffected by any disturbances W (k) present in the system. From (2.6), disturbance decoupling can be achieved if

$$\begin{aligned} &\frac{\partial \Omega (x, W, f)}{\partial W} = 0 \ \forall W \\ &DG (x) = 0 \end{aligned} \tag{2.10}$$

where $\Omega (x, W, f) = B [G (x) W + E (x) f] - [T (F(x, W, f)) - T (F_*(x))]$. The third and final requirement for the functional observer is that the residual must be affected by the input f, so that the fault can be detected by monitoring the residual. For this to be possible the following equations must hold,

$$\begin{bmatrix} \Omega (x, W, f) \\ DE (x) \end{bmatrix} \neq 0 \ \forall f \neq 0 \tag{2.11}$$

Thus, to construct a functional observer (2.2) for fault detection in the process (2.1):

(i) Find a set of constant row vectors $\beta_0, \beta_1, \ldots, \beta_\nu$ that satisfy (2.7)
(ii) Construct T(x) and (A, B, C, D) using (2.8) and (2.9) respectively
(iii) Substitute in (2.10) and (2.11) to see if the disturbance decoupling and fault detectability condition hold. If they hold, the residual generator matrices (A,B,C,D) are given by (2.8), else, look for a different set of vectors in step (i).

Remark 2.1. The conditions (2.7), (2.10) and (2.11) are mathematically restrictive, however they can be applied to a variety of chemical engineering systems. This is because in many chemical processes, dynamic models are composed of conservation equations and inventory rate equations of the form: (Accumulation) = (In) – (Out) + (Generation), with the nonlinearities appearing only in the generation terms. This makes them amenable to the design conditions (2.7), (2.10) and (2.11).

Remark 2.2. In the presence of only sensor disturbances, the disturbance decoupling condition (2.10) becomes

$$\begin{bmatrix} \beta_\nu \\ \beta_{\nu-1} \\ \beta_{\nu-2} \\ \vdots \\ \beta_1 \\ \beta_0 \end{bmatrix} G (x) = 0.$$

Remark 2.3. The row vectors $\beta_0, \beta_1, \ldots, \beta_\nu$ provide information about the measurements that are being used in the residual generator. For example, if the jth element of all of these vectors happens to be 0, it means that the measurement y_j is not used for fault detection since both B and D will have their jth column identically zero.

Remark 2.4. Although dynamic systems are continuous systems, the focus of this work is on designing observers for discrete-time nonlinear systems because fault diagnosis tools use sampled data [8]. Alternatively, one could design a functional observer for the original continuous-time system using methods developed [21], discretize the resulting observer and follow the steps in Sections 3–5.

The disturbance decoupled fault detection approach reviewed in this section can be directly extended to build a fault isolation scheme. For example, in a system with n_f faults, n_f functional observers can be designed, one for each fault, where for functional observer i, fault i is to be detected and all the other $n_f - 1$ faults are disturbances that are decoupled (see Fig. 1).

3. Disturbance decoupled fault detection in the presence of sensor noises

Consider a nonlinear process described by:

$$\begin{aligned} x (k+1) &= F (x (k) , W (k) , f (k)) \\ y(k) &= H (x (k)) + G (x(k)) W (k) + E (x(k)) f(k) + \eta (k) \end{aligned} \tag{3.1}$$

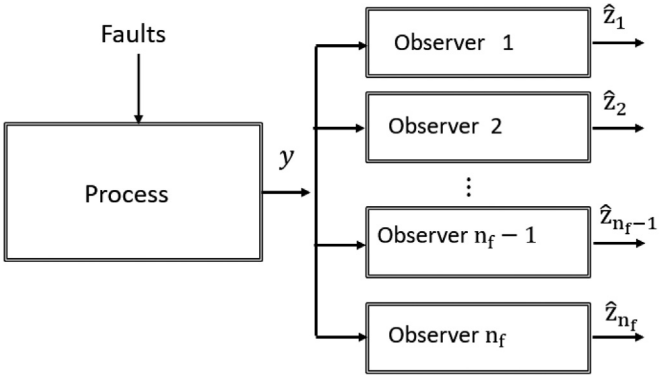

Fig. 1. Fault Isolation scheme, based on a set of observers, one for each fault.

where $x(k) \in \mathbb{R}^n$ denotes the vector of states, $y(k) \in \mathbb{R}^p$ denotes the vector of measured outputs. $f(k) \in \mathbb{R}$ and $W(k) \in \mathbb{R}^m$ are the fault and the disturbances/uncertainties respectively and E (x) , F (x) , G (x) , H (x) , J (x) are smooth functions and η is the Gaussian noise vector with mean 0 and covariance Σ_η^2.

Suppose that we have already built a linear functional observer of the form (2.2) for the noise-free system (2.1). We would like to use the same functional observer, but now driven by the noisy measurement signal y(k) from (3.1):

$$\begin{aligned} \hat{\xi}_\eta(k+1) &= A\hat{\xi}_\eta(k) + By(k) \\ \hat{z}_\eta(k) &= C\hat{\xi}_\eta(k) + Dy(k) \end{aligned} \tag{3.2}$$

where $\hat{\xi}_\eta(k) \in \mathbb{R}^\nu$ and $\hat{z}_\eta(k) \in \mathbb{R}$, and A, B, C and D are given by (2.8), the objective being to detect faults in (3.1). The functional observer's error dynamics with error defined as $e_\eta(k) = \hat{\xi}_\eta(k) - T(x(k))$, with T(x) given by (2.9), is

$$\begin{aligned} e_\eta(k+1) &= Ae_\eta(k) + B[G(x(k))W(k) + E(x(k))f(k) + \eta(k)] \\ &\quad - [T(F(x(k), W(k), f(k))) - T(F_*(x(k)))] \\ \hat{z}_\eta(k) &= Ce_\eta(k) + D[G(x(k))W(k) + E(x(k))f(k) + \eta(k)] \end{aligned} \tag{3.3}$$

We would like to analyze the effect of noise vector η on the observer output $\hat{z}_\eta$. To this end, Eq. (2.6) is subtracted from (3.3) and as a result the following holds:

$$\begin{aligned} e_\eta(k+1) - e(k+1) &= A\left(e_\eta(k) - e(k)\right) + B\eta(k) \\ \hat{z}_\eta(k) - z(k) &= C\left(e_\eta(k) - e(k)\right) + D\eta(k) \end{aligned} \tag{3.4}$$

Now one can express $e_\eta(k) - e(k)$ in terms of the error l time steps before as

$$e_\eta(k) - e(k) = A^l\left(e_\eta(k-l) - e(k-l)\right) + \sum_{i=1}^{l} A^{i-1}B\eta(k-i)$$

If the eigenvalues of A lie in the unit disc, then for large l, $e_\eta(k) - e(k) = \sum_{i=1}^{l} A^{i-1}B\eta(k-i)$ and hence

$$\hat{z}_\eta(k) - z(k) = C\sum_{i=1}^{l} A^{i-1}B\eta(k-i) + D\eta(k) \tag{3.5}$$

From the above equation, one can see that $\hat{z}_\eta(k) - z(k)$ follows a Gaussian distribution with mean $\mu_0 = 0$ and variance

$$\sigma_0^2 = D\left(\Sigma_\eta^2\right)D^T + \hat{h}\hat{\Sigma}_\eta^2\hat{h}^T \tag{3.6}$$

where $\hat{h} = \left[CB\ CAB\ CA^2B \ldots CA^{l-1}B\right]$ and $\hat{\Sigma}_\eta^2$ is the covariance matrix of $\begin{bmatrix} \eta(k-1) \\ \eta(k-2) \\ \vdots \\ \eta(k-l) \end{bmatrix}$ [46].

The foregoing analysis leads to the conclusion that $\hat{z}_\eta(k)$ for large k is a Gaussian distribution with mean equal to the output of the observer (2.2) for the deterministic system (2.1) and variance given by (3.6). Therefore, in the absence of faults (null hypothesis H_0), $\hat{z}_\eta(k)$ follows a Gaussian distribution $N\left(\mu_0 = 0, \sigma_0^2 = D\left(\Sigma_\eta^2\right)D^T + \hat{h}\hat{\Sigma}_\eta\hat{h}^T\right)$. This distribution gives the sensitivity of the residual to sensor noises and to achieve robust fault detection one must be able to accurately distinguish deviations in the residual due to a fault from deviations due to noises. This leads to the next section, where an algorithm using GLR will be presented to detect faults on the basis of changes in the mean of the residual.

4. Generalized likelihood ratio (GLR)

We know from the previous section that the fault free distribution is $N\left(\mu_0, \sigma_0^2\right)$. But, suppose the residual generated from the functional observer designed in the previous section has a $N\left(\mu_1, \sigma_1^2\right)$ distribution. The goal is to detect any fault that produces a shift in mean away from μ_0. To this end, in this section a well-known hypothesis testing technique in statistics known as Generalized Likelihood Ratio [43], that has had significant applications in data driven fault diagnosis of engineering systems [23,31,32], is reviewed.

Say we have k observations, $r_1, \ldots, r_k$, the null hypothesis H_0 is that there has been no mean shift ($\mu = \mu_0$) and under this hypothesis the likelihood function at sample k can be represented as:

$$\begin{aligned} L(\infty, \mu_0 | r_1, r_2, \ldots, r_k) &= (2\pi)^{-\frac{k}{2}}\sigma_0^{-\frac{k}{2}} \\ &\quad \times \exp\left(-\tfrac{1}{2\sigma_0^2}\left(\sum_{i=1}^{k}(r_i - \mu_0)^2\right)\right) \end{aligned} \tag{4.1}$$

For the alternate hypothesis H_1 that a mean shift of some value $\mu_1 \neq \mu_0$ has occurred at some time τ^* between samples τ and $\tau + 1$ where $\tau < k$, the likelihood function at sample k is [23,43]

$$\begin{aligned} L(\tau, \mu_1 | r_1, r_2, \ldots, r_k) &= (2\pi)^{-\frac{k}{2}}\sigma_0^{-\frac{k}{2}} \\ &\quad \times \exp\left(-\tfrac{1}{2\sigma_0^2}\left(\sum_{i=1}^{\tau}(r_i - \mu_0)^2 + \sum_{i=\tau+1}^{k}(r_i - \mu_1)^2\right)\right) \end{aligned} \tag{4.2}$$

If there has been a shift to an unknown μ_1 between samples τ and $\tau + 1$, then the maximum likelihood estimator of μ_1 is

$$\hat{\mu}_{1,\tau,k} = \tfrac{1}{k-\tau}\sum_{i=\tau+1}^{k} r_i \tag{4.3}$$

Then a log likelihood-ratio statistic for determining whether there has in fact been a mean shift is

$$GLR_k = \ln \frac{\max_{0\le\tau<k,-\infty<\mu_1<\infty} L(\tau,\mu_1|r_1,r_2,\ldots,r_k)}{L(\infty,\mu_0|r_1,r_2,\ldots,r_k)} \quad (4.4)$$

$$= \max_{0\le\tau\le k} \frac{(\hat{\mu}_{1,\tau,k}-\mu_0)}{\sigma_0^2} \sum_{i=\tau+1}^{k} \left[(r_i-\mu_0) - \frac{1}{2}(\hat{\mu}_{1,\tau,k}-\mu_0)\right]$$

From the maximum likelihood estimate $\hat{\mu}_{1,\tau,k}$ the above equation reduces to

$$GLR_k = \max_{0\le\tau<k} \frac{k-\tau}{2\sigma_0^2}(\hat{\mu}_{1,\tau,k}-\mu_0)^2 \quad (4.5)$$

We will use the above equation to calculate the GLR statistic at any time k given the data from [0, k].

The following maximum likelihood estimation algorithm is presented to compute the threshold for the GLR statistic for a residual dataset with a given mean and variance of the fault free distribution and a desired false alarm rate.

