

# ELIMINATING THE INITIAL STATE FOR THE GENERALIZED LIKELIHOOD RATIO TEST

David Törnqvist \* Fredrik Gustafsson \*

\* *Department of Electrical Engineering  
Linköpings universitet, SE-58183 Linköping, Sweden  
E-mail: {tornqvist, fredrik}@isy.liu.se*

Abstract: Fault detection based on comparing a batch of data with a model of the system using the *generalized likelihood ratio* test is considered. Careful treatment of the initial state of the model is quite important, in particular for short batch sizes. There are two standard approaches to this problem. One is based on a parity space, where the influence of initial state is removed by projection, and the other on using prior information obtained by Kalman filtering past data. A new idea of *anti-causal Kalman filtering* in the present data batch is introduced and compared to the previous methods. An efficient parameterization of *incipient faults* is given. It is shown in simulations of torque disturbances on a DC-motor that efficient fault profile parameterization and using smoothed estimates of the initial state increase performance considerably. *Copyright ©2006 IFAC*

Keywords: Fault detection; Linear systems; State-space models; Signal processing; State estimation

## 1. INTRODUCTION

A standard approach to model-based change detection is to first compute a residual from a batch of data based on the model, and then check if this residual is significantly different from zero, see (Basseville and Nikiforov 1993), (Gustafsson 2001, Chapters 6, 11). For state space models as considered here, the initial state affects the residual significantly, and its influence has to be removed. The unknown initial state can be approached in the following ways:

- Using a projection that removes the influence of initial conditions as done in the classical parity space approach (Basseville and Nikiforov 1993, Chow and Willsky 1984, Ding *et al.* 1999, Gertler 1997, Gertler 1998).
- Using prior knowledge of the initial state obtained with a causal Kalman filter on past data preceding the data batch under consideration (Willsky and Jones 1976). This gives a correct state estimate if the fault was not already present at that time.
- Using a minimum variance estimator (that is, anti-causal Kalman filter) on the data in the

batch, which gives the correct state distribution if there is no fault in the batch.

- Using a smoothed estimate by combining the two estimates above.

The last two approaches are presumably new. However, it will be shown that the third approach is identical to the first one. The anti-causal estimate is rather to be seen as a tool to derive the smoothed estimate, which will be shown to give the best performance in simulations.

The faults are parameterized using a low-rank orthogonal basis to distinguish them from noise. Depending on the basis function chosen, incipient faults as well as stepwise faults can be modeled. This will show to increase the performance significantly.

The paper is organized as follows. Section 2 introduces the notation and models. Different ways to estimate the initial state are given in Section 3, and used in the GLR-test in Section 4. Section 5 compares the approaches by simulating torque faults in a DC-motor and conclusions are given in Section 6.

## 2. MODELS AND NOTATION

### 2.1 State Space Model to Signal Model

The linear system is defined as the state space model

$$\begin{aligned} x_{t+1} &= F_t x_t + G_{u,t} u_t + G_{f,t} f_t + G_{v,t} v_t \\ y_t &= H_t x_t + H_{u,t} u_t + H_{f,t} f_t + e_t \end{aligned} \quad (1)$$

We separate the following types of input:

- Deterministic known input  $u_t$ . This is common in control applications.
- Deterministic unknown fault input  $f_t$ , which is used in the fault detection literature. The known matrices  $G_{f,t}$  and  $H_{f,t}$  determine which part of the system will be affected by the different faults.
- Stochastic unknown disturbances  $v_t$  and  $e_t$ , process noise and measurement noise, respectively, which are used in the Kalman filter setting. Both will here be assumed to be independent and Gaussian, with zero mean and covariance matrices  $Q_t$  and  $R_t$ , respectively.

