Particle Filtering Based Likelihood Ratio Approach to Fault Diagnosis in Nonlinear Stochastic Systems

Ping Li and Visakan Kadirkamanathan

***Abstract*—This paper presents the development of a particle filtering (PF) based method for fault detection and isolation (FDI) in stochastic nonlinear dynamic systems. The FDI problem is formulated in the multiple model (MM) environment, then by combining the likelihood ratio (LR) test with the PF, a new FDI scheme is developed. The simulation results on a highly nonlinear system are provided which demonstrate the effectiveness of the proposed method.**

***Index Terms*—Extended Kalman filter (EKF), fault diagnosis, likelihood ratio (LR) test, Monte–Carlo technique, nonlinear stochastic system, particle filter (PF).**

I. Introduction

THE PROBLEM of fault detection and isolation (FDI) in dynamic systems has attracted considerable attention world-wide and been theoretically and experimentally investigated with different types of approaches, as can be seen from the survey papers [3], [4], [9], [11], [12], [14], [22], [29], and the books [5], [23], [24]. This development has been mainly stimulated by the trend in automation toward systems with increasing complexity and the growing demands for fault tolerance, cost efficiency, reliability, and security which constitute fundamental design features in modern control systems. The FDI approaches appeared in literature fall into two major categories, i.e., the model-based approaches which make use of the quantitative analytical model of the system to be monitored and the knowledge-based or model-free approaches which do not need full analytical modeling and allow one to use qualitative models based on the available information and knowledge of the system to be monitored. Clearly a perfect analytical model (if available) represents the deepest and most concise knowledge of the system, hence, in the case of information-rich systems where the dynamic behavior of system can be well-described by mathematical models, the analytical model-based methods are by nature the most powerful fault diagnosis methods.

For all model-based approaches, the decision of a fault is based on available input–output (I/O) measurements and a mathematical model of the system to be monitored. One of the main difficulties in FDI of dynamic systems is due to the presence of unknown and unmeasured state variable $\mathbf{x}$. Two approaches are commonly used to deal with them: *estimation* and *elimination*. The estimation of $\mathbf{x}$ is usually performed with observers for deterministic systems, or filters for stochastic systems, which lead to the observer-based and the innovation-based FDI approaches respectively. The elimination of $\mathbf{x}$ directly explores the analytical redundancy embodied in the mathematic model. For linear system, this leads to the well-known parity space (or parity equation)-based FDI approach. However, the literature addressing analytical model-based FDI for nonlinear systems is not extensive, the main reason being that the estimation of the state or measurement vector of a nonlinear system is not easy and analytically performing the manipulations to eliminate $\mathbf{x}$ is often impossible for general nonlinear dynamic systems. The model-based FDI for nonlinear system is known as a difficult problem and very few results are available.

In this paper, we address the FDI problem in general nonlinear stochastic systems, which has been investigated to a lesser extent. When restricted to systems described by a stochastic state space model, much of the development in FDI schemes has relied on the system being linear and the noise and disturbances being Gaussian. In such cases, the Kalman filter is usually employed for state estimation. The innovation from the Kalman filter is then used as the residual, based on which some statistical hypothesis tests are carried out for fault detection (FD)[21], [29]. Fault isolation (FI) is usually achieved by employing the observer/filter scheme [11], or using the multiple model (MM) and generalized likelihood ratio (GLR) methods [27].

The idea used in the linear case mentioned above has been extended to some nonlinear stochastic systems with additive Gaussian noise and disturbance by employing the linearization and Gaussianization techniques, and in this case, the Kalman filter is usually replaced by the extended Kalman filter (EKF) [26], [28], [32]. Although this EKF-based approach appears perfectly straightforward, there are no general results to guarantee that such approximation will work well in most case and the FDI performance of this approach depends very much on the particular application as indicated in [27]. The FDI problems in general nonlinear non-Gaussian stochastic systems are still open.

Recently, the *particle filter,* (PF) a Monte–Carlo technique-based method for nonlinear non-Gaussian state estimation, has attracted much attention [6], [8], [13], [16]. This interest stems from the great advantage of the PF being able to handle any functional nonlinearity and system or measurement noise of any probability distribution. Our early work [15], represents the first attempt to introduce PF into the field of FDI. More recently,

Manuscript received December 30, 2000; revised July 24, 2001. This work was supported by the UK CVCP ORS Award and the Department of Automatic Control and Systems Engineering, University of Sheffield.

The authors are with the Department of Automatic Control and Systems Engineering, University of Sheffield, U.K. (e-mail: l.ping@sheffield.ac.uk; visakan@sheffield.ac.uk).

Publisher Item Identifier S 1094-6977(01)09791-7.

we developed a particle filtering based MM approach to FDI by combining PF with Bayesian inference [20].

In this paper, the PF is employed to develop a new method for solving the FDI problem in general nonlinear and non-Gaussian systems. The paper is organized as follows. In Section II, the FDI problem is formulated in the MM environment followed by a description of the PF in Section III. Then, the particle filtering technique is combined with the likelihood ratio (LR) test and a new approach to FDI in nonlinear non-Gaussian systems is developed in Section IV. Experimental results from simulations are provided in Section V with conclusions and further work in Section VI.