Maximum Likelihood Estimation Algorithm:

a. Obtaining the threshold for the GLR Statistic.

1. Generate random normal distribution of a sufficiently large size (e.g., 10000 observations) using known fault-free mean (μ_0) and variance (σ_0^2).
2. Compute GLR statistic using the random normal distribution that was generated:
   To monitor mean
   $$GLR_k = \max_{0\le\tau<k} \frac{(k-\tau)}{2\sigma_0^2}(\hat{\mu}_{1,\tau,k}-\mu_0)^2$$

Where, $\hat{\mu}_{1,\tau,k}$ is the maximum likelihood estimates (MLEs) of the mean computed utilized the available data. k and τ correspond to the current time instant, and the position in the time window that provides the maximum detection rate for a fixed false alarm rate, respectively.

3. Use the computed generalized likelihood ratio statistic to compute its empirical distribution.
4. Use the desired confidence interval (α), e.g., 99% to obtain the fault detection threshold by extracting the corresponding percentile from the computed cumulative empirical distribution.

   $$GLR_{threshold,\alpha} = ecdf_\alpha\,(GLR_k\,(mean))$$

Note: In this study we are interested in the detecting faults that bring about a change in the mean of the residual. However, one could also use GLR to detect faults that cause changes in the variance of the residual. There exists significant literature on this topic and the interested reader is referred to these papers [32,47].

b. Using the threshold to detect faults

1. Compute the GLR statistic online using the available formula:
   a. To monitor mean:
   $$GLR_{test,k} = \max_{0\le\tau<k} \frac{(k-\tau)}{2\sigma_0^2}(\hat{\mu}_{1,\tau,k}-\mu_0)^2$$
2. Declare fault if: $GLR_{test,k} > GLR_{threshold,\alpha}$

Remark 4.1. In theory, all observations in the interval [0,k] are required to calculate the GLR statistic for sample k. This would require the storage and use of all past data, where at each sampling point k the maximum value in (4.5) over $0 < \tau < k$ needs to be computed. This becomes cumbersome as k increases in size. Thus, one could adopt a modification of GLR_k in which the maximum is taken over a window of past l samples [43]

$$GLR_{l,k} = \max_{k-l\le\tau<k} \frac{(k-\tau)}{2\sigma_0^2}(\hat{\mu}_{1,\tau,k}-\mu_0)^2$$

As a rule of thumb, one could use $l = 400$ as Reynolds and coworkers [43] obtained comparable results are obtained for $l = 400$ and $l = 10000$ for all fault magnitudes.

5. Proposed methodology for fault detection in the presence of sensor noises

In Sections 2 and 3, a model-based approach to generate residuals using functional observers was discussed and the behavior of these residuals in the presence of sensor noises was studied. In the previous section, a well-known statistical hypothesis testing method, Generalized Likelihood Ratio, capable of detecting changes in the mean of the residual, was reviewed. A specific algorithm was presented, including the steps for calculating the GLR threshold and online computation of the GLR statistic.

The following steps should be followed to integrate the model-based and statistical methods presented in the previous sections to detect faults in any physical process affected by sensor noises.

5.1. Offline calculations

I. Using the steps of Section 2, design a functional observer for the deterministic physical model
II. Analytically calculate the distribution of the residual in the presence of sensor noises (using Eq. (3.6))
III. Use the analytical distribution to obtain the GLR threshold (Part a in the algorithm in Section 4)

In the above steps, it is essential to account for the effect of observer eigenvalues on the sensitivity of the residual to faults and noises. The observer eigenvalues are tunable parameters and must be tuned to maximize fault sensitivity and minimize noise sensitivity. This analysis will be used to evaluate the robustness of the fault detection scheme for a variety of possible fault sizes.

5.2. Online calculations

I. Feed the output from the noisy sensor to the designed functional observer and generate the residuals
II. Use the residual data to calculate and track the GLR statistic (Part b in the algorithm in Section 4)
III. Flag faults if GLR threshold is violated (Part b in the algorithm in Section 4).

The steps above constitute a methodology to compute thresholds using residual data and automate fault detection in the presence of sensor noises.

5.3. Evaluation metrics

The fault detection methodology will be evaluated utilizing three metrics

a. False alarm rate (%)- The percentage of fault free observations incorrectly classified as faulty. This metric should be as low as possible ($\le$ 1%) for a good fault detection scheme.
b. Missed detection rate (%) - The percentage of faulty observations incorrectly classified as fault-free. An accurate fault detection scheme should have low missed detection rates.
c. Out-of-control Average Run Length (ARL_1)- Number of observations required since fault occurrence to detect faults, corresponding to the speed of detection. A low ARL_1 means the detection is fast.

The fault detection methodology proposed in this section will be illustrated and tested through a non-isothermal chemical reactor case study in the next section.

6. Case study — fault detection in a non-isothermal CSTR in the presence of sensor noises

Consider a non-isothermal Continuous Stirred Tank Reactor (CSTR) undergoing N-alkyl pyridine oxidation with hydrogen peroxide. The product of the reaction, Alkyl Pyridine N-Oxides is an important intermediate in several important reactions in the pharmaceutical industry, including the production of anti-ulcer and anti-inflammatory drugs [48].

It is assumed the reactor is well-mixed and has constant volume with an inlet stream containing N-methyl pyridine (A) + catalyst Z (assumed to be fully dissolved) and hydrogen peroxide (B). The catalyst is assumed to be completely dissolved in the pyridine stream and its concentration is assumed to be constant in the reactor [48] The dynamics of the reactor [48] is described by:

$$\begin{aligned}
\frac{dC_A}{dt} &= \frac{F}{V}\left(C_{A,in} - C_A\right) - R\left(C_A, C_B, \theta, w\right)\\
\frac{dC_B}{dt} &= \frac{F}{V}\left(C_{B,in} - C_B\right) - R\left(C_A, C_B, \theta, w\right)\\
\frac{d\theta}{dt} &= \frac{F}{V}\left(\theta_{in} - \theta\right) - \frac{US_A}{\rho c_p V}\left(\theta - \theta_j\right) + \frac{-\Delta H_R}{\rho c_p} R\left(C_A, C_B, \theta, w\right)\\
\frac{d\theta_j}{dt} &= \frac{F_j}{V_j}\left(\theta_{j,in} + f_2(t) - \theta_j\right) + \frac{US_A}{\rho_j c_{p_j} V_j}\left(\theta - \theta_j\right)\\
y_1 &= C_A + \eta_1(t) + f_1(t)\\
y_2 &= \theta + \eta_2(t)\\
y_3 &= \theta_j + \eta_3(t)
\end{aligned} \tag{6.1}$$

The state vector $X = \left[C_A, C_B, \theta, \theta_j\right]$ consists of N-methyl pyridine concentration, hydrogen peroxide concentration, reactor temperature and outlet coolant temperature. The noise vector H= $[\eta_1, \eta_2, \eta_3]$ has zero mean and variance

$$\Sigma_\eta^2 = \begin{bmatrix} 0.001\left(\frac{mol}{L}\right)^2 & 0 & 0\\ 0 & 0.01K^2 & 0\\ 0 & 0 & 0.01\,K^2 \end{bmatrix}.$$

$R\left(C_A, C_B, \theta, w\right) = \frac{(A_1+w)e^{-\frac{E_1}{\theta}} A_2 e^{-\frac{E_2}{\theta}} C_A C_B Z}{1+A_2 e^{-\frac{E_2}{\theta}} C_B} + A_3 e^{-\frac{E_3}{\theta}} C_A C_B$ is the reaction rate. Parameters A_1, A_2, A_3 are the pre-exponential factors in the reaction rate and w is the uncertainty/disturbance in A_1. $C_{A,in}, C_{B,in}, \theta_{in}, \theta_{j,in}$ represent the inlet values of the state values. F and F_j are the inlet feeds and coolant flowrates, respectively. V and V_j are the reactor volume and cooling jacket volume, respectively. ΔH_R is the enthalpy of the reaction. ρ, c_p and ρ_j, c_{p_j} are the densities, heat capacities of the reactor contents and cooling fluid, respectively. $f_1(t)$ represents a possible concentration sensor fault and $f_2(t)$ represents a possible fault that could affect the inlet temperature of the coolant.

The parameter values are as follows, $C_{A,in} = 2\frac{mol}{L}$, $C_{B,in} = 1.5\frac{mol}{L}$, $\theta_{in} = 373K$, $\theta_{j,in} = 300$ K, $F = 0.1\frac{L}{min}$, $F_j = 1\frac{L}{min}$, $V = 0.5$ L, $V_j = 3 \times 10^{-2}$ L, $A_1 = e^{8.08}$ $Lmol^{-1}s^{-1}$, $A_2 = e^{28.12}$ $Lmol^{-1}s^{-1}$, $A_3 = e^{25.12}Lmol^{-1}$. $\Delta H_R = -160\frac{kj}{mol}$, $\rho = 1200\frac{g}{L}$, $\rho_j = 1000\ \frac{g}{L}$, $c_{p_j} = 3\frac{J}{gK}$. $c_p = 3.4\ \frac{J}{gK}$ U=0.942 $\frac{W}{m^2K}$, $S_A = 1\,m^2$, Z = 0.0021 $\frac{mol}{L}$, $E_1 = 3952$ K, $E_2 = 7927$ K, $E_3 = 12989$ K, and the initial conditions are $C_A = 0\frac{mol}{L}$, $C_B = 0\frac{mol}{L}$, T = 300K, T_j = 300K.

The model equations are discretized using Euler's method with sampling period $\delta_t = 0.5$ s. This is easy to implement and preserves the nonlinearities of the original continuous-time system. The discretized equations are

$$\begin{aligned}
C_A(k+1) &= C_A(k)\\
&\quad + \delta_t\left(\frac{F}{V}\left(C_{A,in} - C_A(k)\right) - R\left(C_A(k), C_B(k), \theta(k), w(k)\right)\right)\\
C_B(k+1) &= C_B(k)\\
&\quad + \delta_t\left(\frac{F}{V}\left(C_{B,in} - C_B\right) - R\left(C_A(k), C_B(k), \theta(k), w(k)\right)\right)\\
\theta(k+1) &= \theta(k)\\
&\quad + \delta_t\left(\frac{F}{V}\left(\theta_{in} - \theta(k)\right) - \frac{US_A}{\rho c_p V}\left(\theta(k) - \theta_j(k)\right)\right) +\\
&\quad \frac{\delta_t(-\Delta H_R)}{\rho c_p}\left(R\left(C_A(k), C_B(k), \theta(k), w(k)\right)\right)\\
\theta_j(k+1) &= \theta_j(k)\\
&\quad + \delta_t\left(\frac{F_j}{V_j}\left(\theta_{j,in}(k) + f_2(k) - \theta_j(k)\right) + \frac{US_A}{\rho_j c_{p_j} V_j}\left(\theta(k) - \theta_j(k)\right)\right)\\
y_1 &= C_A(k) + f_1(k) + \eta_1(k)\\
y_2 &= \theta(k) + \eta_2(k)\\
y_3 &= \theta_j(k) + \eta_3(k)
\end{aligned} \tag{6.2}$$

The equations are converted to deviation form to remove the constants corresponding to the inlet conditions with $C_A' = C_A - C_{A,ref}$, $\theta' = \theta - \theta_{ref}$ and $\theta_j' = \theta_j - \theta_{j,\,ref}$ (where the subscript ref denotes the reference steady state value of the system in the absence of noises, uncertainties, and faults) representing the deviation variables. The outputs in deviation form are $y_1' = y_1 - C_{A,ref}$, $y_2' = y_2 - \theta_{ref}$ and $y_3' = y_3 - \theta_{j,ref}$.

The following two functional observers are built using the steps presented in Section 2.