Over a sliding window, where the signals are stacked, the model (1) can be written as

$$\mathbb{Y} = \mathcal{O}x_{t-L+1} + \bar{H}_u \mathbb{U} + \bar{H}_v \mathbb{V} + \bar{H}_f \mathbb{F} + \mathbb{E}. \quad (2)$$

Stack  $L$  signal values to define the signal vectors  $\mathbb{Y} = (y_{t-L+1}^T, \dots, y_t^T)^T$ , etc for all signals. Also define the Toeplitz matrices (time indices are omitted for simplicity)

$$\bar{H}_s = \begin{pmatrix} H_s & 0 & \dots & 0 \\ HG_s & H_s & \dots & 0 \\ \vdots & & \ddots & \vdots \\ HF^{L-2}G_s & \dots & HG_s & H_s \end{pmatrix} \quad (3)$$

for all signals  $s = u, f, v$  and the observability matrix

$$\mathcal{O} = \begin{pmatrix} H \\ HF \\ \vdots \\ HF^{L-1} \end{pmatrix}. \quad (4)$$

Since the input  $U$  is known it can always be subtracted from the output and without loss in generality  $\mathbb{Z}$  will be used as output of the system.

$$\mathbb{Z} \triangleq \mathbb{Y} - \bar{H}_u \mathbb{U} = \mathcal{O}x_{t-L+1} + \bar{H}_f \mathbb{F} + \bar{H}_v \mathbb{V} + \mathbb{E} \quad (5)$$

Note that we use the notation  $x$  for  $x_{t-L+1}$  to shorten the notation from here on. We also define  $S = \text{Cov}(\bar{H}_v \mathbb{V} + \mathbb{E})$ .

### 2.2 Structured Faults

To lower the dimension of the space spanned by the fault and thereby get less noisy residuals, the fault pattern is modeled. A regression model is used, so the fault can be parameterized as

$$\mathbb{F} = \Phi^T \theta, \quad (6)$$

where the columns of  $\Phi^T = (\Phi_1^T, \Phi_2^T, \dots)$  are orthogonal vectors generated by a Chebyshev polynomial, see (Abramowitz and Stegun 1965, Rivlin 1974),

and  $\theta$  contains the parameters. The regressors obtained has the form of a constant, linear, quadratic term and so on. This means that a step fault can be described using a one dimensional basis and incipient faults using a low dimensional basis. An example of an incipient fault can be seen in Figure 1 where also the basis is shown. When structured faults are used, simply define  $\bar{H}_\theta \triangleq \bar{H}_f \Phi^T$ .

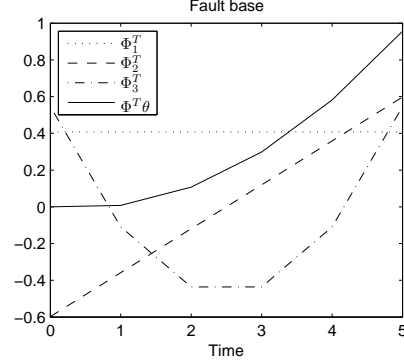


Fig. 1. Basis functions used to describe the behavior of a fault. An example of a parameterized fault is also plotted with  $\theta = (0.5, 0.8, 0.3)$ .

## 3. STATE ESTIMATION

There are several ways of estimating the initial state of the time window. Given the data in the window, the simplest way is to make an orthogonal projection of the measurements onto the signal space. Another approach would be to consider the noise distribution and to make an oblique projection in order to obtain a minimum variance estimate.

### 3.1 Orthogonal Projection

A system (5) without faults is described by

$$\mathbb{Z} = \mathcal{O}x + \bar{H}_v \mathbb{V} + \mathbb{E}. \quad (7)$$

One way of estimating the initial state, without using the knowledge about noise covariances, is to use

$$\hat{x} = \mathcal{O}^\dagger \mathbb{Z}. \quad (8)$$

Here,  $\star^\dagger$  denotes the Moore-Penrose pseudo inverse (see (Golub and van Loan 1996)). This estimate is unbiased since

$$\mathbb{E} \hat{x} = \mathbb{E} \mathcal{O}^\dagger \mathbb{Z} = \mathbb{E}(x + \mathcal{O}^\dagger(\bar{H}_v \mathbb{V} + \mathbb{E})) = x.$$

In presence of faults, the expected value would be

$$\mathbb{E} \hat{x} = \mathbb{E} \mathcal{O}^\dagger \mathbb{Z} = x + \mathcal{O}^\dagger \bar{H}_\theta \theta. \quad (9)$$

Using the estimator in (8), the influence of the initial state on the output is given by

$$\mathcal{O} \hat{x} = \underbrace{\mathcal{O} \mathcal{O}^\dagger}_{\triangleq \mathcal{P}_\mathcal{O}} \mathbb{Z}, \quad (10)$$

where  $\mathcal{P}_\mathcal{O}$  denotes the orthogonal projection onto the column space of  $\mathcal{O}$ , see (Golub and van Loan 1996).