II. Problem Statement

The problem of FD consists of making the decision on the presence or absence of faults in the monitored system and the problem of FI consists of deciding the present faulty mode among a number of possible modes. In this section, the FDI problem in the stochastic nonlinear systems is formulated in the MM environment. The MM method was originally developed for system identification, adaptive state estimation and control [19]. The use of MM method in FDI is reviewed in [29] and [31] (see also [20], [27], and [28]). Throughout this paper, it is assumed that the normal behavior and all possible faults of the physical system to be monitored can be described by a given finite set of nonlinear stochastic state space models indexed by $m = 0, 1, \ldots, M$

$$\mathbf{x}_k^{(m)} = \mathbf{f}_{k-1}^{(m)}(\mathbf{x}_{k-1}^{(m)}, \mathbf{w}_{k-1}^{(m)}) \tag{1}$$

$$\mathbf{y}_k = \mathbf{h}_k^{(m)}(\mathbf{x}_k^{(m)}, \mathbf{v}_k^{(m)}) \tag{2}$$

where

- $\mathbf{x}$ state vector;
- $\mathbf{w}$ zero mean white noise vector independent of past and current state;
- $\mathbf{f}(\cdot, \cdot)$ vector-valued nonlinear state transition function;
- $\mathbf{y}$ output measurement vector;
- $\mathbf{v}$ zero mean white measurement noise vector, independent of past and present states and the system noise $\mathbf{w}$;
- $\mathbf{h}(\cdot, \cdot)$ vector-valued nonlinear measurement function.

The probability density functions (pdfs) of $\mathbf{w}_k$ and $\mathbf{v}_k$ are both assumed to be known. Note that the dimensions of the state vector $\mathbf{x}_k^{(m)}$ may be different for different m (corresponding to assuming that a fault may lead to a change in the order of the system dynamics) and, the disturbance $\mathbf{w}_k$ and noise $\mathbf{v}_k$ need not necessarily enter additively.

After having established the mathematical models of the system, we can now express the FDI problem through a model based approach. Initially, the system works normally and its behavior is governed by the given normal model described as in (1) and (2) (indexed by $m = 0$), but the model may change at an unknown time t_0 subject to the initiation of a fault. Thus the FDI problems can be stated as follows.

Problem 1 (Problem of FD): FD is to decide a model shift, or more specifically, detecting a jump from the normal (fault-free) model (indexed by $m = 0$) to the faulty models (indexed by $m = 1, \ldots, M$) and perhaps estimating the time t_0 at which this jump takes place.

Problem 2 (Problem of FI): Fault isolation is to determine which of the M possible faulty models the system has jumped to.

III. Particle Filters

In this section, we give a brief explanation of the PF that forms the basis for the development of the new FDI method for general nonlinear non-Gaussian systems. A detailed description of the PF can be found in [13] and [16]. We begin with the Bayesian solution to the dynamic state estimation problem which involves the construction of the pdf of the current state $\mathbf{x}_k$, given the measurements up to time k. If $\mathcal{Z}_k$ is denoted to be the set of measurements up to time k, i.e., $\mathcal{Z}_k = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k\}$, then the Bayesian solution would be to calculate the pdf $p(\mathbf{x}_k|\mathcal{Z}_k)$. This pdf will encapsulate all the information about the state $\mathbf{x}_k$ which is contained in the measurements $\mathcal{Z}_k$ and the prior pdf of $\mathbf{x}_0$. Once $p(\mathbf{x}_k|\mathcal{Z}_k)$ is known, the estimates of functions of the state $\mathbf{x}_k$ conditional on measurements $\mathcal{Z}_k$, can be made. For example, the minimum mean squared error estimate of $\mathbf{x}_k$ given $\mathcal{Z}_k$ is

$$\hat{\mathbf{x}}_k = \mathbf{E}[\mathbf{x}_k|\mathcal{Z}_k] = \int \mathbf{x}_k p(\mathbf{x}_k|\mathcal{Z}_k) d\mathbf{x}_k. \tag{3}$$

The key to calculating the conditional pdf $p(\mathbf{x}_k|\mathcal{Z}_k)$ is Bayes theorem, the recursive formulas for the estimation of the pdf $p(\mathbf{x}_k|\mathcal{Z}_k)$ are formed by the following two steps, (e.g., [17] and [25]).

1) *Prediction*: Assuming knowledge of the posterior pdf for the state at time $k-1$: $p(\mathbf{x}_{k-1}|\mathcal{Z}_{k-1})$, the one-step ahead predictive pdf at time $k-1$, $p(\mathbf{x}_k|\mathcal{Z}_{k-1})$ can be obtained by

$$p(\mathbf{x}_k|\mathcal{Z}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathcal{Z}_{k-1}) d\mathbf{x}_{k-1} \tag{4}$$

   where $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is defined by (1) and the known statistics of $\mathbf{w}_{k-1}$.

2) *Filtering*: Based on predictive pdf $p(\mathbf{x}_k|\mathcal{Z}_{k-1})$, the posterior pdf at time k given measurement y_k, $p(\mathbf{x}_k|\mathcal{Z}_k)$ can be computed *via* Bayes rule

$$p(\mathbf{x}_k|\mathcal{Z}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathcal{Z}_{k-1})}{\int p(\mathbf{y}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathcal{Z}_{k-1}) d\mathbf{x}_k} \tag{5}$$

   where the conditional pdf $p(\mathbf{y}_k|\mathbf{x}_k)$ is defined by (2) and the known statistics of $\mathbf{v}_k$.

The above equations for Bayes recursive estimation (4) and (5) can only be analytically solved for a small class of problems, the most important example of such a class of problems is that with linear system and measurement equations, and Gaussian additive noise, in which the pdf can be summarized by the mean and covariance. Then, the Kalman filter is used to propagate and update the mean and covariance of the pdf. For general nonlinear, non-Gaussian systems described by (1) and (2), there is no simple way to proceed.