Functional Observer 1: Detection of the analytical sensor fault f_1 while considering f_2 as an additional disturbance.

$$\begin{aligned}
\beta_0 &= \left[1, \frac{\rho c_p}{(-\Delta H)}, 0\right] \\
\beta_1 &= \left[\left(\frac{\delta_t F}{V} - 1\right), -\frac{\rho c_p}{(-\Delta H)}\left(1 - \frac{\delta_t F}{V} - \frac{US_A\delta_t}{\rho c_p V}\right),\right.\\
&\quad \left. - \frac{US_A}{(-\Delta H)V}\delta_t\right]
\end{aligned} \tag{6.3}$$

The resulting functional observer (with $\alpha_1 = -0.99$) is

$$\begin{aligned}
\hat{\xi}_\eta(k+1) &= -\alpha_1\hat{\xi}_\eta(k) + \left[\frac{\delta_t F}{V} - 1 - \alpha_1\right] y_1'(k)\\
&\quad - \frac{\rho c_p}{(-\Delta H)}\left(1 + \alpha_1 - \frac{\delta_t F}{V} - \frac{US_A\delta_t}{\rho c_p V}\right) y_2'(k) - \frac{US_A\delta_t}{(-\Delta H)V} y_3'(k)\\
\hat{z}_\eta(k) &= \hat{\xi}_\eta(k) + y_1'(k) + \frac{\rho c_p}{(-\Delta H)} y_2'(k)
\end{aligned} \tag{6.4}$$

At large times ($k \to \infty$) the residual follows a Gaussian distribution with mean 0 and variance=0.001 in the absence of faults (calculated analytically using 3.6).

Functional Observer 2: Detection of inlet coolant temperature fault f_2 while considering f_1 as an additional disturbance.

$$\begin{aligned}
\beta_0 &= [0, 0, 1] \\
\beta_1 &= \left[0, -\frac{US_A\delta_t}{\rho_j c_{p_j} V_j}, -\left(1 - \frac{\delta_t F_j}{V_j} - \frac{US_A\delta_t}{\rho_j cp_j V_j}\right)\right]
\end{aligned} \tag{6.5}$$

The resulting functional observer with $\alpha_1 = -0.1$. is:

$$\begin{aligned}
\hat{\xi}_\eta(k+1) &= -\alpha_1\hat{\xi}_\eta(k) - \frac{US_A\delta_t}{\rho_j cp_j V_j} y_2'(k)\\
&\quad - \left[1 + \alpha_1 - \frac{\delta_t F_j}{V_j} - \frac{US_A\delta_t}{\rho_j cp_j V_j}\right] y_3'(k)\\
\hat{z}_\eta(k) &= \hat{\xi}_\eta(k) + y_3'(k)
\end{aligned} \tag{6.6}$$

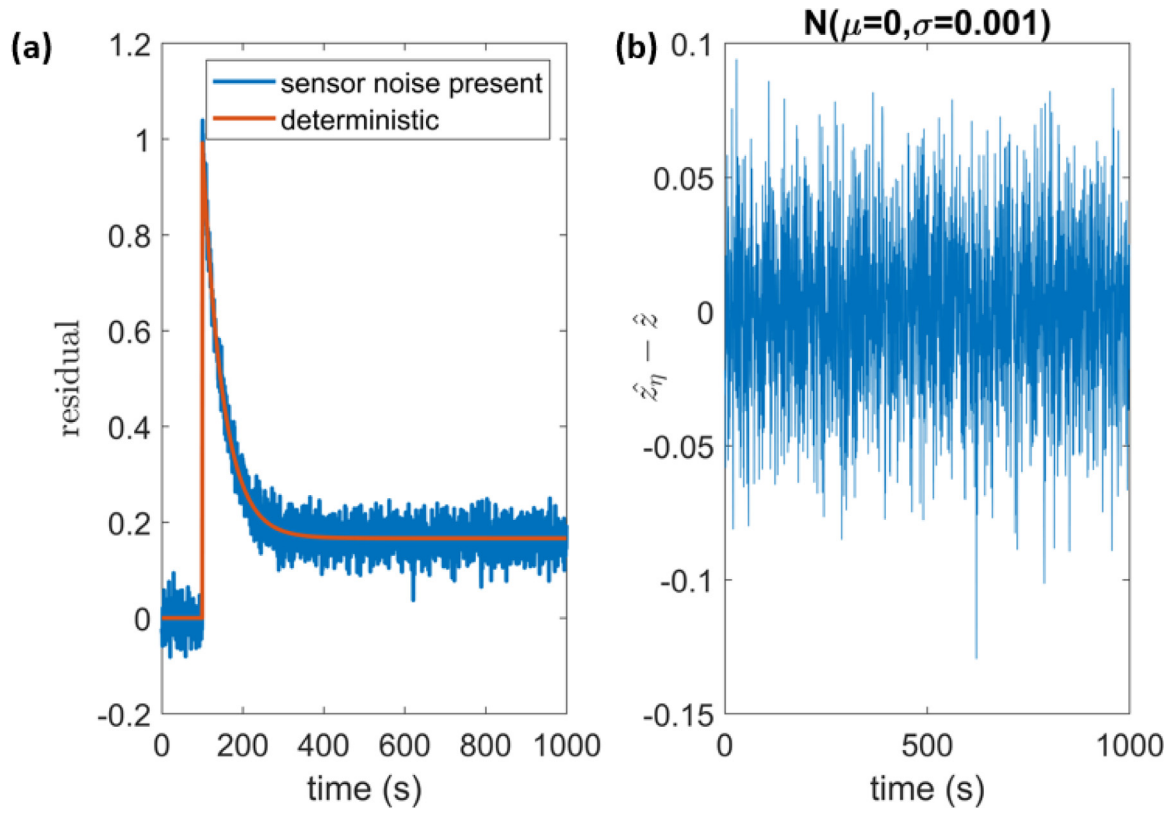

Fig. 2. (a) Residual vs time (seconds) for functional observer 1 (b) Difference between functional observer output in deterministic and stochastic case ($\hat{\mathbf{z}}_\eta - \hat{\mathbf{z}}$) corresponding to a N(0,0.001) distribution (numerically observed).

At large times ($k \rightarrow \infty$) the residual follows a Gaussian distribution with mean 0 and variance =0.01385 in the absence of faults (calculated analytically using 3.6).

The fault scenario is as follows $f_1(t) = \begin{cases} 0, & t < 100s \\ 1, & t \geq 100s, \end{cases}$ $f_2(t) = \begin{cases} 0, & t < 150s \\ 10, & t \geq 150s \end{cases}$, $w(t) = 10^5$. A false alarm rate of less than 1% is desired. It is to be noted when evaluating the fault detection scheme that the out-of-control average run length (ARL_1) is given in terms of the sampling rate. For example, $ARL_1 = 10$ would indicate 10 observations are required, each spaced 0.5s apart, for fault detection.

The proposed fault detection framework is implemented in MATLAB (MathWorks R2019b). The computation is performed on a Dell OptiPlex 9020 with Intel® Core™ i7-4770 CPU @3.40 GHz and 16GB of total memory. The average computation time of the GLR statistic for each iteration is 0.01 s, implying that the algorithm can be implemented in real-time to enable efficient online fault detection. It is important to note that at each iteration the GLR statistic evaluates a window size of 400 as recommended by Reynolds [43].

The residuals generated from the functional observer 1 in the presence and absence (deterministic) of noises is plotted in Fig. 2. Analytically, using Eqs. (3.5) and (3.6), one can show that the difference between the observer output in the stochastic and deterministic cases i.e, $\hat{z}_\eta - \hat{z}$ is a Gaussian distribution with 0 mean and variance 0.001(Note: this is also confirmed numerically see Fig. 2b). In Fig. 3, the residuals generated from functional observer 2 in the presence and absence of noises is plotted. $\hat{z}_\eta - \hat{z}$ in this case is supposed (using Eqs. (3.5) and (3.6)) to be a Gaussian distribution with 0 mean and a variance 0.01385. The plot on Fig. 3b confirms this analytical calculation.

Table 1
Fault detection metrics in the presence of sensor noises.

	Functional observer 1 (sensor fault)	Functional observer 2 (coolant fault)
Missed Detection Rate (%)	0.00	0.06
False Alarm Rate (%)	0.00	0.00
ARL_1	1.00	2.00

6.1. GLR calculation

The fault free distributions of the output of the two functional observers are used to develop thresholds of the GLR statistic (Steps i and ii in algorithm in Section 4). The thresholds for residuals from functional observer 1 and functional observer 2 are 2.193 and 2.203, respectively. The time series evolution of the GLR statistic for the two residuals is plotted in Figs. 4 and 5, respectively. In both cases when the respective fault occurs a sharp increase in the GLR statistic is observed, which then settles to a value higher than the threshold for fault free data in turn facilitating rapid fault detection rate with near 0% missed detection rate.

The fault detection metrics are summarized in Table 1. For sensor and coolant faults, the missed detection rate (percentage of faulty observations flagged fault free) is 0% and 0.06%, respectively, the false alarm rate (percentage of fault free observations flagged faulty) is 0% in both cases. This highlights the ability of the fault diagnosis scheme to accurately detect and isolate faults of interest in the presence of sensor noises. Furthermore, the detection is rapid since the out of control run length average run length (ARL_1) for sensor fault is 1, meaning only 1 observation is required since fault occurrence to detect faults, and for coolant fault the ARL is 2.

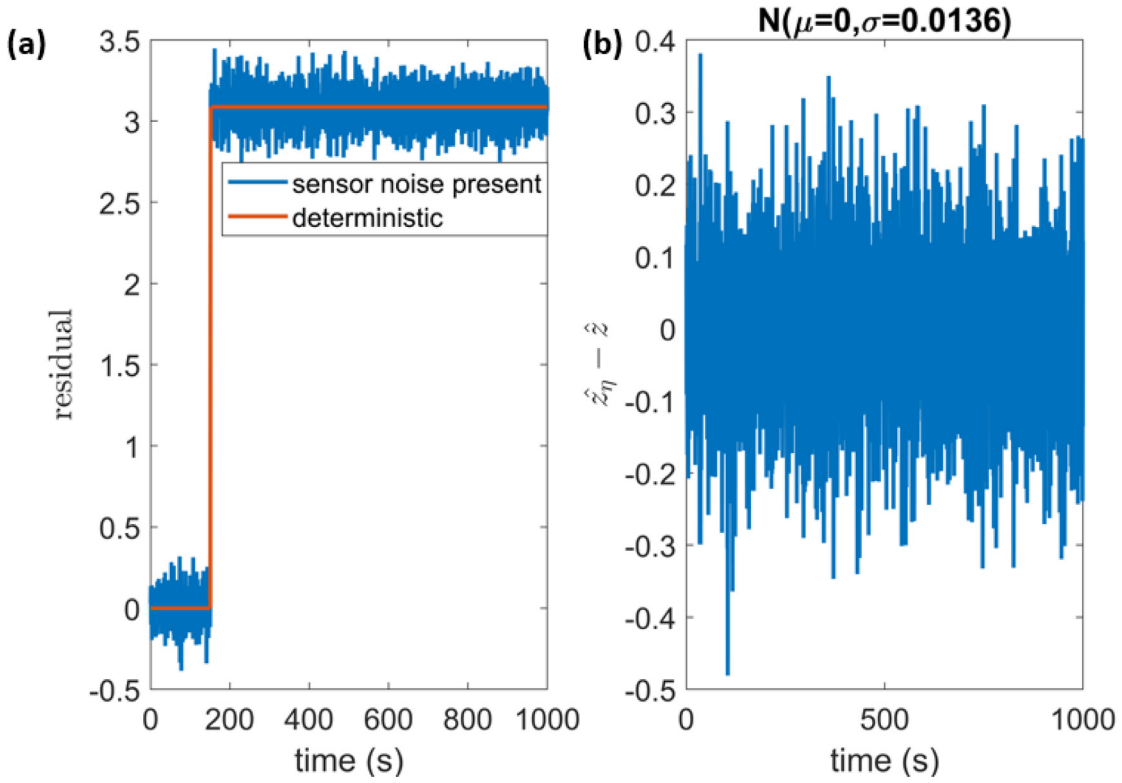

Fig. 3. (a) Residual vs time (seconds) for functional observer 2 (b) Difference between functional observer output in deterministic and stochastic case ($\hat{\mathbf{z}}_\eta - \hat{\mathbf{z}}$) corresponding to a N(0,0.0136) distribution (numerically observed).

