Stochastic Change Detection based on an Active Fault Diagnosis Approach

Poulsen, Niels Kjølstad; Niemann, Hans Henrik

Published in:
Proceedings of the IEEE Conference on Decision and Control

Link to article, DOI:
10.1109/CDC.2007.4434429

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Stochastic Change Detection based on an Active Fault Diagnosis Approach

Niels Kjølstad Poulsen
Informatics and Math. Modelling
Tech. Univ. of Denmark, Build. 321
DK-2800 Kgs. Lyngby, Denmark
nkp@imm.dtu.dk

Henrik Niemann
Ørsted•DTU, Automation
Tech. Univ. of Denmark, Build. 326
DK-2800 Kgs. Lyngby, Denmark
hhn@oersted.dtu.dk

Abstract—The focus in this paper is on stochastic change detection applied in connection with active fault diagnosis (AFD). An auxiliary input signal is applied in AFD. This signal injection in the system will in general allow to obtain a fast change detection/isolation by considering the output or an error output from the system. The classical CUSUM (cumulative sum) method will be modified such that it will be able to detect change in the signature from the auxiliary input signal in the (error) output signal. It will be shown how it is possible to apply both the gain as well as the phase change of the output vector in the CUSUM test.

I. INTRODUCTION

The area of active fault diagnosis has been considered in a number of papers [2], [3], [5], [10], [13], [15], [16], [17] and in the books, [4], [8], [19].

The AFD is based on the inclusion of an auxiliary input signal/vector into the system. As output from the system, a standard residual signal known from the passive FDI (fault detection and isolation) approach is applied, [6]. Using the AFD approach from [12], [13], the auxiliary input is injected in the closed-loop system such that the standard residual vector is decoupled from this auxiliary input in the nominal case and coupled in the case of parameter changes.

There are mainly two approaches in AFD. In the approach derived by Zhang, [19] and has later been investigated extensively in [4], [8], [10], the auxiliary input is designed with respect to a fast fault diagnosis/multi model selection. By using a dedicated design of the auxiliary input gives a fast diagnosis/selection.

The other AFD approach in [12], [13] and applied in [14], a periodic auxiliary input is applied. The signature from this periodic input in the output/residual will also be periodic with the same frequency. A change in the system due to e.g. parametric changes/faults, will give a change in the signature. This change can be in the amplitude and/or the phase of the periodic output.

Using the AFD approach from [12], [13], the auxiliary input is decoupled from the output/residual vector in the nominal case. The detection of parameter changes can then be done by a detection of a signature from the auxiliary input in the residual signal. Another approach is to use a filter/observer to estimate the periodic signature with the known frequency directly. This approach will not be considered in this paper. Instead the classical CUSUM method, [1], [7], will be applied for change detection. The CUSUM method will be modified with respect to detect changed based on the periodic auxiliary input. This modification can be done in different ways. It is possible to let the CUSUM test be based only on the amplitude/gain of the signature in the residual signal from the auxiliary input or it can be based on both the gain and the phase shift in the signature signal. Using both the gain and the phase shift for the change detection, it will also be possible to isolate change in different parameters. From a theoretical point of view, it will be possible to isolate an unlimited number of parameter changes. In practice, however, there will be an upper bound on the number of parameters that can be isolated based on a single periodic auxiliary input. This number will depend strongly on the signal/noise ratio.

Only the SISO (single input single output) case will be considered in this paper, but it is possible to extend the results to the MIMO (multi input multi output) case without any major difficulties. Further, only periodic stationary auxiliary inputs will be applied as considered in [9] in connection with AFD.

