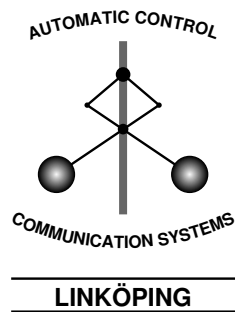


GLR tests for fault detection over sliding data windows

David Törnqvist, Fredrik Gustafsson, Inger Klein

Division of Automatic Control
Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden
WWW: <http://www.control.isy.liu.se>
E-mail: tornqvist@isy.liu.se, fredrik@isy.liu.se,
inger@isy.liu.se

20th October 2004



Report no.: [LiTH-ISY-R-2635](#)

Technical reports from the Control & Communication group in Linköping are available at <http://www.control.isy.liu.se/publications>.

Abstract

The Generalized Likelihood Ratio (GLR) test for fault detection as derived by Willsky and Jones is a recursive method to detect additive changes in linear systems in a Kalman filter framework. Here, we evaluate the GLR test on a sliding window and compare it to stochastic parity space approaches. Robust fault detection defined as being insensitive to faults in the signal space is also studied in the GLR framework.

Keywords: fault detection, statistical signal processing, robust estimation, parity space

GLR TESTS FOR FAULT DETECTION OVER SLIDING DATA WINDOWS

David Törnqvist* Fredrik Gustafsson*
Inger Klein*

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

Abstract: The Generalized Likelihood Ratio (GLR) test for fault detection as derived by Willsky and Jones is a recursive method to detect additive changes in linear systems in a Kalman filter framework. Here, we evaluate the GLR test on a sliding window and compare it to stochastic parity space approaches. Robust fault detection defined as being insensitive to faults in the signal space is also studied in the GLR framework.

Keywords: fault detection, statistical signal processing, robust estimation, parity space

1. INTRODUCTION

The work concerns primarily linear Gaussian models, which over a sliding window can be represented by the following signal model

$$Y = \mathcal{O}x + H_u U + H_v V + H_f F + E, \quad (1)$$

where Y is a vector of outputs during a sliding window time period, x is the initial state of the system, \mathcal{O} is the extended observability matrix, H_u , H_v and H_f are Toeplitz matrices describing how the inputs, U , process noise, V , and faults, F , enter the system and E is additive measurement noise. The sum of two noise terms is a Gaussian noise with $\text{Cov}(H_v V + E) = S$. The details are described in Section 2 and in (Gustafsson, 2001). Fault detection is here seen as the hypothesis test

$$H_0 : F = 0 \quad (2a)$$

$$H_1 : F \neq 0. \quad (2b)$$

Recently, (Desai and Mangoubi, 2003) proposed robust fault detection as only testing for faults that do not lie in the signal space. In our notation, the signal space is spanned by the observability

matrix, which means that we exclude faults F such that $H_f F$ belongs to the column space of \mathcal{O} and thus can be explained by a different state vector. Using the pseudo-inverse H_f^\dagger , this means that F does not belong to the range of $H_f^\dagger \mathcal{O}$ and the hypothesis test becomes

$$H_0 : F \in \mathcal{R}(H_f^\dagger \mathcal{O}) \quad (3a)$$

$$H_1 : F \notin \mathcal{R}(H_f^\dagger \mathcal{O}). \quad (3b)$$

These two tests will be treated in parallel.

The basic residual to use in fault detection is $r = Y - H_u U$, which depends on the state x . The classic parity space approach (Basseville and Nikiforov, 1993; Chow and Willsky, 1984; Ding *et al.*, 1999; Gertler, 1997; Gertler, 1998), computes a projection of this residual to the null space of the observability matrix, so $W_1^T r = W_1^T (H_f F + H_v V + E)$ is independent of the state. The original parity space approach is a deterministic one, so this residual is non-zero for non-zero faults. The extension to process and measurement noise was done in (Gustafsson, 2002), and analytical results on detectability of a all-one fault vector F was derived.

The first extension here is to arbitrary fault vectors, including incipient faults (slowly increasing) and general time-varying fault profiles. The GLR approach (Willsky and Jones, 1976) maximizes the likelihood ratio function over all faults, and the explicit GLR test statistic will be given. It turns out that this in both hypotheses tests (2) and (3) can be expressed as a certain projection of the residual.

A second approach is to estimate the state rather than projecting the residual to the parity space.

2. NOTATION

2.1 State space model to signal model

The linear system is defined as the state space model

$$\begin{aligned} x_{t+1} &= A_t x_t + B_{u,t} u_t + B_{f,t} f_t + B_{v,t} v_t \\ y_t &= C_t x_t + D_{u,t} u_t + D_{f,t} f_t + e_t \end{aligned} \quad (4)$$

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 $B_{f,t}$ and $D_{f,t}$ determines 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.