### 3.2 Minimum Variance Estimation in Sliding Window

The estimator derived in Section 3.1 gives an unbiased estimate of the initial state  $x$ . This section derives an unbiased estimator with minimum variance of the estimation error. Such an estimator is often referred to as the *Best Linear Unbiased Estimator* (BLUE).

For the equation system  $\mathcal{O}x = \mathbb{Z}$ , the least squares (LS) solution  $\hat{x}^{LS} = \mathcal{O}^\dagger \mathbb{Z}$  is the minimum variance estimate if and only if  $\text{Cov}(\mathbb{Z}) = I$ . However, if  $\text{Cov}(\mathbb{Z}) = S$ , then  $\text{Cov}(S^{-1/2}\mathbb{Z}) = I$ , so pre-multiplying with  $S^{-1/2}$  is a whitening operation. The minimum variance estimate can now be written as

$$\hat{x} = (S^{-1/2}\mathcal{O})^\dagger S^{-1/2}\mathbb{Z}. \quad (11)$$

This estimate will be distributed according to

$$\hat{x} \sim \mathcal{N}\left(x + (S^{-1/2}\mathcal{O})^\dagger S^{-1/2}\bar{H}_\theta\theta, (\mathcal{O}^T S^{-1}\mathcal{O})^{-1}\right). \quad (12)$$

The basis for  $x$  in the measurement space is  $\mathcal{O}$ , so estimating the term  $\mathcal{O}x$  can be done as

$$\mathcal{O}\hat{x} = \underbrace{\mathcal{O}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2}}_{\triangleq \mathcal{P}} \mathbb{Z}, \quad (13)$$

where  $\mathcal{P}$  is an oblique projection onto  $\mathcal{R}(\mathcal{O})$  along  $\mathcal{R}(S^{1/2}(S^{-1/2}\mathcal{O})^\perp)$ , see (Meyer 2000, p.387).

### 3.3 Smoothed Estimate

The initial state of the window can be known from prior information or be estimated with a Kalman filter from old data. The information about the initial state is assumed to be a Gaussian distributed estimate

$$\hat{x}^{(1)} \sim \mathcal{N}(x, P^{(1)}). \quad (14)$$

Then, the smoothed initial state is estimated by forming the joint state estimate of (11) and (14) by the standard sensor fusion formula, see for instance (Gustafsson 2001). Denote the estimate from (11) with  $\hat{x}^{(2)}$  and its covariance  $P^{(2)}$ , then

$$\hat{x} = P\left(P^{(1)-1}\hat{x}^{(1)} + P^{(2)-1}\hat{x}^{(2)}\right), \quad (15)$$

where

$$P \triangleq \left(P^{(1)-1} + P^{(2)-1}\right)^{-1}. \quad (16)$$

## 4. GLR TESTS

Fault detection is here considered as detecting whether the fault is zero or not. This approach corresponds to the hypothesis test

$$\mathcal{H}_0 : \theta = 0 \quad (17a)$$

$$\mathcal{H}_1 : \theta \neq 0. \quad (17b)$$

This section will derive different test statistics based on the Generalized Likelihood Ratio, GLR, for the hypothesis test above.

### 4.1 Parity Space Approach

The parity space equations can be derived if the initial state is estimated by an orthogonal projection as in 3.1. The prediction error is here computed as

$$\varepsilon = \mathbb{Z} - \mathcal{O}\hat{x} = (I - \mathcal{P}_\mathcal{O})\mathbb{Z} = \mathcal{P}_{\mathcal{O}^\perp}\mathbb{Z}. \quad (18)$$

The projector can be computed as

$$\mathcal{P}_{\mathcal{O}^\perp} = \mathcal{B}_{\mathcal{O}^\perp}\mathcal{B}_{\mathcal{O}^\perp}^T, \quad (19)$$

where  $\mathcal{B}_{\mathcal{O}^\perp}$  is an orthonormal basis for  $\mathcal{R}(\mathcal{O}^\perp)$ .