PFs were proposed as a new way of representing and recursively generating an approximation to the conditional pdf $p(\mathbf{x}_k|\mathcal{Z}_k)$ [13], [16]. The key idea is to represent the required pdf by a swarm of points called "particles," rather than by a function over the state space. For example, the predictive pdf $p(\mathbf{x}_k|\mathcal{Z}_{k-1})$ is expressed or approximated by a set of N particles $\{\mathbf{x}_{k|k-1}(i)\colon i=1,\ldots,N\}$, and $p(\mathbf{x}_k|\mathcal{Z}_k)$ is approximated by a set of N particles $\{\mathbf{x}_k(i)\colon i=1,\ldots,N\}$. These particles can be considered as the realizations or random samples from the required pdfs and, as the number of particles increases, they effectively provide a good approximation to the required pdf.

It can be shown that these particles can be obtained recursively by the following filtering algorithm [13], [16].

1) Assume that there is a set of random samples (particles) $\{\mathbf{x}_{k-1}(i)\colon i=1,2,\ldots,N\}$ from the pdf $p(\mathbf{x}_{k-1}|\mathcal{Z}_{k-1})$.
2) Prediction: Sample N values $\{\mathbf{w}_{k-1}(i)\colon i=1,2,\ldots,N\}$ from the pdf of system noise $\mathbf{w}_{k-1}$. Use these samples to generate new swarm of points $\{\mathbf{x}_{k|k-1}(i)\colon i=1,2,\ldots,N\}$ which approximate the predicted pdf $p(\mathbf{x}_k|\mathcal{Z}_{k-1})$ where

$$\mathbf{x}_{k|k-1}(i)=\mathbf{f}_{k-1}(\mathbf{x}_{k-1}(i),\mathbf{w}_{k-1}(i)). \tag{6}$$

3) Update: Assign each $\mathbf{x}_{k|k-1}(i)$ a weight $w_k(i)$ for $i=1,2,\ldots,N$, after measurement $\mathbf{y}_k$ is received. The weights are given by

$$w_k(i)=\frac{p(\mathbf{y}_k|\mathbf{x}_{k|k-1}(i))}{\sum_{j=1}^{N}p(\mathbf{y}_k|\mathbf{x}_{k|k-1}(j))}. \tag{7}$$

This defines a discrete distribution over $\{\mathbf{x}_{k|k-1}(i)\colon i=1,2,\ldots,N\}$, which assigns probability mass $w_k(i)$ to the element $\mathbf{x}_{k|k-1}(i)$ and results in the posterior pdf $p(\mathbf{x}_k|\mathcal{Z}_k)$ being represented in terms of weighted samples (particles).
4) Resample: Resample independently N times from the above discrete distribution. The resulting particles $\{\mathbf{x}_k(i)\colon i=1,2,\ldots,N\}$ which satisfies

$$\Pr\{\mathbf{x}_k(i)=\mathbf{x}_{k|k-1}(j)\}=w_k(j) \qquad \text{for all } i \tag{8}$$

form an appropriate sample (with equal weight to each element) from the posterior pdf $p(\mathbf{x}_k|\mathcal{Z}_k)$.
5) The prediction, update and resample steps form a single iteration and is recursively applied at each time k.

IV. FDI Via Particle Filtering and Likelihood Ratio Approach

The starting point for the LR approach is the logarithm of the likelihood ratio (LLR), which is a function of random variable y, defined by

$$s(y)=\ln\frac{p_{\theta_1}(y)}{p_{\theta_0}(y)} \tag{9}$$

where $p_{\theta_i}(y)(i=0,1)$ is a pdf parameterized by θ_i. The key statistical property of this ratio is as follows [2]. Let $\mathbf{E}_{\theta_0}$ and $\mathbf{E}_{\theta_1}$ denote the expectations of the random variables with distributions p_{θ_0} and p_{θ_1} respectively, then

$$\mathbf{E}_{\theta_0}(s)<0 \quad \text{and} \quad \mathbf{E}_{\theta_1}(s)>0.$$

In other words, any change in parameter θ is reflected as a change in the sign of the mean value of the LLR. If the observations $y_k(k=1,2,\ldots)$ with a pdf $p_\theta(y)$ are independent of each other, the joint LLR for the observations from y_j to y_k can be expressed as

$$S_j^k=\sum_{i=j}^{k}s_i \quad \text{and} \quad s_i=\ln\frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)}. \tag{10}$$

Suppose $\theta=\theta_0$ before change, and $\theta=\theta_1\neq\theta_0$ after change, then the typical behavior of this joint or cumulative LLR S_1^k shows, on average, a negative drift before change, and a positive drift after change. This behavior can be used for detecting any change between two known pdf p_{θ_0} and p_{θ_1}, and several detection algorithms based on the LLR test have been developed, see, e.g., [2], [7], and [30].