To establish the correspondence of models (4) and (1), stack L signal values to define the signal vectors $Y_t = (y_{t-L+1}^T, \dots, y_t^T)^T$, etc for all signals. We here use the time index t to note that fault detection is a recursive task. Also define the Hankel matrices (time indices are omitted for simplicity)

$$H_s = \begin{pmatrix} D_s & 0 & \dots & 0 \\ CB_s & D_s & \dots & 0 \\ \vdots & & \ddots & \vdots \\ CA^{L-2}B_s & \dots & CB_s & D_s \end{pmatrix} \quad (5)$$

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

$$\mathcal{O} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{pmatrix}. \quad (6)$$

Equation (4) can then be written as

$$\begin{aligned} Y_t - H_u U_t &= \\ \mathcal{O} x_{t-L+1} + H_f F_t + H_v V_t + E_t. \end{aligned} \quad (7)$$

which is (1). The definition of $S = \text{Cov}(H_v V_t + E_t)$ is straightforward.

2.2 Fault basis

The unknown fault vector F_t is a vector with arbitrary elements. Since H_f generally has full column rank, F_t can be used to explain the noise part. In particular, in the square H_f case, we can take $F_t = H_f^{-1}(H_v V_t + E_t)$. Since this does not make sense, we have to put certain restrictions on the fault vector. The basic idea is that the fault profile is a smooth function in time, so we can define smooth basis functions $m^i(t)$ and let \bar{F}_t^i be the coordinates, so

$$F_t = \sum_{i=1}^n m^i(t) \bar{F}_t^i = M \bar{F}_t$$

leading to

$$H_f F_t = H_f M \bar{F}_t = \bar{H}_f \bar{F}_t. \quad (8)$$

This is just a re-parameterization of the fault into a new known matrix \bar{H}_f with a low order basis \bar{F}_t . As a special case, we cover the model used in (Gustafsson, 2002) by using $n = 1$ and $m(t) = 1$. A polynomial basis might be suitable to model incipient faults. In the sequel, we will understand that the fault part in (1) is given by (8).

2.3 Projections and whitening operations

The basic tools in the derivation are the following:

- The pseudo-inverse operation is defined as $A^\dagger = (A^T A)^{-1} A^T$.
- Projection operator. A projection on the space spanned by the columns in A is given by $P_A = A(A^T A)^{-1} A^T = A A^\dagger$, with the obvious property $P_A A = A$. P_A is a basis for the range space of the columns in A .
- Projection on null space. To remove the state dependence in (1), the orthogonal projection $I - P_{\mathcal{O}}$ is used, with the obvious property $(I - P_{\mathcal{O}})\mathcal{O} = 0$. $I - P_{\mathcal{O}}$ is a basis for the null space of the columns in \mathcal{O} .
- Whitening. Assume that $\text{Cov}(r) = P$, then $\text{Cov}(P^{-1/2} r) = I$, so pre-multiplying with $P^{-1/2}$ is a whitening operation.
- Minimum variance (MV) estimation. For the equation system $Ax = r$, the least squares (LS) solution $\hat{x}^{LS} = A^\dagger r$ is the minimum variance estimate if and only if $\text{Cov}(r) = I$. That is, using pre-whitened residual, we have

$$\begin{aligned} \hat{x}^{MV} &= (P^{-1/2} A)^\dagger P^{-1/2} r \\ &= (A^T P^{-1} A)^{-1} A^T P^{-1} r. \end{aligned}$$

- GLR test. To test whether $F = 0$ or not when $r \in \mathbf{N}(HF, S)$, where r is a L dimensional vector, the log likelihood ratio of the two hypotheses is first formed:

$$\begin{aligned}\lambda(F) &= \log \frac{(2\pi)^{-L/2}(\det(S))^{-1/2}}{(2\pi)^{-L/2}(\det(S))^{-1/2}} \times \\ &\quad \times \frac{e^{-\frac{1}{2}(r-HF)^T S^{-1}(r-HF)}}{e^{-\frac{1}{2}r^T S^{-1}r}} = \\ &= \frac{1}{2} (r^T S^{-1}r - (r-HF)^T S^{-1}(r-HF)).\end{aligned}$$

The likelihood ratio is then maximized over the unknown parameter F to get the GLR test statistic

$$\begin{aligned}\lambda &= \max_F r^T S^{-1}r - (r-HF)^T S^{-1}(r-HF) \\ &= r^T S^{-1}r - (r-HH^\dagger r)^T S^{-1}(r-HH^\dagger r) \\ &= r^T S^{-1}r - r^T (I-P_H)S^{-1}(I-P_H)r \\ &= r^T (S^{-1} - (I-P_H)S^{-1}(I-P_H))r.\end{aligned}$$

The test statistic is χ^2 distributed, and a suitable threshold is based on this.