*Residual* Since the prediction error is computed by a non full rank projection, the information in the prediction error can be represented by a vector of lower dimension. This vector is here denoted residual and since  $\mathcal{B}_{\mathcal{O}^\perp}^T \mathcal{B}_{\mathcal{O}^\perp} = I$ , where the dimension is the rank of the projection in (19), the residual becomes

$$r = \mathcal{B}_{\mathcal{O}^\perp}^T \varepsilon = \mathcal{B}_{\mathcal{O}^\perp}^T \mathbb{Z}, \quad (20)$$

which has the covariance

$$\text{Cov}(r) = \mathcal{B}_{\mathcal{O}^\perp}^T S \mathcal{B}_{\mathcal{O}^\perp}. \quad (21)$$

The normalized residual is then

$$\bar{r} = \underbrace{(\mathcal{B}_{\mathcal{O}^\perp}^T S \mathcal{B}_{\mathcal{O}^\perp})^{-1/2} \mathcal{B}_{\mathcal{O}^\perp}^T}_{\triangleq \bar{W}_1^T} \mathbb{Z}. \quad (22)$$

*Test Statistic* In the hypothesis test, (17), the question is whether the fault is zero or not. This can be rearranged to a hypothesis test based on the residual as

$$\mathcal{H}_0 : \bar{r} \sim \mathcal{N}(0, I) \quad (23a)$$

$$\mathcal{H}_1 : \bar{r} \sim \mathcal{N}(\bar{W}_1^T \bar{H}_\theta \theta, I). \quad (23b)$$

To decide between the hypotheses, a GLR test will be performed with test statistic

$$L = 2 \log \left( \frac{\sup_\theta p(\bar{r} - \bar{W}_1^T \bar{H}_\theta \theta)}{p(\bar{r})} \right) = 2 \sup_\theta \log \frac{e^{-\frac{1}{2}\|\bar{r} - \bar{W}_1^T \bar{H}_\theta \theta\|_2^2}}{e^{-\frac{1}{2}\|\bar{r}\|_2^2}}. \quad (24)$$

This function is maximized for  $\theta = (\bar{W}_1^T \bar{H}_\theta)^\dagger \bar{r}$ , which gives

$$L = \bar{r}^T \mathcal{P}_{\bar{W}_1^T \bar{H}_\theta} \bar{r}. \quad (25)$$

### 4.2 Minimum Variance State Estimation in Sliding Window

The prediction error can be computed in minimum variance sense with the state estimate from (13) as

$$\varepsilon = \mathbb{Z} - \mathcal{O}\hat{x} = (I - \mathcal{P})\mathbb{Z}. \quad (26)$$

The matrix  $I - \mathcal{P}$  is the complementary projector to  $\mathcal{P}$  which can be shown to have the Singular Value Decomposition, SVD

$$I - \mathcal{P} = (\mathcal{B}_{S^{1/2}(S^{-1/2}\mathcal{O})^\perp} \mathcal{B}_{S^{-1}\mathcal{O}}) \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{B}_{\mathcal{O}^\perp}^T \\ \mathcal{B}_{\mathcal{O}^T} \end{pmatrix}, \quad (27)$$

where  $\mathcal{B}_*$  is an orthogonal basis in  $\mathcal{R}(\star)$ . To construct the residual, the dimension of  $\varepsilon$  should be reduced to the rank of  $I - \mathcal{P}$ . By inspection of (27), the natural pre-multiplier would be  $\Sigma^{-1}\mathcal{B}_{S^{1/2}(S^{-1/2}\mathcal{O})^\perp}^T$ . Since

$$\Sigma^{-1}\mathcal{B}_{S^{1/2}(S^{-1/2}\mathcal{O})^\perp}^T(I - \mathcal{P}) = \mathcal{B}_{\mathcal{O}^\perp}^T, \quad (28)$$

the residual will be

$$r = \mathcal{B}_{\mathcal{O}^\perp}^T \mathbb{Z}. \quad (29)$$

This coincides with the residual in the parity space approach and the hypothesis test therefore coincides with the one in Section 4.1. It is thereby proved that if the initial state is estimated from data in the window, the GLR test will be equal to the parity space method.