A. Change Detection Based on LLR

To develop the particle filtering based LLR approach to FDI, let us first consider a simple change detection problem in observation $y_k(k=1,2,\ldots)$. The detection problem, given the observations up to time k, consists of testing between two hypotheses which can be written as

No change hypothesis $\mathbf{H}_0\colon j>k$
Change hypothesis $\mathbf{H}_1\colon j\leq k$

where j is the unknown change time. The LLR between these two hypotheses is defined by (10). Replacing the unknown change time j by its maximum likelihood estimate (MLE) under $\mathbf{H}_1$, i.e.,

$$\hat{j}_k=\arg\max_{1\leq j\leq k}\left[\prod_{i=0}^{j-1}p_{\theta_0}(y_i)\prod_{i=j}^{k}p_{\theta_1}(y_i)\right]=\arg\max_{1\leq j\leq k}S_j^k \tag{11}$$

the following change detector can be obtained

$$g_k\triangleq S_{\hat{j}_k}^k=\max_j S_j^k \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \lambda \tag{12}$$

where g_k is the decision function and $\lambda>0$ is a threshold. In other words, decide $\mathbf{H}_1$ whenever g_k exceeds λ, and $\mathbf{H}_0$ otherwise. The fault alarm is set at the time t_a determined by

$$t_a=\min\{k\colon g_k>\lambda\}=\min\left\{k\colon \max_{1\leq j\leq k}S_j^k>\lambda\right\} \tag{13}$$

and the MLE of change onset time t_0 after a change has been detected is equal to the time j at which the maximum in (12) is reached. This estimate can be computed as

$$\hat{t}_0=\arg\max_{1\leq j\leq t_a}S_j^{t_a}. \tag{14}$$

If the parameter θ_1 after change is unknown, then the cumulative LLR defined in (10) is a function of two unknown independent parameters, namely the unknown change time j and the value of the parameter θ_1 after change. In this case, (10) should be written as

$$S_j^k = S_j^k(\theta_1) = \sum_{i=j}^{k} \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)}. \tag{15}$$

One of the solution to the above change detection problem is to replace θ_1 by its MLE which results in the generalized likelihood ratio (GLR) algorithm. Thus, the decision function of the GLR change detector, which involves the double maximization, is given by

$$g_k = \max_{1\le j\le k} \sup_{\theta_1} S_j^k(\theta_1). \tag{16}$$

The corresponding detection rule (or the way to determine the alarm time t_a) is the same as in (13), and the conditional MLEs of the parameter θ_1 and change onset time t_0 after detection are given by

$$(\hat{t}_0, \hat{\theta}_1) = \arg \max_{1\le j\le t_a} \sup_{\theta_1} S_j^{t_a}(\theta_1). \tag{17}$$

B. Particle Filtering-Based LLR Approach to FDI

Willsky and Jones [30] combined the LLR test with Kalman filter and developed the so-called *generalized likelihood ratio* (GLR) method for detecting and isolating the faults which are modeled as additive changes in linear dynamical systems described by linear state space models. Motivated by this, here we combine the LLR test with PF and present a particle filtering based LLR method for FDI in general nonlinear non-Gaussian systems.

As stated in Section II, the actual system is governed by one of $M+1$ models given by (1) and (2), where $m = 0$ corresponds to normal operation situation and $m = 1, \ldots, M$ correspond to M faulty situations respectively, and a PF described in previous section is implemented for each of these models to obtain the sample-based posterior pdf of the states. For simplicity, it is also assumed that the measurement noise $\mathbf{v}$ has the same dimensionality as the measurement $\mathbf{y}$ and, for each model, given the state $\mathbf{x}_k^{(m)}$ and the measurement $\mathbf{y}_k$, the measurement noise $\mathbf{v}_k^{(m)}$ is uniquely determined by $\mathbf{v}_k^{(m)} = \mathbf{g}_k^{(m)}(\mathbf{y}_k, \mathbf{x}_k^{(m)})$, where $\mathbf{g}_k^{(m)}(\cdot, \cdot)$ is the vector-valued inverse function of $\mathbf{h}_k^{(m)}(\cdot, \cdot)$ with respect to $\mathbf{v}_k^{(m)}$ and has a *Jacobian* denoted by $\partial \mathbf{g}_k^{(m)}/\partial \mathbf{y}_k$.

The key idea of our method is to compute the joint likelihood of the observations conditional on each hypothesized model through Monte–Carlo estimation which uses the complete sample-based pdf information provided by PF, and then activating in parallel M LLR tests for $\mathbf{H}_m (m = 1, 2, \ldots, M)$ versus $\mathbf{H}_0$. More specifically, the joint LLR to be computed in the present case is as follows.

$$S_j^k(m) = \sum_{r=j}^{k} \ln \frac{p(\mathbf{y}_r|\mathbf{H}_m, \mathcal{Z}_{r-1})}{p(\mathbf{y}_r|\mathbf{H}_0, \mathcal{Z}_{r-1})} \tag{18}$$

where the likelihood of the observation $\mathbf{y}_r$ give its past values $\mathcal{Z}_{r-1}$, i.e., $p(\mathbf{y}_r|\mathbf{H}_m, \mathcal{Z}_{r-1}) (m = 0, 1, \ldots, M)$ is precisely the one step output prediction density based on $\mathbf{H}_m$ which is defined by the mth measurement model and the known statistics of $\mathbf{v}_r^{(m)}$. If the pdf of $\mathbf{v}_r^{(m)}$ is denoted by $q_r^{(m)}(\mathbf{v}_r^{(m)})$, the $p(\mathbf{y}_r|\mathbf{H}_m, \mathcal{Z}_{r-1})$ can then be expressed as

$$\begin{aligned} l_r^{(m)} &= p(\mathbf{y}_r|\mathbf{H}_m, \mathcal{Z}_{r-1}) \qquad (m = 0, 1, \ldots, M) \\ &= q_r^{(m)}\left(\mathbf{g}_r^{(m)}\left(\mathbf{y}_r, \mathbf{x}_{r|r-1}^{(m)}\right)\right) \left|\frac{\partial \mathbf{g}_r^{(m)}}{\partial \mathbf{y}_r}\right| \end{aligned} \tag{19}$$