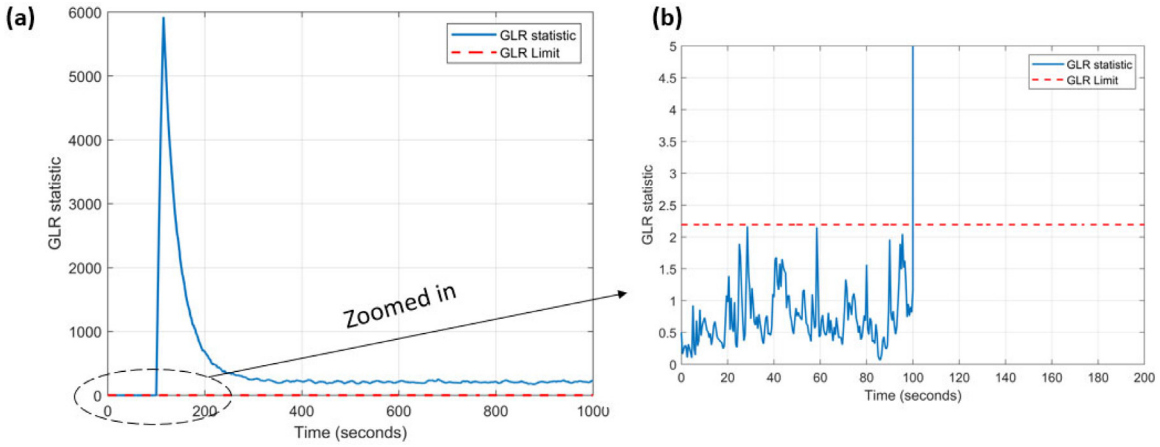

Fig. 4. (a) GLR statistic vs time for functional observer 1 in the presence of sensor noises (b) GLR statistic vs time zoomed in to show the behavior when fault occurs at t=100s.

7. Eigenvalue analysis for sensor noises

The eigenvalues of the observer play an important role in the determining the sensitivity of the observer to noises, and fault of interest. For the example considered in the previous section, the effect of the choice of eigenvalue ($-\alpha_1$) on the fault free residual variance for functional observer and the residual mean after fault occurrence is studied in this section.

7.1. Functional observer 1

In Fig. 6, the fault free residual variance for functional observer 1 vs eigenvalue ($-\alpha_1$) is plotted analytically (using Eq. (3.6)) and numerically (by calculating the variance of the simulated observer output). In both the cases, variance decreases monotonously with eigenvalue. Therefore, to minimize fault sensitivity, one is better off with a slower eigenvalue (close to 1).

To study the sensitivity of the functional observers to the fault of interest, the mean of the residual is calculated for different eigenvalues a long time after the fault has occurred and the residual has settled to its random distribution. In Fig. 7, the plot of the mean of the residual after fault occurrence vs eigenvalue is shown for two fault values ($f_1 = 1\frac{\text{mol}}{\text{L}}, 10\frac{\text{mol}}{\text{L}}$). In both the cases, the mean of the residual increases with eigenvalue and this increase is more pronounced for eigenvalues close to 1. Therefore, to maximize fault sensitivity it is advised to choose an eigenvalue close to 1.

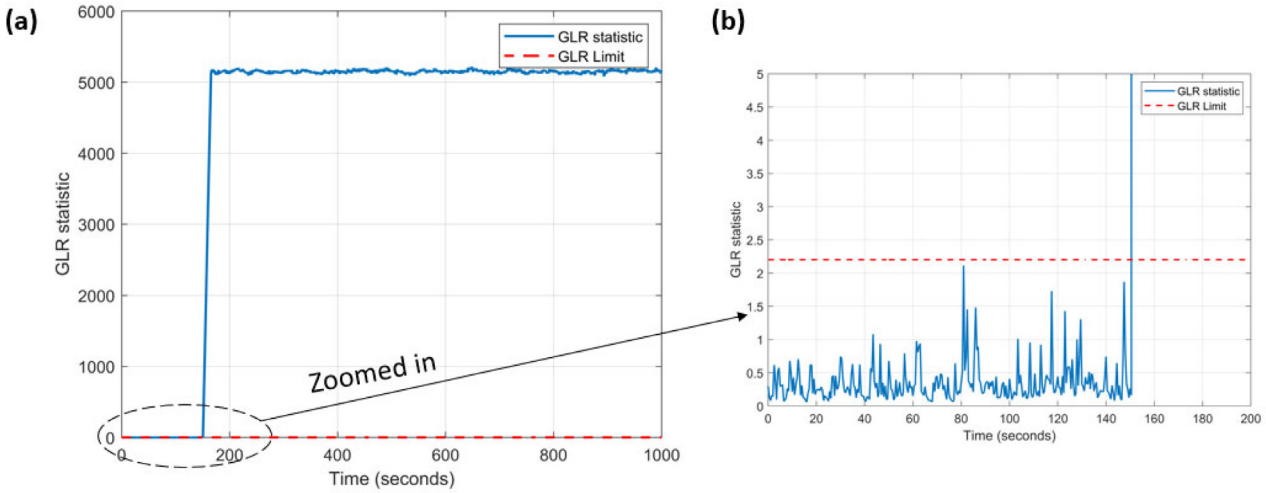

Fig. 5. (a) GLR statistic vs time for functional observer 2 in the presence of sensor noises (b) GLR statistic vs time zoomed in to show the behavior when fault occurs at t=150s.

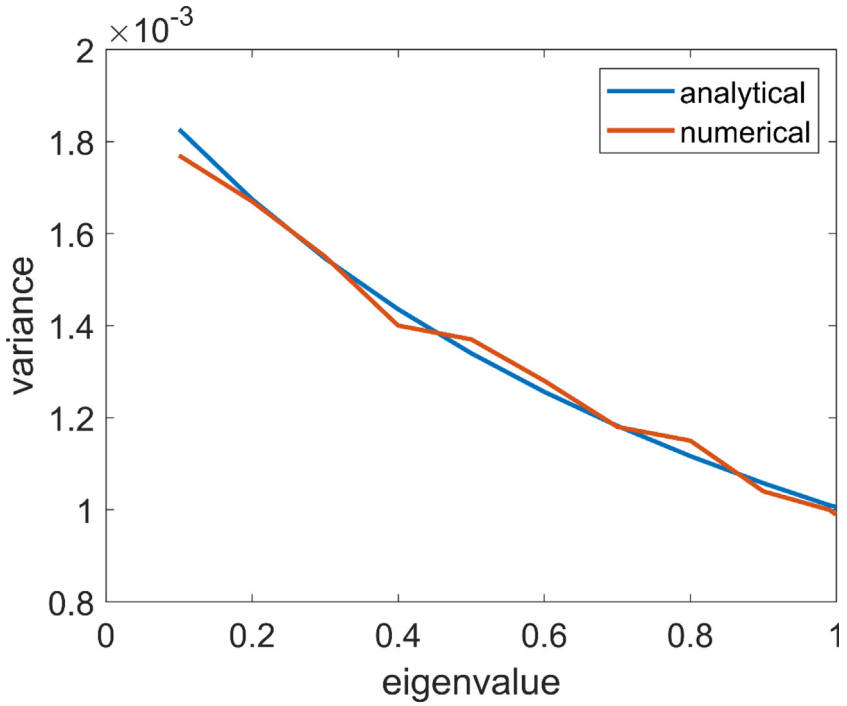

Fig. 6. Fault free variance vs eigenvalue (functional observer 1) in the presence of sensor noises for eigenvalues in [0.1,0.999].

In summary, for functional observer 1 a slower eigenvalue will have low noise sensitivity and high fault sensitivity. Thus, fault detection for the same fault value will be easier when the eigenvalue is close to 1. Once can also see this in Fig. 8 where the evolution of the residual vs time is plotted for two eigenvalues ($\alpha_1 = -0.5$ and -0.99, and fault $= 1$ mol/L). For $\alpha_1 = -0.5$, the residual after fault occurrence settles in the same region as the pre-fault residual making fault detection difficult but for $\alpha_1 = -0.99$, there is a clear shift in the mean compared to the pre-fault region.

7.2. Functional observer 2

The fault free residual variance from functional observer 2 vs eigenvalue is plotted in Fig. 9. Here, the variance decreases with the eigenvalue initially but increases as the eigenvalue approaches 1 (Fig. 9a) and shoots up rapidly very close to 1 (Fig. 9b). The mean of the residual vs eigenvalue for two different fault sizes is plotted in Fig. 10. In both the cases, mean of the residual increases with eigenvalue in this increase is profound for slower eigenvalues (closer to 1). It is to be noted that even though, noise sensitivity and fault sensitivity increase with eigenvalue, noise sensitivity (fault free variance) is much lower than fault sensitivity (residual mean at long times) and therefore, a faster eigenvalue can be used to detect faults robustly. In the previous section $\alpha_1 = -0.1$ was used to in the functional observer and the missed detection rate was much below 1%. Thus, in general, a careful study of the noise and fault sensitivities of the residual must be performed before choosing the eigenvalue.

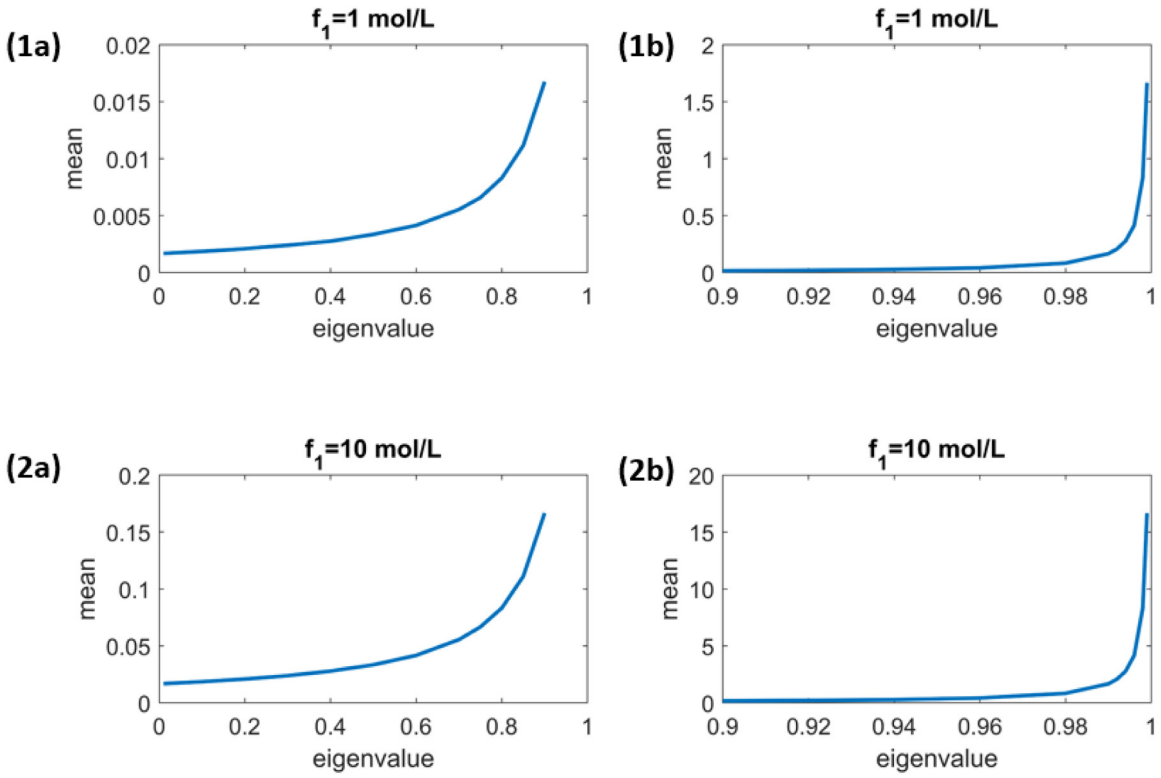

Fig. 7. Mean of the residual from functional observer 1 vs eigenvalue. (1a) For $\mathbf{f_1} = \mathbf{1}\frac{\mathbf{mol}}{\mathbf{L}}$ with eigenvalues in [0.01,0.9] (1b) $\mathbf{f_1} = \mathbf{1}\frac{\mathbf{mol}}{\mathbf{L}}$ with eigenvalues in [0.9,0.999] (2a) $\mathbf{f_1} = \mathbf{10}\frac{\mathbf{mol}}{\mathbf{L}}$ with eigenvalues in [0.01,0.9] (2b) $\mathbf{f_1} = \mathbf{10}\frac{\mathbf{mol}}{\mathbf{L}}$ with eigenvalues in [0.9,0.999].