### 4.3 State Estimation by Smoothing

The prediction error is formed by using the smoothed estimate from Section 3.3 as

$$\varepsilon = \mathbb{Z} - \mathcal{O}\hat{x} = \mathcal{O}(x - \hat{x}) + \bar{H}_\theta\theta + \bar{H}_v\nabla + \mathbb{E}, \quad (30)$$

then the expected value of the prediction error is obtained as

$$\begin{aligned} \mathbb{E}\varepsilon &= \bar{H}_\theta\theta - \mathcal{P}\mathcal{P}^{(2)-1}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2}\bar{H}_\theta\theta \\ &= \underbrace{(I - \mathcal{O}\mathcal{P}\mathcal{O}^T S^{-1})}_{\triangleq \bar{W}_2^T} \bar{H}_\theta\theta. \end{aligned} \quad (31)$$

To make the test statistic, it is desirable to have  $\varepsilon$  as a function of the measurements,  $\mathbb{Z}$ , and the state estimate from the Kalman filter,  $\hat{x}^{(1)}$ . Thus,

$$\begin{aligned} \varepsilon &= \mathbb{Z} - \mathcal{O}\hat{x} = \mathbb{Z} - \mathcal{O}\mathcal{P}\left(P^{(1)-1}\hat{x}^{(1)} + P^{(2)-1}\hat{x}^{(2)}\right) \\ &= \left(I - \mathcal{O}\mathcal{P}\mathcal{P}^{(2)-1}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2}\right)\mathbb{Z} \\ &\quad - \mathcal{O}\mathcal{P}\mathcal{P}^{(1)-1}\hat{x}^{(1)} \\ &= W_2^T \mathbb{Z} - \mathcal{O}\mathcal{P}\mathcal{P}^{(1)-1}\hat{x}^{(1)} \end{aligned} \quad (32)$$

and the covariance is

$$\begin{aligned} \text{Cov}(\varepsilon) &= W_2^T S W_2 + \mathcal{O} \underbrace{\mathcal{P}\mathcal{P}^{(1)-1}\mathcal{P}}_{\triangleq Q} \mathcal{O}^T \\ &= \begin{pmatrix} W_2^T & \mathcal{O} \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} W_2 \\ \mathcal{O}^T \end{pmatrix}. \end{aligned} \quad (33)$$

It is important to note that the matrix  $W_2^T$  is time varying due to the dependence of the covariance of the initial state estimate, see (31), (14) and (15). However, the Kalman filter will become stationary and then the covariance converges to a matrix which can be computed by solving the Riccati equation.

*Lemma 1.* Assume that  $S$  and  $Q$  are positive definite matrices, then the matrix in (33) is positive definite.

**PROOF.** See Appendix A.

Since  $Q = \mathcal{P}\mathcal{P}^{(1)-1}\mathcal{P}$  where  $P^{(1)}$  and  $P$  are positive definite matrices,  $Q$  is also positive definite. Then, according to Lemma 1, the covariance of  $\varepsilon$  is positive

definite and thus nonsingular. The normalized residual can therefore be expressed as

$$\begin{aligned} \bar{r} &= \text{Cov}(\varepsilon)^{-1/2}\varepsilon = \underbrace{\text{Cov}(\varepsilon)^{-1/2}W_2^T}_{\triangleq \bar{W}_2^T} \mathbb{Z} \\ &\quad - \text{Cov}(\varepsilon)^{-1/2}\mathcal{O}\mathcal{P}\mathcal{P}^{(1)-1}\hat{x}^{(1)}. \end{aligned} \quad (34)$$

*Test Statistic* Using the normalized residual (34), the hypotheses test (17) can be written as

$$\begin{aligned} \mathcal{H}_0 &: \bar{r} \sim \mathcal{N}(0, I) \\ \mathcal{H}_1 &: \bar{r} \sim \mathcal{N}(\bar{W}_2^T \bar{H}_\theta\theta, I). \end{aligned}$$

This yields the log-likelihood ratio

$$\begin{aligned} L &= \sup_\theta \log \frac{e^{-\frac{1}{2}\|\bar{r} - \bar{W}_2^T \bar{H}_\theta\theta\|_2^2}}{e^{-\frac{1}{2}\|\bar{r}\|_2^2}} \\ &= \sup_\theta -\left(\|\bar{r} - \bar{W}_2^T \bar{H}_\theta\theta\|_2^2 - \|\bar{r}\|_2^2\right), \end{aligned} \quad (35)$$

which is maximized for  $\theta = (\bar{W}_2^T \bar{H}_\theta)^\dagger \bar{r}$ . Then

$$\begin{aligned} L &= -\left(\|I - \underbrace{\bar{W}_2^T \bar{H}_\theta (\bar{W}_2^T \bar{H}_\theta)^\dagger}_{\mathcal{P}_{\bar{W}_2^T \bar{H}_\theta}} \bar{r}\|_2^2 - \|\bar{r}\|_2^2\right) \\ &= \bar{r}^T \mathcal{P}_{\bar{W}_2^T \bar{H}_\theta} \bar{r} \end{aligned} \quad (36)$$