where $\mathbf{x}_{r|r-1}^{(m)}$ is the one step state prediction given $\mathcal{Z}_{k-1}$ and based on mth model. In the linear Gaussian case, the quantity defined by (19) is just the innovation likelihood which can be derived from the Kalman filter equations based on the mth model [27]. For the general nonlinear non-Gaussian model (1) and (2), there is no general analytical means to perform the calculation. However, with the PF, this quantity can be estimated by utilizing the complete pdf information of the predicted state $\mathbf{x}_{r|r-1}^{(m)}$ represented by a swarm of particles, this is achieved by reusing the likelihood of each predicted state particle computed during particle filtering, more specifically, since $\{\mathbf{x}_{r|r-1}^{(m)}(i)\colon i = 1, \ldots, N\}$ can be considered as the samples from $p(\mathbf{x}_r|\mathbf{H}_m, \mathcal{Z}_{r-1})$, the required quantity defined by (19) can be computed *via* the Monte–Carlo integration as follows:

$$\begin{aligned} l_r^{(m)} &= p(\mathbf{y}_r|\mathbf{H}_m, \mathcal{Z}_{r-1}) \\ &= \int p(\mathbf{y}_r|\mathbf{H}_m, \mathbf{x}_r) p(\mathbf{x}_r|\mathbf{H}_m, \mathcal{Z}_{r-1})\, d\mathbf{x}_r \\ &\approx \frac{1}{N} \sum_{i=1}^{N} p\left(\mathbf{y}_r|\mathbf{x}_{r|r-1}^{(m)}(i)\right) \end{aligned} \tag{20}$$

where the likelihood of each predicted state sample $\mathbf{x}_{r|r-1}^{(m)}(i) (i = 1, \ldots, N)$ from PF is given by

$$p\left(\mathbf{y}_r|\mathbf{x}_{r|r-1}^{(m)}(i)\right) = q_r^{(m)}\left(\mathbf{g}_r^{(m)}\left(\mathbf{y}_r, \mathbf{x}_{r|r-1}^{(m)}(i)\right)\right) \left|\frac{\partial \mathbf{g}_r^{(m)}}{\partial \mathbf{y}_r}\right| \quad (i = 1, 2, \ldots, N). \tag{21}$$

The decision function for FD is then given by

$$g_k = \max_{1\le j\le k} \max_{1\le m\le M} S_j^k(m). \tag{22}$$

The fault alarm time t_a is determined by (13) where the threshold $\lambda > 0$ is chosen to provide a reasonable tradeoff between false and missing alarms. Fault isolation is achieved by finding out the faulty model index m which, along with the MLE $\hat{t}_0$ of fault onset time, is given by

$$(\hat{m}, \hat{t}_0) = \arg \max_{1\le j\le t_a} \max_{1\le m\le M} S_j^{t_a}(m). \tag{23}$$

The full implementation of the above particle filtering-based LLR detector requires a linearly growing number of calculations, as $S_j^k(m)$ must be calculated for $m = 1, \ldots, M$, and *all* possible fault onset times up to the present, i.e., $j = 1, \ldots, k$. The standard method to avoid this problem is to constrain the

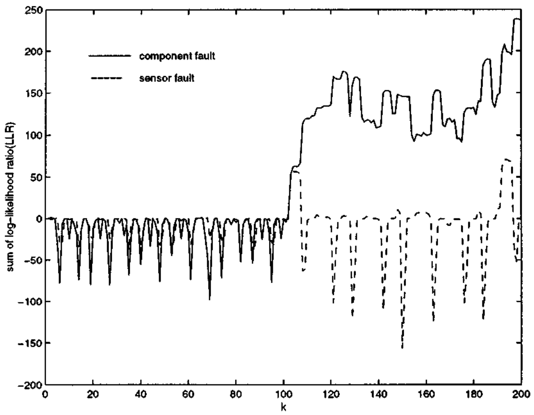

Fig. 1. Sum of LLR computed with the new algorithm proposed in this paper.

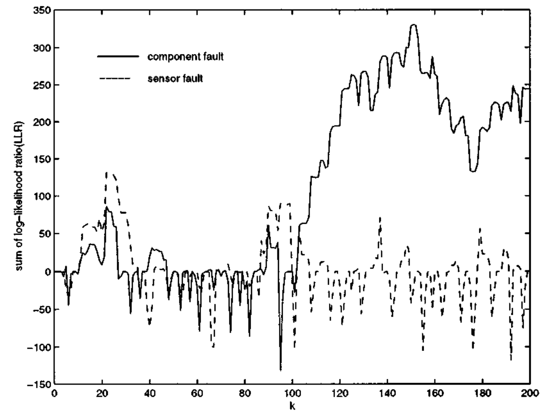

Fig. 2. Sum of LLR computed with the EKF-based GLR algorithm.

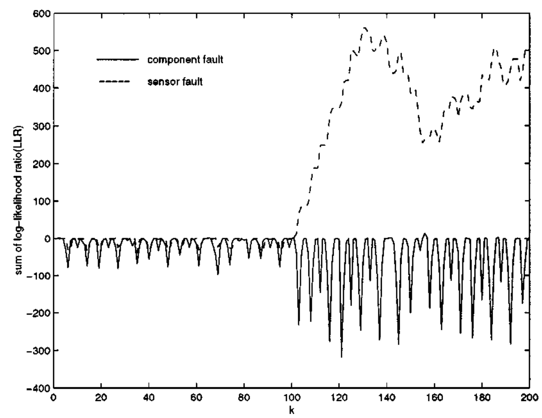

Fig. 3. Sum of LLR computed with the new algorithm proposed in this paper.

search in a fixed width (say W) "sliding window" of the most recent past observations, the decision function is then given by

$$g_k = \max_{k-W+1 \le j \le k} \max_{1 \le m \le M} S_j^k(m). \tag{24}$$

If the window is sufficiently wide to insure detection and identification of all important faults, this approximation avoids the aforementioned difficulty.