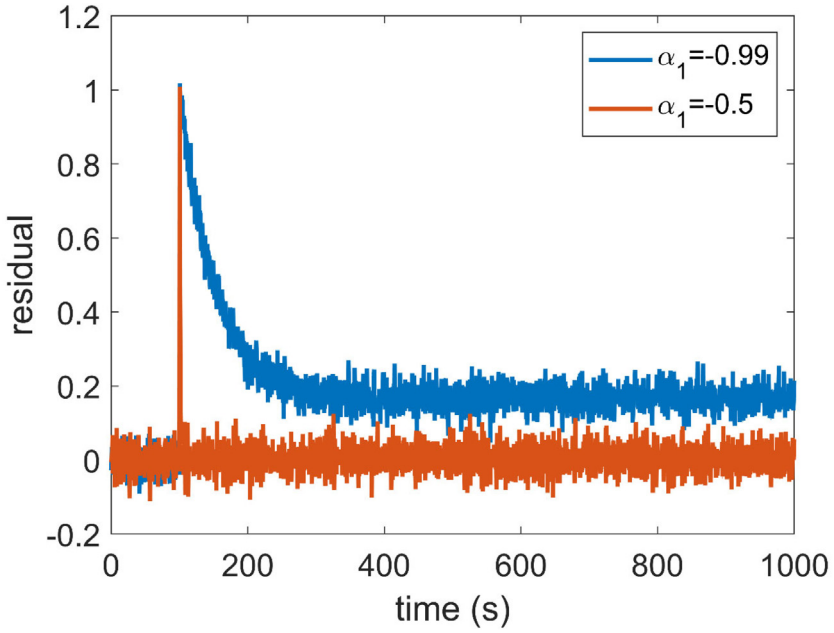

Fig. 8. Residual vs time from functional observer 1 for two eigenvalues (0.99 and 0.5).

8. Fault detection metrics for different fault magnitudes

The fault detection capabilities of the scheme presented in this paper is tested for different sensor fault magnitudes in the presence of sensor noises. The following scenario is assumed to occur: a sensor noise vector of zero mean and variance $\Sigma_\eta^2 = \begin{bmatrix} 0.001\left(\frac{\text{mol}}{\text{L}}\right)^2 & 0 & 0 \\ 0 & 0.01\text{K}^2 & 0 \\ 0 & 0 & 0.01\,\text{K}^2 \end{bmatrix}$ and a fault $\text{f}_1 = \begin{cases} 0, & \text{t} < 0 \\ \text{M}, & \text{x} \geq 100\text{s} \end{cases}$, $M \epsilon [0.01, 0.1]$ mol/L. Functional observer 1 of Section 6 with eigenvalue $\alpha_1 = -0.99$ is used and false alarm rate of below 1% is desired. In Fig. 11, the missed detection rate for different fault sizes is plotted. As expected, for extremely small

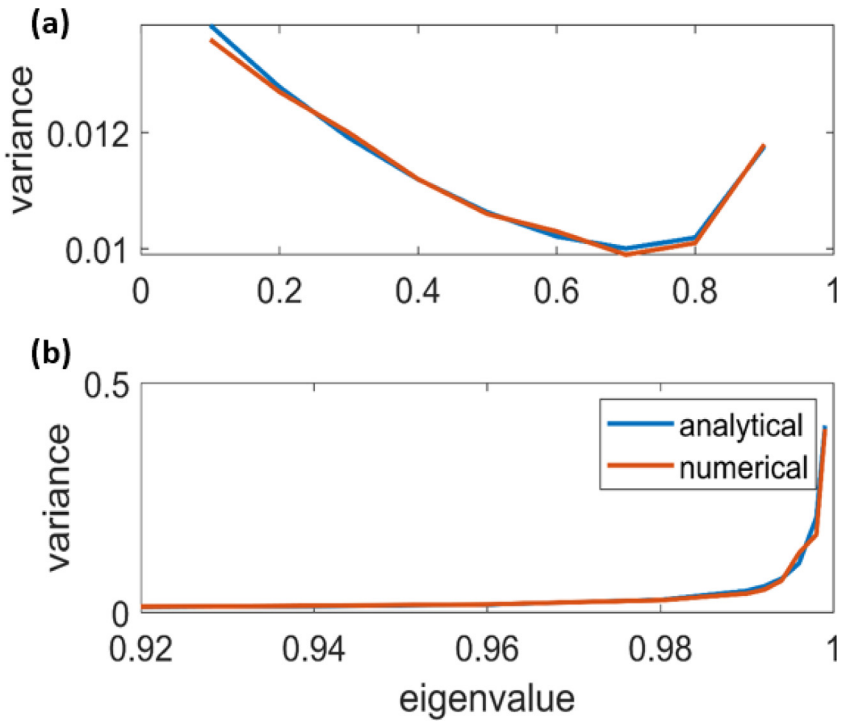

Fig. 9. Fault free variance vs eigenvalue (functional observer 2) in the presence of sensor noises. (a) For eigenvalues in [0.1,0.9] (b) For eigenvalues in [0.9,0.999].

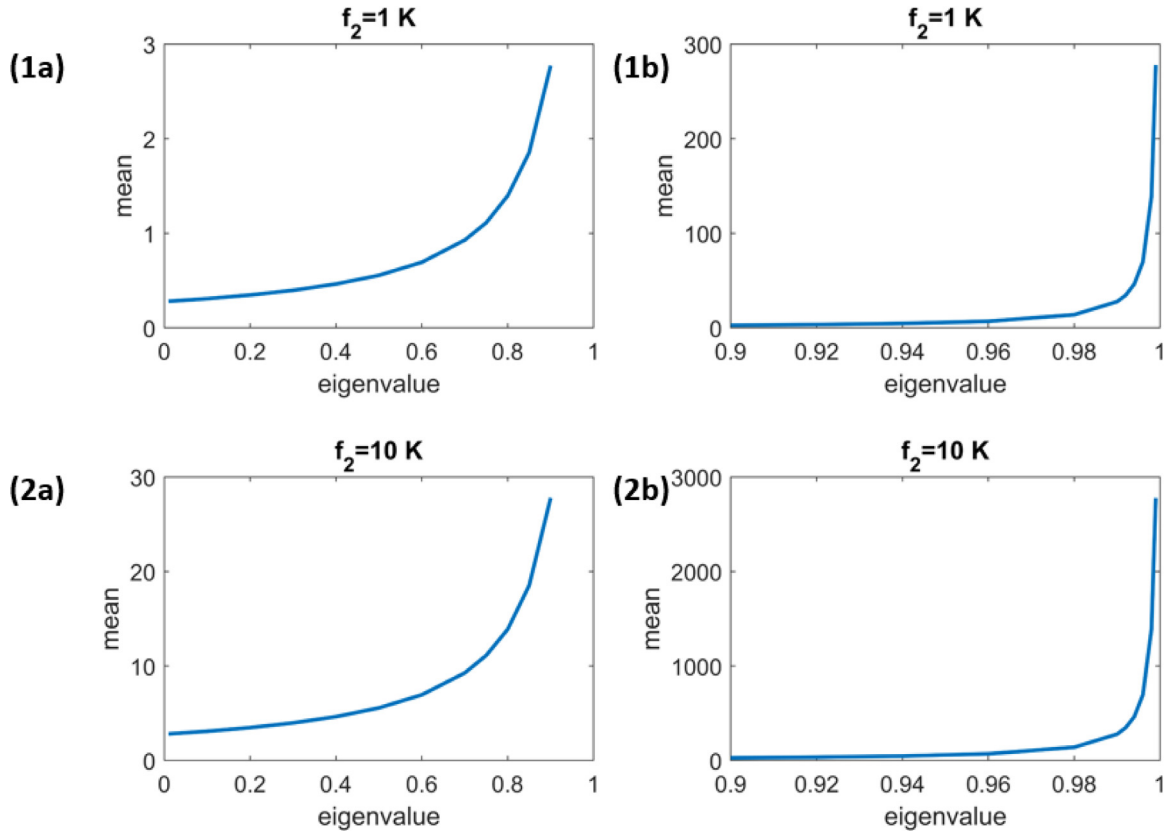

Fig. 10. Mean of the residual from functional observer 2 vs eigenvalue. (1a) For $\mathbf{f_2 = 1K}$ with eigenvalues in [0.01,0.9] (1b) $\mathbf{f_1 = 1\,K}$ with eigenvalues in [0.9,0.999] (2a) $\mathbf{f_1 = 10\,K}$ with eigenvalues in [0.01,0.9] (2b) $\mathbf{f_1 = 10\,K}$ with eigenvalues in [0.9,0.999].

fault sizes (< 0.02 mol/L), the missed detection rate is very high (> 80%). However, as the fault size increases a drastic decrease in the missed detection rate is observed and for fault sizes greater than 0.05 mol/L, the missed detection rate goes to 0.

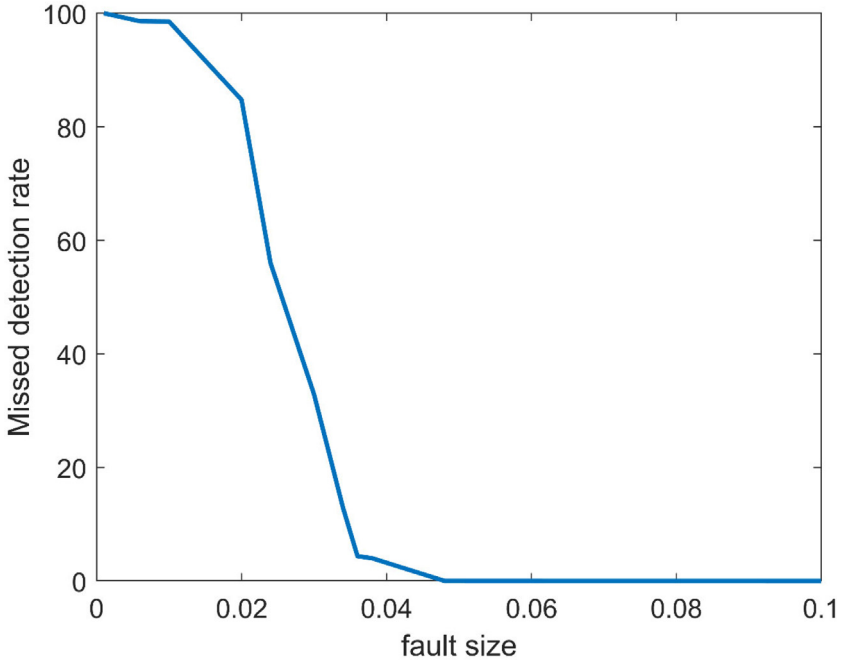

Fig. 11. Missed detection rate (%) vs sensor fault size.