II. SYSTEM SET-UP

Let a general system be given by:

\[ \Sigma_{P\theta} : \begin{cases} e_t = G_{ed}(\theta)d_t + G_{ed}(\theta)u_t \\ y_t = G_{yd}(\theta)d_t + G_{yd}(\theta)u_t \end{cases} \quad (1) \]

where \( d_t \in \mathcal{R}^p \) is a disturbance signal vector, \( u_t \in \mathcal{R}^m \) the control input signal vector, \( e_t \in \mathcal{R}^n \) is the external output signal vector to be controlled and \( y_t \in \mathcal{R}^p \) is the measurement vector. \( \theta \) is given by

\[ \theta = (\theta_1, \ldots, \theta_i, \ldots, \theta_k)^T \]

represents the \( k \) parametric faults in the system. It will in many cases be possible to give an explicit expression of the connection between the system and the parametric faults as described in [13], [14].
Let the system be controlled by a stabilizing feedback controller given by:

\[ \Sigma_C : \{ u_t = Ky_t \} \]  

(2)

A. Coprime Factorization

Let a coprime factorization of the nominal system \( G_{y_n}(0) \) from (1) and the stabilizing controller \( K \) from (2) be given by:

\[ G_{y_n}(0) = \frac{N}{M} K = \frac{U}{V} \]  

(3)

where the four stable transfer functions in (3) must satisfy the Bezout equation given by, see [18]:

\[ MV - NU = 1 \]  

(4)

III. AFD Set-Up

Now, let’s consider the AFD set-up described in [13], [14]. The set-up is shown in Fig. 1. The diagram include the residual vector \( \varepsilon_t \) and an auxiliary input vector \( \eta_t \). The residual vector \( \varepsilon_t \) for \( \Sigma_{P_0} \) in (1) is given by:

\[ \varepsilon_t = M y_t - N u_t \]  

(5)

is the same residual generator applied in connection with the passive fault diagnosis, [6]. A more detailed discussion of the applied AFD set-up is given in [13], [14].

The transfer function from the input vector \( \eta_t \) to the residual vector \( \varepsilon_t \) is equal to the dual YJBK (after Youla, Jarb, Bongiorno and Kucera) transfer function, [12], [14]. An important thing in this connection is that the dual YJBK transfer function is equal to zero in the nominal case. In [12]-[14], the dual YJBK transfer function has been named as the fault signature matrix in connection to AFD and it will be denoted by \( S(\theta) = P_{\eta}(\theta) \) in the following. An explicit equation for \( S(\theta) \) has been derived in [11] - [13].

IV. Change Detection

The implementation of the AFD set-up is not unique, because the coprime factorization is not unique. This will give an extra freedom in the implementation of the AFD set-up. It is e.g. possible to based the coprime factorization of \( G_{y_n} \) on a Kalman filter. This will give a residual vector with well defined disturbances properties in the nominal case. The design freedom introduced in connection with the coprime factorization of the system and controller will not be discussed further in this paper. It will be assume that the residual vector is the innovation vector from a Kalman filter.

In a passive scheme the detection is often based on a change in the statistics (mean, variance, correlation or spectral properties) of the residual sequence. In an active scheme an auxiliary signal is introduced and the residual vector in (6) takes then the following form:

\[ \varepsilon_t = S(\theta) \eta_t + \xi_t \]  

(8)

where \( \xi_t \in N(0, \sigma^2_\xi) \) in the nominal case. From (7), [12]-[14], we have that \( S(\theta) \) is zero in the nominal case, i.e.

\[ S(0) = 0 \]  

(9)

It is clear from this observation, that \( S(\theta) \) is very important in connection with active change detection (or active fault detection). A direct consequence of (9) is the following condition:

\[ S(\theta) = 0 \quad \text{for} \quad \theta = 0 \]

\[ S(\theta) \neq 0 \quad \text{for} \quad \theta \neq 0 \]  

(10)

The detection (isolation) of parameter variations can then be based on the following hypothesis, [1], [7]:

\[ H_0 : S(\theta) = 0 \]

\[ H_1 : S(\theta) \neq 0 \]  

(11)

Test of the above two hypothesis can be done by on-line evaluation of the residual signal with respect to the signature from the auxiliary input in the residual \( \varepsilon_t \). Consequently the auxiliary signal is chosen such that the signature in the residual is distinctive. This is contradiction to methods in which the objective is a change in statistics (normally the mean and variance) of the residual. For this (and others explained later) reasons the auxiliary input is chosen as being a periodic signal given by

\[ \eta_t = a_\eta \sin(\omega_\eta t) \]  