### 4.4 Statistics

To choose suitable thresholds for the test statistics above, it is necessary to compute their distributions. While having Gaussian noise, the test statistics will be Chi-Square distributed variables. For a thorough review of statistics in signal processing, see (Kay 1998). The normalized residual is distributed as

$$\bar{r} \sim \mathcal{N}(\bar{W}_\star^T \bar{H}_\theta\theta, I),$$

where  $\star$  is 1 or 2 (method) and  $\theta = 0$  under the null hypothesis (17a). The test statistic is then distributed as the non-central Chi-Square distribution

$$L = \bar{r}^T \mathcal{P}_{\bar{W}_\star^T \bar{H}_\theta} \bar{r} \sim \chi_\nu^2(\lambda) \quad (37)$$

where  $\nu = \text{rank}(W_\star^T \bar{H}_\theta)$  and

$$\begin{aligned} \lambda &= (W_\star^T \bar{H}_\theta\theta)^T \mathcal{P}_{\bar{W}_\star^T \bar{H}_\theta} W_\star^T \bar{H}_\theta\theta \\ &= (W_\star^T \bar{H}_\theta\theta)^T W_\star^T \bar{H}_\theta\theta. \end{aligned}$$

## 5. EXAMPLE

To show the performance of the fault detection algorithms, a DC-motor has been simulated. The DC-motor can have a fault which is interpreted as a torque disturbance or a fault in the voltage supply.

### 5.1 Modeling

The continuous transfer function from input voltage to axis angle is

$$G(s) = \frac{1}{s(s+1)}. \quad (38)$$

With state variables  $x_1$  (angle) and  $x_2$  (angular velocity), a sampled version of the DC-motor can be written as

$$x_{t+1} = Fx_t + G_u u_t + G_f f_t + G_w w_t \quad (39a)$$

$$y_t = Hx_t + e_t, \quad (39b)$$

with the system matrices

$$F = \begin{pmatrix} 1 & 1 - e^{-T} \\ 0 & e^{-T} \end{pmatrix}, \quad G_u = \begin{pmatrix} T - (1 - e^{-T}) \\ 1 - e^{-T} \end{pmatrix},$$

$$G_w = G_f = G_u, \quad H = (1 \ 0),$$

where  $\text{Cov}(w_t) = (\frac{\pi}{180})^2$ ,  $\text{Cov}(e_t) = (\frac{\pi}{180})^2$  and the sample interval used is  $T = 0.4$  s.

## 5.2 Simulations

In the simulation, 200 samples in time have been simulated with the unit step as input. After 100 fault free samples, a constant fault is introduced with the magnitude of  $\frac{2\pi}{180}$ , which is present until the last sample. The fault can be seen as a torque disturbance which causes a drop in the angular velocity,  $x_2$ , and a slope change in the angle,  $x_1$ . With this setup, 2000 Monte Carlo (MC) simulations have been made. One instance of the state trajectory realizations is shown in Figure 2 where also the fault free trajectory is shown using the same noise realization. The drop in angular velocity when the fault is present can be seen in Figure 2b. The angle, which is measured, has a small change in the slope which is hardly noticeable in Figure 2a. The fault

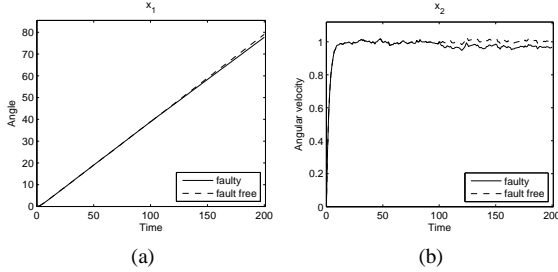


Fig. 2. Plots of the state trajectories with and without fault influence.

detection method using smoothing, includes a time-varying term dependent on the estimation covariance from the Kalman filter. After a transient period the Kalman filter will become stationary which means that the estimation covariance converges. Figure 3 shows  $\sqrt{\text{tr } P^{(1)}}$ , a measure of the state covariance,  $P^{(1)}$ . Note that  $P^{(1)}$  converges already after approximately 15 samples and the smoothing method can then be considered time invariant.