9. Disturbance decoupled fault detection in the presence of process noises

Consider a nonlinear process described by:

$$\begin{aligned} x(k+1) &= F(x(k), W(k), f(k)) + \zeta(k) \\ y(k) &= H(x(k)) + G(x(k))\, W(k) + E(x(k))\, f(k) \end{aligned} \tag{9.1}$$

where $x(k) \in \mathbb{R}^n$ denotes the vector of states, $y(k) \in \mathbb{R}^p$ denotes the vector of measured outputs. $f(k) \in \mathbb{R}$ and $W(k) \in \mathbb{R}^m$ are the fault and the disturbances/uncertainties respectively and $E(x)$, $F(x)$, $G(x)$, $H(x)$, $J(x)$ are smooth functions and ζ is the process noise vector with mean

$$\begin{aligned} \hat{\xi}_\zeta(k+1) &= A\hat{\xi}_\zeta(k) + By(k) \\ \hat{z}_\zeta(k) &= C\hat{\xi}_\zeta(k) + Dy(k) \end{aligned} \tag{9.2}$$

where $\hat{\xi}_\zeta(k) \in \mathbb{R}^\nu$, $\hat{z}_\zeta(k) \in \mathbb{R}$, and A, B, C and D are given by (2.8) to detect faults in (9.1).

It is assumed that the response of the residual in the absence of faults at large times follows a Gaussian distribution of unknown mean and variance. A fault-free scenario needs to be simulated for the functional observers and the functional observer output will constitute the training data for obtaining the threshold for the GLR statistic (replaces step 1 in part a of the algorithm in Section 4). Once the fault-free random distribution is found the same steps of the algorithm in Section 4 for sensor noises for can be followed.

In what follows the non-isothermal CSTR example will be considered again. Simulations will be performed in the presence of process noises.

Consider the CSTR seen in the previous section

$$\begin{aligned} C_A(k+1) &= C_A(k) \\ &\quad + \delta_t \left(\frac{F}{V} \left(C_{A,in} - C_A(k) \right) - R\left(C_A(k), C_B(k), \theta(k), w(k) \right) \right) + \zeta_1(k) \\ C_B(k+1) &= C_B(k) \\ &\quad + \delta_t \left(\frac{F}{V} \left(C_{B,in} - C_B \right) - R\left(C_A(k), C_B(k), \theta(k), w(k) \right) \right) + \zeta_2(k) \\ \theta(k+1) &= \theta(k) + \delta_t \left(\tfrac{F}{V} \left(\theta_{in} - \theta(k) \right) - \tfrac{US_A}{\rho c_p V} \left(\theta(k) - \theta_j(k) \right) \right) \\ &\quad + \tfrac{\delta_t(-\Delta H_R)}{\rho c_p} \left(R\left(C_A(k), C_B(k), \theta(k), w(k) \right) \right) + \zeta_2(k) \\ \theta_j(k+1) &= \theta_j(k) + \delta_t \left(\frac{F_j}{V_j} \left(\theta_{j,in}(k) + f_2(k) - \theta_j(k) \right) \right. \\ &\quad \left. + \frac{US_A}{\rho_j c_{p_j} V_j} \left(\theta(k) - \theta_j(k) \right) \right) + \zeta_3(k) \\ y_1 &= C_A(k) + f_1(k) \\ y_2 &= \theta(k) \\ y_3 &= \theta_j(k) \end{aligned} \tag{9.3}$$

The sampling period is again $\delta_t = 0.5$s. The model is converted to deviation form and the same functional observers in Section 6 are used here. The noise vectors $N = [\zeta_1, \zeta_2, \zeta_3]$ are Gaussian with zero mean and variance

$$\Sigma_\zeta^2 = \begin{bmatrix} 0.0001 \left(\frac{mol}{Ls} \right)^2 & 0 & 0 \\ 0 & 0.001 \left(\frac{K}{s} \right)^2 & 0 \\ 0 & 0 & 0.001 \left(\frac{K}{s} \right)^2 \end{bmatrix}.$$

The fault scenario is as follows

$$f_1(t) = \begin{cases} 0, & t < 100s \\ 1, & t \geq 100s, \end{cases} \quad f_2(t) = \begin{cases} 0, & t < 150s \\ 10, & t \geq 150s \end{cases}, \quad w(t) = 10^5.$$

A false alarm rate of less than 1% is desired.

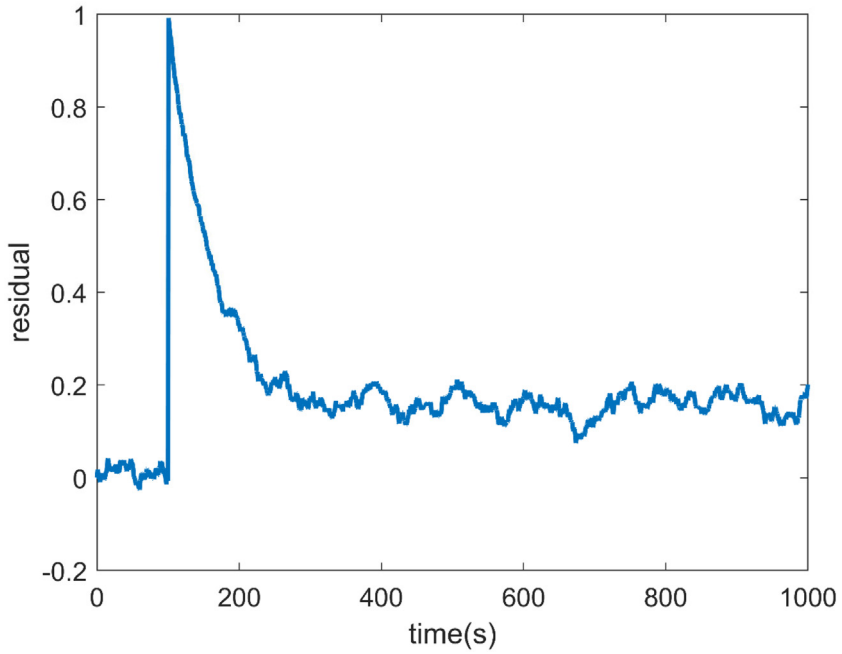

Fig. 12. Residual vs time for functional observer 1 in the presence of process noises.

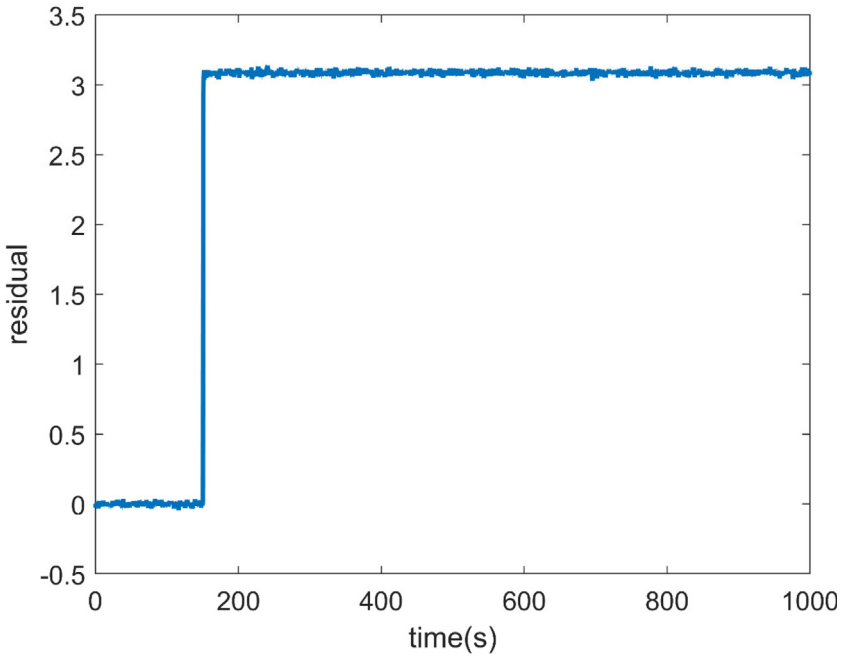

Fig. 13. Residual vs time for functional observer 2 in the presence of process noises.

9.1. GLR calculation

The residuals generated from the functional observer 1 and functional observer 2 are plotted in Figs. 12 and 13. The fault free distributions of the output of the two functional observers are used to develop thresholds of the GLR statistic (Part a of the algorithm in Section 4). The thresholds for residuals from functional observer 1 and functional observer 2 are 33.01 and 3.205, respectively. The time series evolution of the GLR statistic for the two residuals is plotted in Figs. 14 and 15. In both cases when the respective fault occurs as sharp increase GLR statistic is observed which then settles to a value higher than the threshold for fault free data in turn facilitating rapid fault detection with 0% missed detection.

The fault detection metrics are summarized in Table 2. For sensor and coolant faults, both the missed detection rate and the false alarm rate is 0% and average run length (ARL) is 1 for the sensor fault and 2 for the coolant fault. Therefore, even for process noises, the fault diagnosis scheme is able to both rapidly and accurately detect and isolate faults of interest.

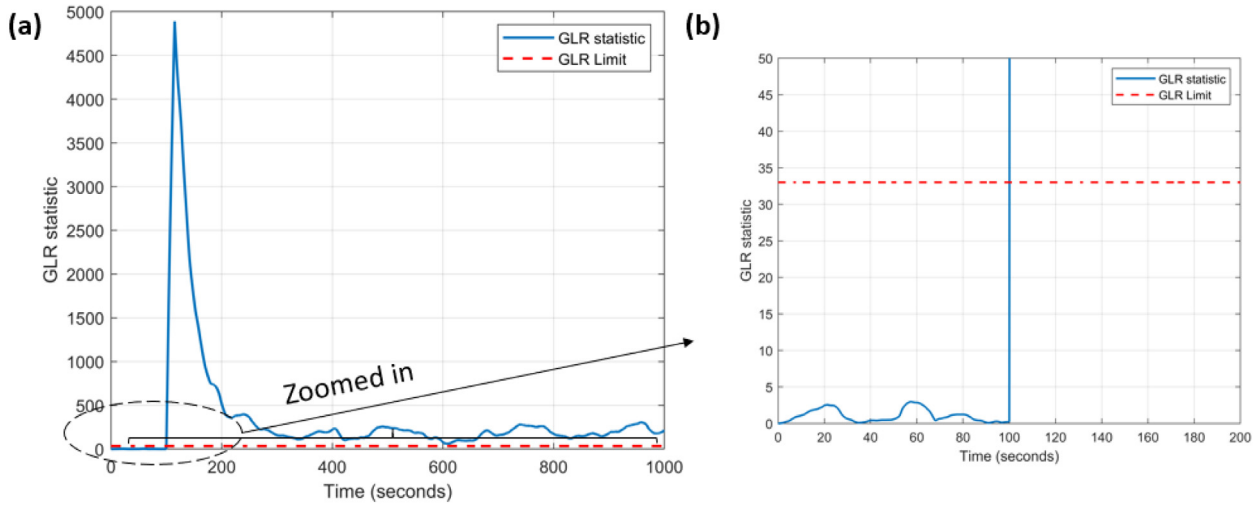

Fig. 14. (a) GLR statistic vs time for functional observer 1 in the presence of process noises (b) GLR statistic vs time zoomed in to show the behavior when fault occurs at t=100s.