It is most difficult to make any precise and provable statement on the optimal properties of the above proposed method in the general nonlinear non-Gaussian case. However, for a closely related change detector, the window-limited GLR detector introduced by Willsky and Jones [30] in the context of detecting additive changes in linear Gaussian state space models, Lai and Shan [18] have recently shown that, with an appropriate choice of the threshold and the window size, this GLR detector is asymptotically optimal. The detector proposed in this paper can be considered as an extention of the Willsky–Jones' GLR detector to the general nonlinear non-Gaussian case through particle filtering and Monte-Carlo integration.

V. Numerical Example and Simulation Results

A. System Description

To illustrate the operation of the proposed particle filtering-based LLR algorithm for FDI developed in this paper, an example is presented in this section. In the following simulation study, the data are artificially generated by a set of $M+1$ univariate state space models described below with different types of system and measurement noise

$$\begin{aligned} x_k &= \frac{1}{2} x_{k-1} + a_1^{(m)} \frac{x_{k-1}}{(1+x_{k-1}^2)} + 8\cos(1.2(k-1)) + w_{k-1} \\ y_k &= a_2^{(m)} x_k^2 + v_k \qquad (m = 0, 1, \ldots, M) \end{aligned} \tag{25}$$

where w_k and v_k are system and measurement noise respectively and they are assumed to be uncorrelated. The parameter values for nominal system model (indexed by $m = 0$) are $a_1^{(0)} = 25$ and $a_2^{(0)} = 0.05$ which were taken from [13] and [17]. Two kinds of fault are considered (i.e., $M = 2$), the component fault (indexed by $m = 1$) is modeled by a jump in the parameter of system state equation, in which $a_1^{(0)}$ is shifted to $a_1^{(1)} = (1/2)a_1^{(0)} = 12.5$ while the parameter in measurement equation remains unchanged, i.e., $a_2^{(1)} = a_2^{(0)} = 0.05$. The sensor fault ($m = 2$) is modeled as a jump in the parameter of measurement equation, in which $a_2^{(0)}$ is shifted to $a_2^{(2)} = 2a_2^{(0)} = 0.1$ while $a_1^{(2)} = a_1^{(0)} = 25$. Two types of simulation are conducted. In the first type of simulations, w_k and v_k are assumed to be Gaussian so as to facilitate the performance comparison between the proposed method and the EKF based GLR method, whereas in the second type of simulations, w_k and v_k are assumed to be non-Gaussian which demonstrate the broad applicability of the proposed method.

B. Experimental Results with Gaussian Noise

In this type of simulations, w_k and v_k are assumed to be zero mean Gaussian white noise with variance $Q_w = 0.1$ and $Q_v = 1$, respectively. Two Monte–Carlo simulation experiments have been carried out. In the first experiment, the component fault is simulated to occur at time $k = 101$ at which the system model is shifted from $m = 0$ (nominal model) to $m = 1$ (component fault). In the second experiment, the sensor fault is simulated to occur at time $k = 101$ at which the system model is shifted from $m = 0$ to $m = 2$ (sensor fault). The PF-based LLR algorithm for FDI developed in this paper is used to detect and isolate these two faults. For comparison, the well-known EKF-based GLR method [27], [28] is also applied.

The sum of LLR computed by PF-based method and by EKF-based method in these two experiments are shown in Figs. 1 and 3 and Figs. 2 and 4, respectively. In the calculations of the deci-

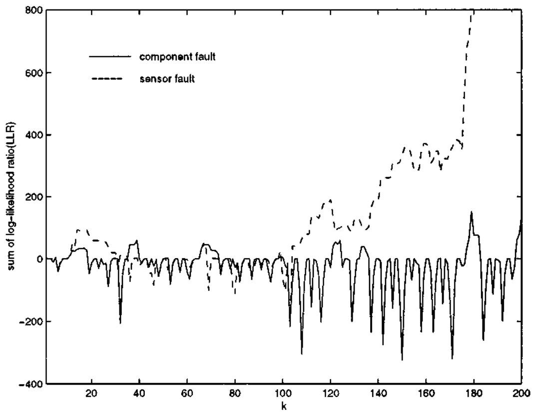

Fig. 4. Sum of LLR computed with the EKF-based GLR algorithm.

sion function g_k, the search for the discrete maximization over the fault onset time j is constrained within a sliding window of fixed width $W = 25$. The results from these simulations show that the PF-based algorithm is able to detect the faults in time and to identify the faults correctly. We can see, from Figs. 1 and 3, the sum of LLR computed with the new algorithm proposed in this paper remains steady below zero before the onset of the fault, and jumps significantly over zero after the fault occurs. While the sum of LLR computed by EKF-based algorithm (see Figs. 2 and 4) changes frequently over zero before the fault actually takes place which indicates a high false alarm rate in fault detection.