## 5.3 Fault detection

Fault detection on the simulated data has been made both with and without making use of the known fault structure. Since the fault is a step disturbance a basis of dimension one can be used to describe its behavior. For fault detection a sliding window of 8 samples is

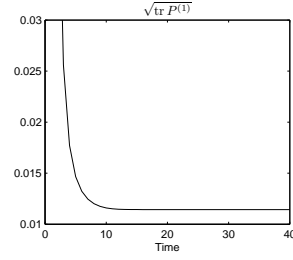


Fig. 3.  $\sqrt{\text{tr } P^{(1)}}$  is used as a measure of the state estimation covariance.

used and for the smoothed method the Kalman filter estimate from previous data is used as well. For the discussions below, some definitions have to be made. The alarm rate is calculated as

$$a_r(t) = \frac{\# L(t) > \gamma}{\# \text{MC-simulations}},$$

where  $L(t)$  is the test statistic and  $\gamma$  the threshold for detection. The probability of detection,  $P_D$ , is defined as the alarm rate during presence of fault. Probability of false alarm,  $P_{FA}$ , is defined as the alarm rate during non-faulty periods.

**A. No Knowledge about Fault Structure** First, the fault detection methods are used without assuming knowledge about the structure of the fault. The results can be seen in Figures 4–6. The average over the MC-simulations of the test statistics ((25) and (36)) is shown in Figure 4. Note that the difference in average during the fault-free part between the methods are due to different degrees of freedom of the test statistic. The parity space method has 6 degrees of freedom whereas the smoothed method has 7. The alarm rate versus time is shown in Figure 5. The false alarm rate has been set to 1% and it and the actual alarm rate during the fault free parts of the simulation corresponds to this value. The method with smoothing outperforms the parity space approach, but the comparison is unfair since the estimation is based on a different amount of data. In Figure 6, the Receiver Operation Characteristics (ROC) is plotted. The ROC-curve shows how the probability of detection for a fault varies versus the probability of false alarm. Guessing corresponds to a straight line between (0, 0) and (1, 1). In this case the values for  $P_D$  is estimated one window after the entrance of the fault. This is done since the Kalman filter estimate, used in the smoothed method, will be influenced by the fault afterward. More about ROC-curves can be found in (Kay 1998).

**B. Using the Structure of the Faults** Secondly, the structure of the fault is used to enhance the fault detectability. Since it is here assumed that the fault is a step, a one dimensional basis function is used to describe the fault. It is discussed in Section 4.4 that the average of the test statistics should equal the degrees of freedom for the test statistic. Due to the parameterization, the degree of freedom is 1 and the effect of this can be seen in Figure 4 where the average of the test statistics ((25) and (36)) are 1 in the fault-free part. The lower average during the fault-

free part makes the relative raise to the faulty part larger and the probability of detection is therefore higher than without structure. Figure 5 shows that the probability of detection is considerably increased for both methods due to the use of the fault structure. The ROC-curves in Figure 6 has therefore also shifted towards higher detection rates.

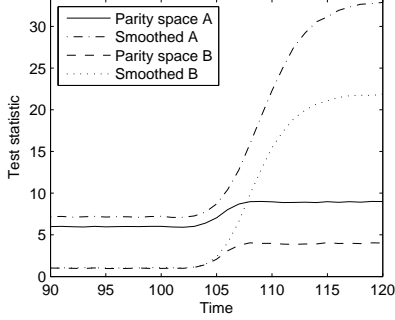


Fig. 4. The average of the test statistics (25) and (36). A and B means with and without using the structure of the fault.

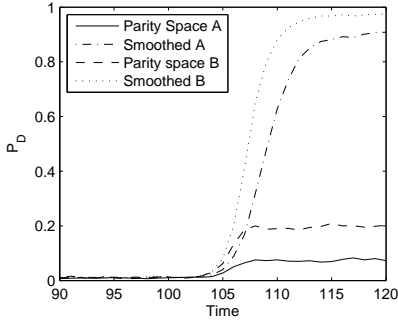


Fig. 5. The alarm rate. A and B means with and without using the structure of the fault.