(12)
where the amplitude $a_\eta$ and the frequency $\omega_0$ are the tuning parameters in the auxiliary input. The specific signature in the residual of this signal is particular easy. Using the auxiliary input given by (12), the residual signal is given by

$$\xi_t = \xi + \xi \in N(0, \sigma^2)$$

in the nominal case and

$$\xi_t = a_\eta|S| \sin(\omega_0 t + \phi) + \xi \in N(m, \sigma^2)$$

in the case of parameter changes. For brevity we have omitted the dependency of $\theta$ and $\omega_0$ in $S = S(\theta, \omega_0)$, $\phi = \phi(\theta, \omega_0)$, $m = m(\theta)$ and $\sigma_1 = \sigma_1(\theta)$, $m$ will in general be zero. Both the amplitude and the phase of the periodic signal in $\xi_t$ depend on $\theta$ and on the chosen frequency, $\omega_0$. The periodic signal in $\xi_t$ is the signature of the periodic auxiliary input $\eta_t$.

Detection of parameter changes is then based on a detection of the signature from $\eta_t$ in $\xi_t$. Further, isolation of parameter changes might be possible from an investigation of the amplitude and phase of the signature in $\xi_t$. It might be needed to include more than a single periodic signal in $\eta_t$ to be able to isolate different parameter changes. This is the case when the amplitude of $S(\theta)$ is insensitive to change of one parameter and sensitive to change of another parameter at a specific frequency. Here, we will only consider a single periodic auxiliary input signal due to the fact that at the chosen $\omega_0$ it is possible to isolate changes in the system parameters by also using the phase of $S$.

The selection of $a_\eta$ and $\omega_0$ need to done with respect to a number of conditions. The choice of amplitude is given by the tolerated increase in power in $\xi_t$ due to the auxiliary signal in the normal case. It is clear that a higher amplitude will increase the speed of detection and enable the algorithm to detect smaller parameter changes. The selection of $\omega_0$ need to be done with respect to the following conditions:

1) Maximize the signal to noise ratio between the signature from $\eta_t$ on the residual signal $\xi_t$ in the faulty case and the effect from disturbance input $d_t$ in the residual signal.
2) Minimize the effect from $\eta_t$ on the external output $e_t$ in the fault free case.
3) The selection of the frequency must be done with respect to obtain (if possible) a possibility for discriminate the different type of changes from the signature from $\eta_t$ in residual signal $\xi_t$.

Using the closed-loop transfer function from the AFD set-up, the above conditions can be formulated as follows:

1) is equivalent with: $\max_\omega \left[ \frac{\left| S(\theta, \omega) \eta_t \right|}{\left| P_{l\eta}(\theta, \omega) \right|} \right]$.
2) is equivalent with: $\min_\omega \left| P_{l\eta}(\theta, \omega) \right| \eta_t$.
3) specifies that the signature from $\eta_t$ in $\xi_t$ given by (14) must be different for different faults. This is satisfied if the amplitude and the phase change of the periodic signature in $\xi_t$ is unique for the single faults.

Condition 1) and 2) are related to fault detection whereas 3) is only related to fault isolation.

### A. Parameter Change Detection

Assume that an auxiliary input signal has been selected, i.e. the amplitude $A$ and the frequency $\omega_0$ has been specified. In order to detect if the signature of the auxiliary signal is present in the residual, the following two signals are formed:

$$s_t = \xi_t \sin(\omega_0 t) \quad c_t = \xi_t \cos(\omega_0 t)$$

where according to (14)

$$s_t = |S|^2 \left( -\frac{\cos(\phi)}{2} \right) + \xi_t \sin(\omega_0 t)$$

$$c_t = |S|^2 \left( \frac{\sin(\phi)}{2} \right) + \xi_t \cos(\omega_0 t)$$

From this it is clear that in the fault free situation

$$s_t = \xi_t \sin(\omega_0 t) \in N(0, \sigma^2 \sin^2(\omega_0 t))$$

$$c_t = \xi_t \cos(\omega_0 t) \in N(0, \sigma^2 \cos^2(\omega_0 t))$$

Additionally, the two signals are (when a filter parameterization is applied) white. The time average variance is equal to $3/2 \sigma^2_0$.