In addition to the correct fault detection and isolation, the new method is also able to give a relatively accurate estimate of fault onset time $\hat{t}_0$. In our simulation experiments, the threshold for decision making is chosen as $\lambda = 10$, with the new method, the component fault is detected at $t_a = 108$ in the first experiment, and the sensor fault is detected at $t_a = 102$ in the second experiment, the estimates of the fault onset time are $\hat{t}_0 = 104$ in first experiment and $\hat{t}_0 = 101$ in the second experiment (true fault onset times in both experiments are $t_0 = 101$).

C. Experimental Results With Non-Gaussian Noise

Simulations are now performed with non-Gaussian system noise w_k and non-Gaussian measurement noise v_k. Two Monte–Carlo simulation experiments have been carried out, both with the component fault occurring at time $k = 101$ as described previously. In the first experiment, w_k and v_k are assumed to have the following distributions:

$$w_k \sim 0.5\mathcal{N}(0, 0.1) + 0.5\mathcal{N}(0, 1) \quad v_k \sim \mathcal{N}(0, 1)$$

where w_k is a mixture of Gaussian noise. The sum of LLR computed with the new algorithm is shown in Fig. 5. In the second experiment, w_k and v_k are distributed as follows:

$$w_k \sim \mathcal{N}(0, 0.1) \quad v_k \sim 0.5\mathcal{N}(0, 1) + 0.5\mathcal{N}(0, 10)$$

where v_k is a mixture of Gaussian noise. The corresponding sum of LLR is shown in Fig. 6. These figures clearly shown that, in both cases, the fault can be detected and isolated with the new method. In the implementation of both simulation experiments, the threshold for decision making is chosen as before ($\lambda = 10$),

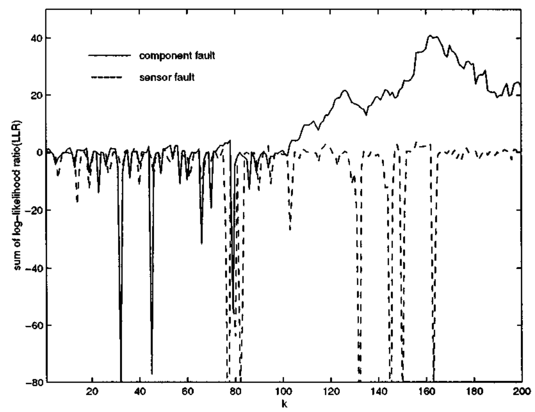

Fig. 5. Sum of LLR computed with a non-Gaussian (Gaussian mixture) system noise w_k.

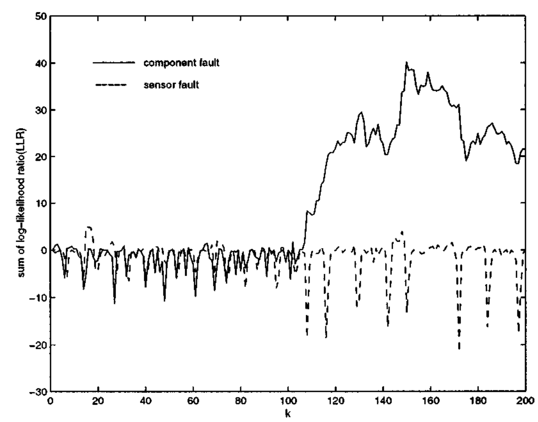

Fig. 6. Sum of LLR computed with a non-Gaussian (Gaussian mixture) measurement noise v_k.

the component fault is detected at $t_a = 112$ in both experiments, the estimates of the fault onset time are $\hat{t}_0 = 103$ in first experiment and $\hat{t}_0 = 105$ in the second experiment.