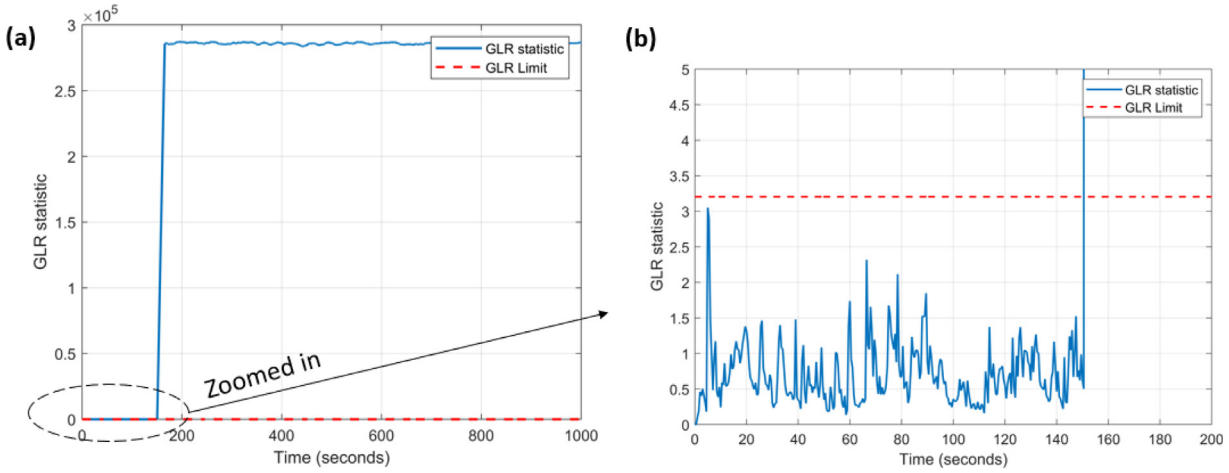

Fig. 15. (a) GLR statistic vs time for functional observer 2 in the presence of process noises (b) GLR statistic vs time zoomed in to show the behavior when fault occurs at t=150s.

Table 2
Fault detection metrics in the presence of process noises.

	Observer 1	Observer 2
Missed Detection Rate (%)	0.0000	0.0000
False Alarm Rate (%)	0.0000	0.0000
ARL_1	1.0000	2.0000

9.2. Eigenvalue tuning for process noises

In Fig. 16, the fault free residual variance for functional observer 1 vs eigenvalue ($-\alpha_1$) numerically (by calculating the variance of the simulated observer output). The fault free residual variance increases with eigenvalue and this increase is drastic as the eigenvalue tends to 1. Therefore, to minimize fault sensitivity one is better off with a faster eigenvalue. In Fig. 17, the plot of the mean of the residual after fault occurrence vs eigenvalue is shown for two fault values ($f_1 = 1\frac{mol}{L}, 10\frac{mol}{L}$). In both the cases, the mean of the residual increases with eigenvalue. However, the increase in the residual mean is much higher than the increase in noise sensitivity for eigenvalues close to 1. Therefore, to have low missed detection rate it is advisable to choose an eigenvalue closer to 1.

For functional observer 2, the fault free residual variance for functional observer 2 vs eigenvalue ($-\alpha_1$) is plotted in Fig. 18. The fault free residual variance increases with eigenvalue and it shoots up as the eigenvalue tends to 1. Thus, to minimize fault sensitivity one is better off with a faster eigenvalue. In Fig. 19, the plot of the mean of the residual after fault occurrence vs eigenvalue is shown for two fault values ($f_1 = 1K, 10K$). In both the cases, the mean of the residual increases with eigenvalue. It is to be noted that even though noise sensitivity and fault sensitivity increase with eigenvalue, noise sensitivity (fault free variance) is much lower than fault sensitivity (residual mean) and therefore, a faster eigenvalue can be used to detect faults robustly. For example, in functional observer 2, $\alpha_1 = -0.1$ was used and the missed detection rate was significantly below 1%. In general, a careful study of the noise and fault sensitivities of the residual must be performed before choosing the eigenvalue.

10. Conclusions

This work tackled the problem of detecting faults in nonlinear systems in the presence of noises. A methodology based on designing model-based functional observers to generate residuals for fault diagnosis is proposed. The behavior of the residuals in

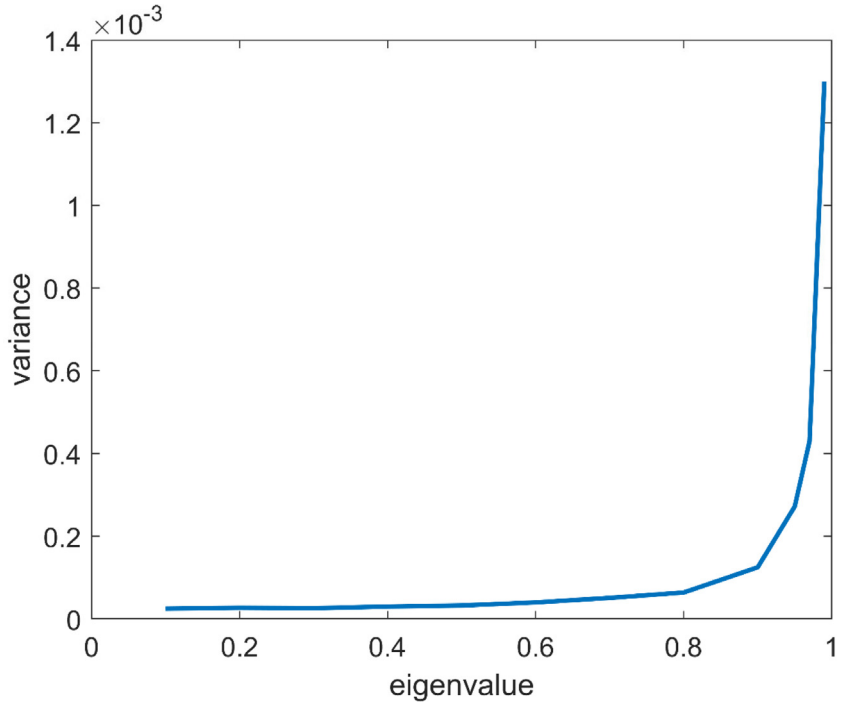

Fig. 16. Residual variance (fault free) vs eigenvalue for functional observer 1 in the presence of process noises.

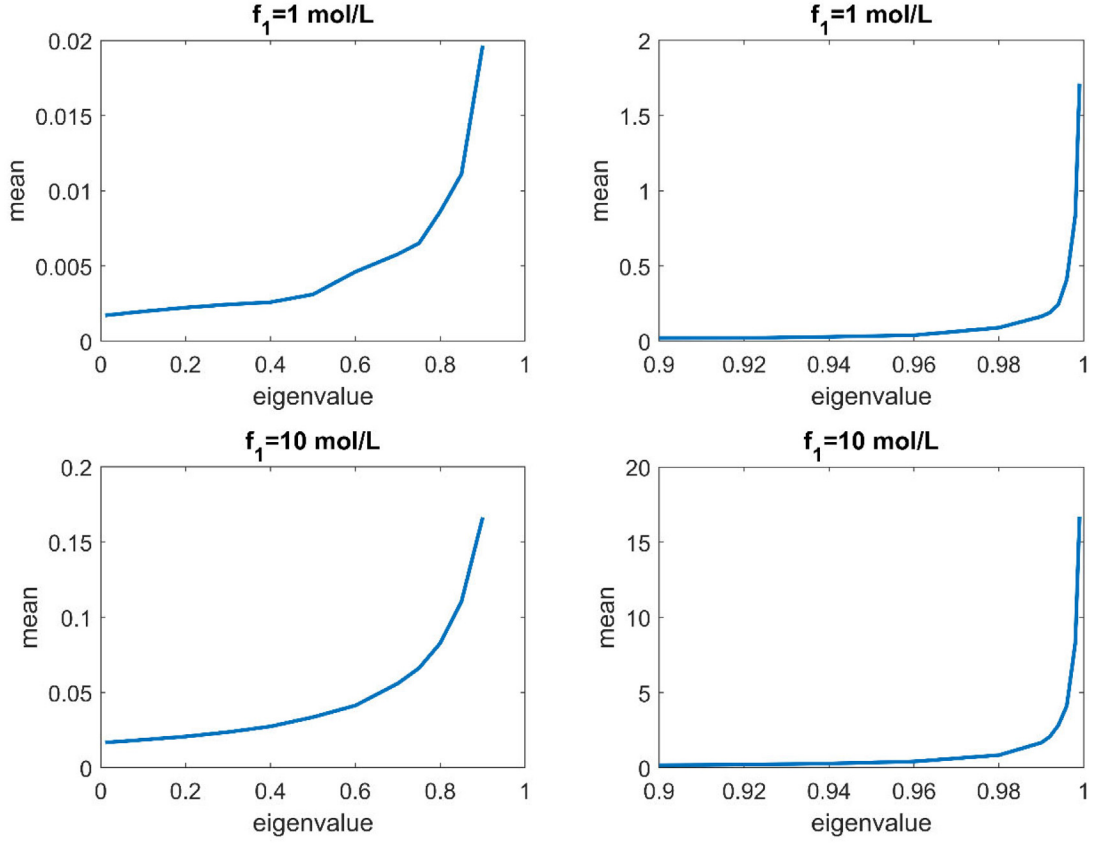

Fig. 17. Functional observer 1 residual mean vs eigenvalue for different sensor fault magnitudes in the presence of process noises.

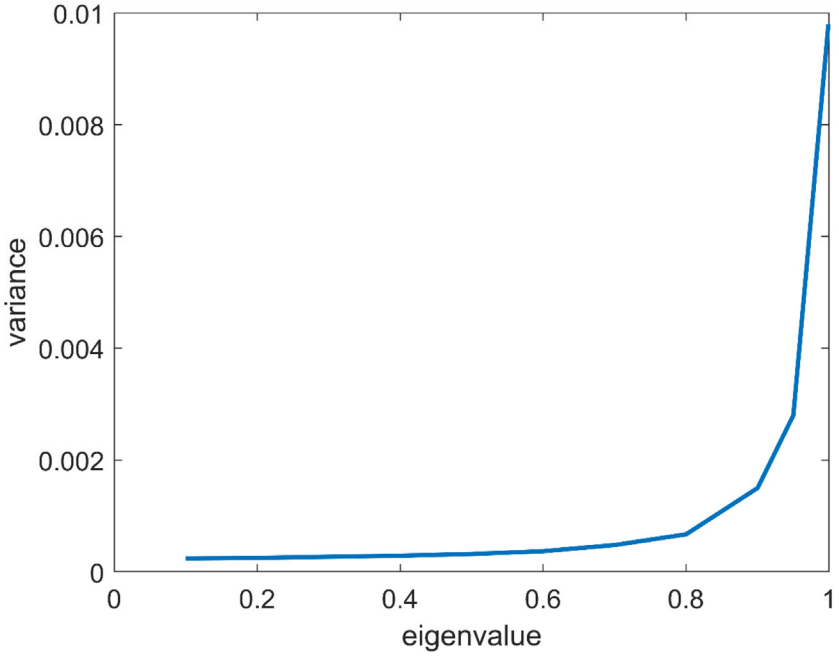

Fig. 18. Fault free variance vs eigenvalue for functional observer 2 in the presence of process noises.