In a faulty situation the fault signature matrix, $S$, is different from zero and the two detection signals, $s_t, c_t$, will have a deterministic component

$$m(S(\theta), a_\eta) = \begin{bmatrix} m_s \\ m_c \end{bmatrix} = \left| S \right|^2 \frac{a_\eta}{2} \left[ \frac{\cos(\phi)}{\sin(\phi)} \right]$$

This component can be used for fault detection.

Besides the mentioned component the detector signals will have a time varying deterministic component

$$|S|^2 \left( -\frac{\cos(2\omega_0 t + \phi)}{\sin(2\omega_0 t + \phi)} \right)$$

The effect of this component can be eliminated by means of an average or integration technique such as in the CUSUM methodology.

The detection can be implemented as a CUSUM detection given by

$$z_{t+1} = \max(0, z_t + \left[ \frac{\delta_t}{\sigma^2 + \frac{1}{2} \gamma} \right])$$

where

$$\delta_t = \begin{bmatrix} s_t \\ c_t \end{bmatrix}, \quad \sigma^2 = \frac{1}{2} \sigma^2_0$$

The $H_0$ is accepted if $z_t$ is smaller than the threshold, i.e. $z_t \leq \frac{\log(B)}{\gamma} = h$.
where the inequality is to be understood element wise. The tuning parameters in this CUSUM detector is $B$, which is related to the average length between false detections, $\gamma$ which forms a typical lower limit of changes to be detected. The latter quantity is off course related to the lower limits of detection for the individual parameter changes. Furthermore note that the time average variance of $c_i$ and $s_j$ has been used in (16).

The time distance from the last zero crossing of the elements in $z_j$ forms an estimate of the fault time instance, $T_d$.

**B. Parameter Change Isolation**

The phase information can be utilized in the process of isolate the type of parameter changes. For each type of error (and for fixed $\omega_0$), the fault signature matrix, $S(\omega_0, \theta)$ forms, as illustrated in Fig. 2 (for a two parameter problem) a curve in the complex plane which pass trough the origin for $\theta = 0$. Let us for brevity denotes these as fault curves. The fault isolation can then be done by estimating the fault signature matrix, $S(\omega_0, \theta)$ and match with the possible values. However, due to stochastic disturbances an estimate of $S$ will inherently be uncertain. Instead the estimate should be matched with the nearest (e.g. in a least squares sense) fault curve. These curves will then divide the complex plane into double conic areas each related to each type of parameter change. The isolation procedure will then be a classification in which of the areas an estimate of the $S$ belong to.

This classification can for small parameter change easily be mechanized by assigning a (unit) vector, $v_i$, $i \in \{1, \ldots, k\}$, to each type of parameter changes. The vectors are parallel to the tangent of $S(\omega_0, \theta)$ in origin. Let $T_d$ denoted the estimate of the fault time. The vector

$$v = \int_{T_d}^{\tau} \left[ \begin{array}{c} s_T \\ c_T \end{array} \right] d\tau$$

(17)

is an estimate of $S$. The classification is then simply to find the maximal projection among the considered types of parameter changes, i.e.

$$i = \arg \max_{i \in \{1, \ldots, k\}} v^T v_i$$

(18)

**V. Evaluation of Fault Detectors**

It is relevant to evaluate the fault detectors based on the AFD by using a number of the standard performance measures. Some of these performance measures are: mean time between false alarms (MTFA) (or similar false alarm rate (FAR)) and mean time to detect (MTD). These performance measures can be determined from the average run length (ARL) function, which in general cannot be calculated exactly. Instead, approximations of the performance measures can be derived, see e.g. [1], [7].