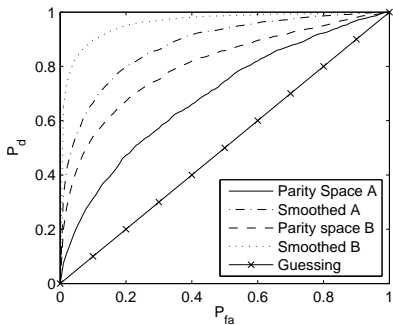


Fig. 6. A plot of the ROC-curve. A and B means with and without using the structure of the fault.

## 6. CONCLUSIONS

This paper studied the initial state dependency for the GLR test for fault detection based on a batch of data. Previous approaches are either based on eliminating the initial state by projection (the parity space approach) or by using a prior computed from passed data (the sliding window approach). We here derived the

minimum variance estimate of the initial state from the data batch. This, in itself, was shown in Section 4.2 to be algebraically identical to the parity space approach. However, in combination with the sliding window approach, it opens up the possibility to form a smoothed estimate. It was demonstrated in simulations that this improved performance considerably. Another contribution was to point out how a smooth parameterization of the fault profile can improve detection performance.

## Appendix A. PROOF OF LEMMA 1

If (33) is a positive definite matrix, the following must hold

$$x^T \begin{pmatrix} W_2^T & \mathcal{O} \\ 0 & Q \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} W_2 \\ \mathcal{O}^T \end{pmatrix} x > 0 \quad \forall x \neq 0. \quad (\text{A.1})$$

Since we know that  $S$  and  $Q$  is positive definite, it suffices to show that

$$x^T (W_2^T \mathcal{O}) \neq 0 \quad \forall x \neq 0. \quad (\text{A.2})$$

The matrix  $W_2^T$  can be written as  $I - \mathcal{O}M$  where  $M$  is an arbitrary matrix. If  $x \in \mathcal{N}(\mathcal{O}^T)$  and  $x \neq 0$  then

$$x^T (W_2^T \mathcal{O}) = (x^T \mathcal{O}) \neq 0. \quad (\text{A.3})$$

If  $x \notin \mathcal{N}(\mathcal{O}^T)$  and  $x \neq 0$  then

$$x^T (W_2^T \mathcal{O}) = (x^T - x^T \mathcal{O}M x^T \mathcal{O}) \neq 0 \quad (\text{A.4})$$

since at least the second element is nonzero. Thus, we have shown that there is no  $x \neq 0$  such that (A.2) is invalidated.

## REFERENCES

- Abramowitz, M. and Stegun, I. A. (Eds.) (1965). *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Publications.
- Basseville, M. and I. V. Nikiforov (1993). *Detection of Abrupt Changes: Theory and Application*. Information and system sciences series. Prentice Hall. Englewood Cliffs, NJ.
- Chow, A. Y. and A. S. Willsky (1984). Analytical redundancy and the design of robust failure detection systems. *IEEE Transactions on Automatic Control* **29**(7), 603–614.
- Ding, X., L. Guo and T. Jansch (1999). A characterization of parity space and its application to robust fault detection. *IEEE Transactions on Automatic Control* **44**(2), 337–343.
- Gertler, J. (1997). Fault detection and isolation using parity relations. *Control Engineering Practice* **5**(5), 653–661.
- Gertler, J. J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, Inc.
- Golub, G. H. and C. F. van Loan (1996). *Matrix Computations*. 3 ed. John Hopkins University Press.
- Gustafsson, F. (2001). *Adaptive filtering and change detection*. 2 ed. John Wiley & Sons, Inc.
- Kay, S. M. (1998). *Fundamentals of Statistical Signal Processing: Detection Theory*. Vol. 2. Prentice-Hall, Inc.
- Meyer, C. (2000). *Matrix Analysis and Applied Linear Algebra*. SIAM.
- Rivlin, T. J. (1974). *The Chebyshev Polynomials*. John Wiley & Sons, Inc.
- Willsky, A. and H. Jones (1976). A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control* **21**, 108–112.