VI. Conclusion

By combining the particle filtering algorithm with the log-likelihood ratio test in the multiple model environment, we have proposed a new FDI scheme for general nonlinear non-Gaussian dynamic systems, the FDI performance of the proposed methods is compared with that of the well-known EKF-based GLR method on a highly nonlinear system. The results from simulation experiments show clearly that the FDI performance of this new method is superior to the EKF-based one. This result stems from the fact that the complete pdf information of the estimated state is utilized for FDI in the new method, whereas, only *approximate* mean and covarance are used in EKF-based method. Furthermore, the proposed schemes provide an uniform framework for FDI in general nonlinear systems with non-Gaussian noise and disturbance. Further work is being carried out to investigate the robustness with the proposed method.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[2] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes—Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, Apr. 1993.
[3] M. Basseville, "Statistical approaches to industrial monitoring problems—Fault detection and isolation," in *Proc. IFAC Syst. Identification*, Fukuoka, Japan, 1997, pp. 413–432.
[4] ——, "Detecting changes in signals and systems—A survey," *Automatica*, vol. 24, no. 3, pp. 309–326, 1988.
[5] M. Basseville and A. Benveniste, "Detection of abrupt changes in signals and dynamical systems," in *Lecture Notes in Control and Information Sciences*. Berlin, Germany: Springer-Verlag, Dec. 1985, vol. 77.
[6] E. Bolviken *et al.*, "Monte–Carlo filters for nonlinear state estimation," *Automatica*, vol. 37, no. 2, pp. 177–183, 2001.
[7] T. T. Chien and M. B. Adams, "A sequential failure detection technique and its application," *IEEE Trans. Automat. Control*, vol. AC-21, pp. 750–757, Oct. 1976.
[8] A. Doucet, S. Godsill, and C. Ardrieu, "On sequential Monte–Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, pp. 197–208, 2000.
[9] P. M. Frank, "Analytical and qualitative model-based fault diagnosis—A survey and some new results," *Eur. J. Contr.*, vol. 1, no. 2, pp. 26–28, 1996.
[10] ——, "On-line fault detection in uncertain nonlinear systems using diagnostic observers: A survey," *Int. J. Syst. Sci.*, vol. 25, no. 12, pp. 2129–2154, 1994.
[11] ——, "Fault diagnosis in dynamic systems using analytical and knowledge based redundancy—A survey and new results," *Automatica*, vol. 26, pp. 459–474, 1990.
[12] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Syst. Mag.*, vol. 8, pp. 3–11, Dec. 1988.
[13] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *Proc. Inst. Elect. Eng. F*, vol. 140, no. 2, pp. 107–113, 1993.
[14] R. Isermann, "Process fault detection based on modeling and estimation methods—A survey," *Automatica*, vol. 20, pp. 387–404, 1984.
[15] V. Kadirkamanathan *et al.*, "A sequential Monte–Carlo filtering approach to fault detection and isolation in nonlinear systems," in *Proc. IEEE Conf. Decision Control*, Sydney, Australia, Dec. 2000, pp. 4341–4346.
[16] G. Kitagawa, "Monte–Carlo filter and smoother for non-Gaussian nonlinear state space models," *J. Computat. Graph. Statist.*, vol. 5, no. 1, pp. 1–25, 1996.
[17] ——, "Non-Gaussian state-space modeling of nonstationary time series (with discussion)," *J. Amer. Statist. Assoc.*, vol. 82, no. 400, pp. 1032–1063, 1987.
[18] T-L. Lai and J. Z. Shan, "Efficient recursive algorithms for detection of abrupt changes in signals and control systems," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 952–966, Oct. 1999.
[19] D. G. Lainiotis, "Partitioning: A unifying framework for adaptive system, I: Estimation," *Proc. IEEE*, vol. 64, pp. 1126–1142, Aug. 1976.
[20] P. Li and V. Kadirkamanathan, "Particle filtering based multiple-model approach to fault diagnosis in nonlinear stochastic systems," in *Proc. Euro. Control Conf (ECC)*, Porto, Portugal, Sept. 2001, pp. 1378–1383.
[21] R. K. Mehra and J. Peschon, "An innovations approach to fault detection and diagnosis in dynamic systems," *Automatica*, vol. 7, pp. 637–640, 1971.
[22] R. J. Patton and J. Chen, "Robust fault detection and isolation (FDI) systems," in *Control and Dynamic Systems*. New York: Academic, 1996, vol. 74, pp. 171–224.
[23] R. J. Patton, P. M. Frank, and R. Clark, *Fault Diagnosis in Dynamic Systems—Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[24] A. D. Pouliezos and G. S. Stavrakakis, *Real Time Fault Monitoring for Industrial Processes*. Boston, MA: Kluwer, 1994.
[25] H. W. Sorenson, "Recursive estimation for nonlinear dynamic systems," in *Bayesian Analysis of Time Series and Dynamic Models*, J. C. Spall, Ed. New York: Marcel Dekker, 1988.
[26] P. B. Usoro, I. C. Schick, and S. Negahdaripour, "An innovation-based methodology for HVAC system fault detection," *J. Dyn., Syst., Meas., Contr.*, vol. 107, pp. 284–289, Dec. 1985.
[27] A. S. Willsky, "Detection of abrupt changes in dynamic systems," in *Lecture Notes in Control and Information Sciences*. Berlin, Germany: Springer-Verlag, Dec. 1985, vol. 77.
[28] A. S. Willsky *et al.*, "Dynamic model-based techniques for the detection of incidents on freeways," *IEEE Trans. Automat. Contr.*, vol. AC-25, June 1980.
[29] A. S. Willsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, pp. 601–611, 1976.
[30] A. S. Willsky and H. L. Jones, "A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 108–112, Feb. 1976.
[31] Y. Zhang and X. R. Li, "Detection and diagnosis of sensor and actuator failures using IMM estimator," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 1293–1313, Oct. 1998.
[32] D. H. Zhou, Y. G. Xi, and Z. J. Zhang, "Nonlinear adaptive fault detection filter," *Int. J. Syst. Sci.*, vol. 22, pp. 2563–2571, 1991.

Ping Li received the B.S. and M.S. degrees in electrical engineering from Northwestern Polytechnical University, Xian, China, in 1983 and 1993, respectively. He is currently pursuing the Ph.D. degree at the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, U.K.

He was with the Department of Automatic Control and Systems Engineering, University of Sheffield, as a Research Visitor in 1998. His research interests include fault detection and isolation, optimal filtering, estimation and identification, time series analysis, and their applications to signal processing.

Visakan Kadirkamanathan was born in Jaffna, Sri Lanka, in 1962. He received the B.A. degree in electrical and information sciences with first class honors and the Ph.D. degree both from Cambridge University, Cambridge, U.K., in 1987 and 1992, respectively.

Since 1993, he has been a Senior Lecturer in the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, U.K. Prior to that, he held Postdoctoral Research Assistant positions at the University of Surrey, Surrey, U.K., and in the Engineering Department of Cambridge University. His research interests include adaptive estimation and control, fault detection and diagnosis, particle filters, neural networks, intelligent systems and time-series analysis for signal processing. He has published a number of technical papers and is a coauthor of the book *Functional Adaptive Control: An Intelligent Systems Approach* (New York: Springer-Verlag, 2001). He serves on the Editorial Board of the *International Journal of Condition Monitoring and Diagnostic Engineering Management*.

Dr. Kadirkamanathan is an Associate Editor for the IEEE TRANSACTIONS ON NEURAL NETWORKS and serves on the Editorial Board for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS–PART B (Electronic Section).