![Fig. 2. For each type of errors (and for fixed $\omega_0$), the fault signature matrix, $S(\omega_0, \theta)$ forms a curve (shown dashed lines) in the complex plane. The individual type of parameter changes forms a double coned area (shown shaded for type 2 of parameter change).](image)

Let $\mu_\alpha$ and $\sigma_\alpha^2$ be the mean an the variance of each of the component in the increment

$$\alpha_t = \frac{\delta_t}{\sigma_t} - \frac{1}{2} \gamma$$

in the CUSUM test. An approximative solution to ARL function is given by, [1], [7]:

$$L(\mu_\alpha, \sigma_\alpha, h) = \exp(-2(\frac{\mu_\alpha}{\sigma_\alpha^2} + \frac{\mu_\alpha}{\sigma_\alpha^2} \beta)) - 1 + 2(\frac{\mu_\alpha}{\sigma_\alpha^2} + \frac{\mu_\alpha}{\sigma_\alpha^2} \beta)$$

$$\frac{2\beta}{\sigma_\alpha^2}$$

(19)

where $h = \frac{\log(\beta)}{2}$ is the detection threshold and $\beta = 1.166$. This approximation is based on $\alpha_t$ being white, which is satisfied in the fault free situation (and when a filter parametrization is applied). In the faulty situation this is only satisfied to a certain approximation.

Let $\alpha_t^j, j = 1, \ldots, 4$ denote the components of the CUSUM increments. In the fault free situation we have

$$\alpha_t^j \in N \left( -\frac{\gamma}{2}, 1 \right)$$

and the mean time between false alarms, $\hat{\tau}_{MTFA}$, can be assessed through:

$$\hat{\tau}_{MTFA} = \hat{L}(-\frac{\gamma}{2}, 1, h)$$

In a faulty situation

$$\alpha_t^j \in N \left( \frac{|S(\theta)|A^j}{2\sigma_1} - \frac{1}{2} \gamma \frac{\sigma_j^2}{\sigma_0^2} \right)$$

Here

$$l^T = [\cos(\phi(\theta)) \sin(\phi(\theta)) - \cos(\phi(\theta)) - \sin(\phi(\theta))]$$

where $l^j$ is the j'th component of $l$. The mean time for detection, $\hat{\tau}_{MTD}$, can be assessed from

$$\hat{\tau}_{MTD} = \min_j \hat{L}(\frac{|S(\theta)|A^j}{2\sigma_1} - \frac{1}{2} \gamma \frac{\sigma_j^2}{\sigma_0^2}, h)$$

(20)
An important thing with the AFD set-up used in this paper is that it is possible to change $\hat{t}_{MTD}$ and $\hat{t}_{MT FA}$ by the design of of the auxiliary input signal $\eta_k$. The mean values of $s_t$ and $e_t$ are directly proportional with the amplitude of $\eta_k$ when faults has occurred in the system.

VI. EXAMPLE

Let’s consider a sampled version of a simple second order system given by

$$G(s) = \frac{k}{s^2 + 2\zeta\omega_s + \omega^2} = \frac{1}{s^2 + 0.2s + 1}$$

influenced by stochastic disturbances. Variations in the three parameters $k$, $\zeta$ and $\psi$ will be considered as parametric faults in the system.

In discrete time ($T_s = 0.01$ sec) and in state space the system is given by

$$x_{t+1} = Ax_t + Bu_t + Bd_t, \quad y_t = Cx_t + e_t$$

where the noise processes are zero mean white noise sequences and

$$\text{Var} \left\{ \begin{bmatrix} d_t \\ e_t \end{bmatrix} \right\} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.01 \end{bmatrix}$$

The process noise is here an input disturbance, but the methods are by no means restricted to this type. The control is based on a state estimate obtained by means of a stationary Kalman filter and the control is an ordinary LQ controller which aim at minimizing the objective function

$$J = \mathbb{E} \left\{ \sum_{k=0}^{\infty} x_t^T Q x_t + u_t^T R u_t \right\} \quad Q = I_2 \quad R = 0.2$$

This design results in a controller given by:

$$V = \frac{z^2 - 1.931z + 0.9332}{z^2 - 1.957z + 0.9581} \quad U = \frac{-0.2664z + 0.2661}{z^2 - 1.957z + 0.9581}$$

and a model parameterized through

$$N = \frac{5.05z + 5.046}{z^2 - 1.957z + 0.9581} \times 10^{-5} \quad M = \frac{z^2 - 1.998z + 0.998}{z^2 - 1.957z + 0.9581}$$