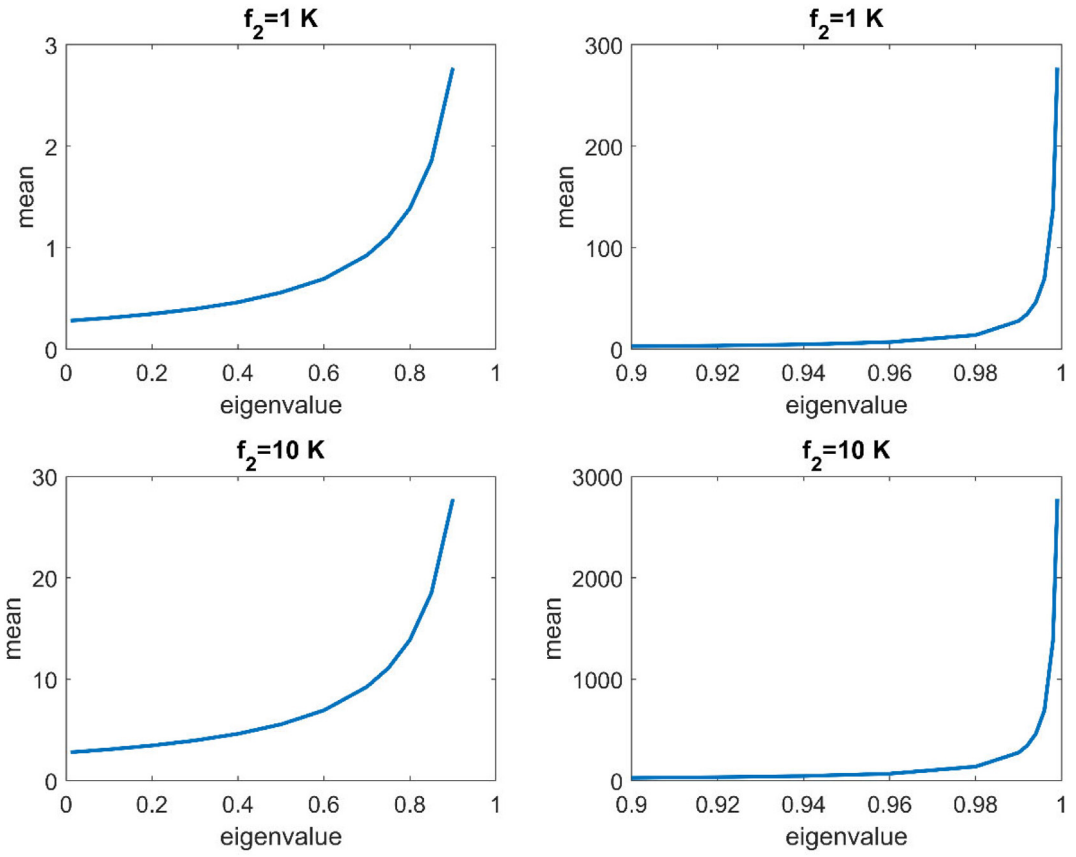

Fig. 19. Functional observer 2 residual mean vs eigenvalue for different coolant fault magnitudes in the presence of process noises.

the presence of Gaussian sensor noises can be analyzed analytically and the resulting probability distributions can be derived. Once the fault free mean and variance have been derived, a well-known statistical hypothesis testing method known as Generalized Likelihood Ratio, can be used to detect changes in the mean of the residual to enable fault detection. The applicability of the proposed scheme was illustrated using a non-isothermal CSTR case-study to detect and isolate sensor and inlet coolant temperature faults. The results show that the fault diagnosis scheme is able to quickly and accurately detect faults. The case study also highlighted the crucial role the functional observer's eigenvalue plays in the observer's ability to detect faults and highlights the need to carefully tune this parameter to ensure maximum fault sensitivity and minimal noise sensitivity. Finally, using an entirely empirical approach, the methodology was successfully extended to process noises.

CRediT authorship contribution statement

Sunjeev Venkateswaran: Conceptualization, Methodology, Formal analysis, Writing – original draft, Software. **M. Ziyan Sheriff:** Methodology, Formal analysis, Writing – original draft, Software. **Benjamin Wilhite:** Supervision. **Costas Kravaris:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Financial support from the National Science Foundation, United States through the grant CBET-1706201 is gratefully acknowledged.

References

[1] A.E. Theis, Case study: T2 Laboratories explosion, J. Loss Prev. Process Ind. 30 (2014) 296–300.
[2] J.A. Venkidasalapathy, T2 laboratories explosion 10th year anniversary, 2007.
[3] R.J. Willey, H.S. Fogler, M.B. Cutlip, The integration of process safety into a chemical reaction engineering course: kinetic modeling of the T2 incident, Process Saf. Prog. 30 (2011) 39–44.
[4] CSB, Williams Olefins plant explosion and fire, 2016.
[5] L. Grim, CSB investigation: Williams Geismar Olefins plant reboiler rupture and fire, in: ASSE Professional Development Conference and Exposition, American Society of Safety Engineers, 2017.
[6] V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri, A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies, Comput. Chem. Eng. 27 (2003) 313–326.
[7] V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri, K. Yin, A review of process fault detection and diagnosis: Part III: Process history based methods, Comput. Chem. Eng. 27 (2003) 327–346.
[8] V. Venkatasubramanian, R. Rengaswamy, K. Yin, S.N. Kavuri, A review of process fault detection and diagnosis: Part I: Quantitative model-based methods, Comput. Chem. Eng. 27 (2003) 293–311.
[9] S.X. Ding, Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools, Springer Science & Business Media, 2008.
[10] P.M. Frank, X. Ding, Survey of robust residual generation and evaluation methods in observer-based fault detection systems, J. Process Control 7 (1997) 403–424.
[11] M.P. Frank, X.S. Ding, B. Koppen-Seliger, Current developments in the theory of FDI, IFAC Proc. Vol. 33 (2000) 17–28.
[12] J. Gertler, Analytical redundancy methods in fault detection and isolation-survey and synthesis, IFAC Proc. Vol. 24 (1991) 9–21.
[13] P.M. Frank, Advanced fault detection and isolation schemes using nonlinear and robust observers, IFAC Proc. Vol. 20 (1987) 63–68.
[14] P.M. Frank, Fault diagnosis in dynamic systems via state estimation-a survey, in: System Fault Diagnostics, Reliability and Related Knowledge-Based Approaches, Springer, 1987, pp. 35–98.
[15] P.M. Frank, Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results, Automatica 26 (1990) 459–474.
[16] R.V. Beard, Failure Accomodation in Linear Systems Through Self-Reorganization, Massachusetts Institute of Technology, 1971.
[17] H.L. Jones, Failure Detection in Linear Systems, Massachusetts Institute of Technology, 1973.
[18] P.S. Teh, H. Trinh, Design of unknown input functional observers for nonlinear systems with application to fault diagnosis, J. Process Control 23 (2013) 1169–1184.
[19] Y. Yang, S.X. Ding, L. Li, On observer-based fault detection for nonlinear systems, Systems Control Lett. 82 (2015) 18–25.
[20] H. Han, Y. Yang, L. Li, S.X. Ding, Observer-based fault detection for uncertain nonlinear systems, J. Franklin Inst. B 355 (2018) 1278–1295.
[21] S. Venkateswaran, Q. Liu, B.A. Wilhite, C. Kravaris, Design of linear residual generators for fault detection and isolation in nonlinear systems, Internat. J. Control (2020) 1–17.
[22] S. Venkateswaran, B.A. Wilhite, C. Kravaris, Functional observers with linear error dynamics for discrete-time nonlinear systems, in: Proceedings 60th IEEE Conference on Decision and Control, IEEE, 2021, See also arXiv: 2106.02185.
[23] M.Z. Sheriff, M. Mansouri, M.N. Karim, H. Nounou, M. Nounou, Fault detection using multiscale PCA-based moving window GLRT, J. Process Control 54 (2017) 47–64.
[24] M. Basseville, Model-based statistical signal processing and decision theoretic approaches to monitoring, IFAC Proc. Vol. 36 (2003) 1–12.
[25] F. Kiasi, J. Prakash, S. Shah, Detection and diagnosis of incipient faults in sensors of an LTI system using a modified GLR-based approach, J. Process Control 33 (2015) 77–89.
[26] F. Kiasi, J. Prakash, S. Patwardhan, S. Shah, A unified framework for fault detection and isolation of sensor and actuator biases in linear time invariant systems using marginalized likelihood ratio test with uniform priors, J. Process Control 23 (2013) 1350–1361.
[27] N. Basha, M.Z. Sheriff, C. Kravaris, H. Nounou, M. Nounou, Multiclass data classification using fault detection-based techniques, Comput. Chem. Eng. 136 (2020) 106786.
[28] F. Harrou, M.N. Nounou, H.N. Nounou, M. Madakyaru, PLS-based EWMA fault detection strategy for process monitoring, J. Loss Prev. Process Ind. 36 (2015) 108–119.
[29] S. Yin, X. Zhu, O. Kaynak, Improved PLS focused on key-performance-indicator-related fault diagnosis, IEEE Trans. Ind. Electron. 62 (2014) 1651–1658.
[30] S.W. Choi, C. Lee, J.-M. Lee, J.H. Park, I.-B. Lee, Fault detection and identification of nonlinear processes based on kernel PCA, Chemometr. Intell. Lab. Syst. 75 (2005) 55–67.
[31] M.Z. Sheriff, C. Botre, M. Mansouri, H. Nounou, M. Nounou, M.N. Karim, Process monitoring using data-based fault detection techniques: Comparative studies, in: Fault Diagnosis and Detection, InTech, 2017, pp. 237–261.
[32] M.Z. Sheriff, M.N. Karim, H.N. Nounou, M.N. Nounou, Process monitoring using PCA-based GLR methods: A comparative study, J. Comput. Sci. 27 (2018) 227–246.
[33] C. Botre, M. Mansouri, M.N. Karim, H. Nounou, M. Nounou, Multiscale PLS-based GLRT for fault detection of chemical processes, J. Loss Prev. Process Ind. 46 (2017) 143–153.
[34] C. Botre, M. Mansouri, M. Nounou, H. Nounou, M.N. Karim, Kernel PLS-based GLRT method for fault detection of chemical processes, J. Loss Prev. Process Ind. 43 (2016) 212–224.
[35] T. Sorsa, H.N. Koivo, H. Koivisto, Neural networks in process fault diagnosis, IEEE Trans. Syst. Man Cybern. 21 (1991) 815–825.
[36] V. Venkatasubramanian, K. Chan, A neural network methodology for process fault diagnosis, AIChE J. 35 (1989) 1993–2002.
[37] L. Wen, X. Li, L. Gao, Y. Zhang, A new convolutional neural network-based data-driven fault diagnosis method, IEEE Trans. Ind. Electron. 65 (2017) 5990–5998.
[38] S. Heo, J.H. Lee, Fault detection and classification using artificial neural networks, IFAC-PapersOnLine 51 (2018) 470–475.
[39] H. Wu, J. Zhao, Deep convolutional neural network model based chemical process fault diagnosis, Comput. Chem. Eng. 115 (2018) 185–197.
[40] H. Shao, H. Jiang, H. Zhao, F. Wang, A novel deep autoencoder feature learning method for rotating machinery fault diagnosis, Mech. Syst. Signal Process. 95 (2017) 187–204.
[41] F. Lv, C. Wen, Z. Bao, M. Liu, Fault diagnosis based on deep learning, in: 2016 American Control Conference, ACC, IEEE, 2016, pp. 6851–6856.
[42] F. Gustafsson, Stochastic fault diagnosability in parity spaces, IFAC Proc. Vol. 35 (2002) 41–46.
[43] M.R. Reynolds Jr., J. Lou, An evaluation of a GLR control chart for monitoring the process mean, J. Qual. Technol. 42 (2010) 287–310.

[44] Q. Zhang, M. Basseville, A. Benveniste, Early warning of slight changes in systems, Automatica 30 (1994) 95–113.
[45] C. Kravaris, S. Venkateswaran, Functional observers with linear error dynamics for nonlinear systems, Systems Control Lett. (2021) in press. See also arXiv:2101.11148.
[46] H. Stark, J.W. Woods, Probability and Random Processes with Applications To Signal Processing: International Edition, Pearson Higher Ed, 2014.
[47] M.R. Reynolds Jr., J. Lou, J. Lee, S. Wang, The design of GLR control charts for monitoring the process mean and variance, J. Qual. Technol. 45 (2013) 34–60.
[48] X. Cui, M.S. Mannan, B.A. Wilhite, Towards efficient and inherently safer continuous reactor alternatives to batch-wise processing of fine chemicals: CSTR nonlinear dynamics analysis of alkylpyridines N-oxidation, Chem. Eng. Sci. 137 (2015) 487–503.