A simple analysis of this closed system results in $\sigma_k^2 = 1.04e^{-3}$. As mentioned in the precious section the auxiliary signal was chosen to be a harmonic function, which has a distinct signature in the residual signal if a fault is present. The frequency was chosen by investigating the variation of $S(\omega, \theta)$ (see Fig. 3) in relation to $P_{\eta}$ and $P_{ed}$ over a range of frequencies and type of faults. It has been selected to use the same harmonic function for both detection and isolation. It is therefor also relevant to consider the variation of $S(\omega, \theta)$ in the complex plane for different frequencies and fault parameters. Based on this analysis, the frequency was chosen to be $\omega_0 = 2.5 rad/sec$. The amplitude was chosen to be 0.64 which is equivalent of having an power increase to a level ten times the stochastic variance. In Fig. 4, $S(\omega, \theta)$ is shown in the complex plane for different parametric faults for $\omega_0 = 2.5 rad/sec$.

As described in section IV-B the complex plane is divided into three double coned areas with top in the origin. Each type of parameter change are assigned a designated unit vector in the complex plane (see Table I).

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Designated vector directions</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\zeta$</th>
<th>$\psi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9830</td>
<td>0.1359</td>
<td>0.9934</td>
</tr>
<tr>
<td>-0.1834</td>
<td>0.9881</td>
<td>0.1144</td>
</tr>
</tbody>
</table>

The parameters in the CUSUM detector was chosen to be:

$$\gamma = 0.01 \quad B = 50$$

The choice of $\sigma_1$ was based on the knowledge of $\sigma_k^2$. This is related to the fact that for these choices $S_{MTFA} = 9181$.

Consider now an initial (at $t=0$) change in each of the three parameters: $k$, $\zeta$ and $\psi$. The detector signals are plotted in the Fig. 5 for a a change in $\zeta$ from 0.1 to 0.15. Additionally,
an equivalent deterministic simulation is plotted as well. The results of the three simulations for all three parameters are summarized in Table II. Each row in this table is related to one type of parameter change (in \( k \), \( \zeta \) and \( \psi \)). The first column gives the channel number which alarm for a fault. Second column contains the time instant of detection, \( T_d \), and the third column contains the assessment in (20) of \( \xi_{MTD} \).

![CUSUM signals for a fault in \( \zeta \).](image)

**TABLE II**

<table>
<thead>
<tr>
<th>Channel</th>
<th>( T_d ) (sec)</th>
<th>( \xi_{MTD} ) (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Fig. 5</td>
<td>58.60</td>
<td>49.34</td>
</tr>
<tr>
<td>—</td>
<td>115.60</td>
<td>115.01</td>
</tr>
</tbody>
</table>

When a fault has been detected and the fault instant, \( T_d \), has been estimated, data from \( T_d \) to \( T_d + t_d \) is used according to (17) to estimate the fault signature matrix, \( \hat{S}(\omega, \Theta) \). The estimate of the time difference between fault occurrence, \( \hat{t}_d \), and detection, \( t_d \) are listed as the second column in Table III. This is illustrated in Fig. 6, where the vector \( v \) easily is recognized as the line from origin. The fault isolation is carried out as given by (18) which is a mechanization of finding the nearest fault curve. The results are summarized in Table III where each row corresponds to one fault (and one simulation). The column wise data (the last 3 column) is the projection of \( v \) on each \( v_i \), \( i \in \{1, 2, 3\} \) in percent (with sign). As could be predicted from Fig. 4 it is in the table clear that it is harder to isolate changes in \( k \) and \( \psi \) than from changes in \( \zeta \).

**TABLE III**

<table>
<thead>
<tr>
<th></th>
<th>( t_d - t_f ) (sec)</th>
<th>( k )</th>
<th>( \zeta )</th>
<th>( \psi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>—, fault in ( k )</td>
<td>58.47</td>
<td>48.77</td>
<td>3.24</td>
<td>47.99</td>
</tr>
<tr>
<td>—, fault in ( \zeta )</td>
<td>115.55</td>
<td>8.69</td>
<td>76.99</td>
<td>14.31</td>
</tr>
<tr>
<td>—, fault in ( \psi )</td>
<td>115.59</td>
<td>44.44</td>
<td>9.81</td>
<td>45.75</td>
</tr>
</tbody>
</table>

**REFERENCES**