- Likelihood with unknown noise scaling. Suppose above that $r \in \mathbf{N}(0, \alpha S)$, where α is an unknown scaling of the noise covariance. This scaling is for instance present in a otherwise properly tuned Kalman filter, since the state estimates are insensitive to a common scaling factor in process noise and measurement noise (and initial state covariance).

$$\begin{aligned}\log \left((2\pi)^{-L/2} \det(S)^{-1/2} e^{-\frac{1}{2\alpha} r^T S^{-1} r} \right) &= \\ &= C - \frac{L}{2} \log(\alpha) - \frac{1}{2\alpha} r^T S^{-1} r \quad (9)\end{aligned}$$

which is maximized for $\hat{\alpha} = r^T S^{-1}r/L$.

3. GLR TEST STATISTICS

This section derives the main results, which are summarized in Table 1.

Table 1. The test statistic $\lambda = r^T P_M r$, this table shows the projection matrix for each case discussed in this section. The matrices \bar{W}_2^T and \bar{W}_3^T are given by equations (19) and (30) respectively.

Case	Conv. P_M	Robust P_M
1	$M = S^{-1/2}H_f$	$M = S^{-1/2}(I-P_{\mathcal{O}})H_f$
2	$M = \bar{W}_2^T H_f$	$M = \bar{W}_2^T H_f$
3	$M = \bar{W}_3^T H_f$	$M = \bar{W}_3^T (I-P_{\mathcal{O}})H_f$

3.1 Case 1: state estimation in sliding window

In this case, x is estimated by minimum variance estimation from data in the time window as

$$\begin{aligned}\hat{x}^{(2)} &= (S^{-1/2}\mathcal{O})^\dagger S^{-1/2}(Y - H_u U) \\ &\in \mathbf{N}(x + (S^{-1/2}\mathcal{O})^\dagger S^{-1/2}H_f F, (\mathcal{O}^T S^{-1}\mathcal{O})^{-1}),\end{aligned} \quad (10)$$

then the prediction error (or residual) becomes

$$\begin{aligned}r_t^{(1)} &= Y - H_u U - \mathcal{O}\hat{x}^{(2)} = \\ &= (I - \mathcal{O}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2})(Y - H_u U) \\ &\in \mathbf{N}((I - \mathcal{O}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2})H_f F, \text{Cov}(r_t^{(1)})).\end{aligned} \quad (11)$$

The covariance becomes

$$\begin{aligned}\text{Cov}(r_t^{(1)}) &= (I - \mathcal{O}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2})S \\ &= (I - \mathcal{O}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2})^T.\end{aligned} \quad (12)$$

To determine the likelihood ratio with a hypothesis test, first normalize the residual.

$$\begin{aligned}\bar{r}_t^{(1)} &= (\text{Cov}(r_t^{(1)}))^{-1/2} r_t^{(1)} \\ &= (\text{Cov}(r_t^{(1)}))^{-1/2} E(Y - H_u U - \mathcal{O}\hat{x}^{(2)}) \\ &\in \mathbf{N}(S^{-1/2}H_f F, I)\end{aligned} \quad (13)$$

3.1.1. Conventional GLR Test The hypotheses in (2) are here

$$\begin{aligned}H_0 &: \bar{r}_t^{(1)} \in \mathbf{N}(0, I) \\ H_1 &: \bar{r}_t^{(1)} \in \mathbf{N}(S^{-1/2}H_f F, I).\end{aligned}$$

This gives the log-likelihood ratio

$$\lambda_{1,c} = \max_F \log \frac{e^{-\frac{1}{2}\|\bar{r}_t^{(1)} - S^{-1/2}H_f F\|_2^2}}{e^{-\frac{1}{2}\|\bar{r}_t^{(1)}\|_2^2}}, \quad (14)$$

which is maximized when $F = (S^{-1/2}H_f)^\dagger \bar{r}_t^{(1)}$. Then,

$$\begin{aligned}\lambda_{1,c} &= \\ &= -\frac{1}{2} \left(\|\bar{r}_t^{(1)} - \underbrace{S^{-1/2}H_f(S^{-1/2}H_f)^\dagger}_{P_{S^{-1/2}H_f}} \bar{r}_t^{(1)}\|_2^2 - \|\bar{r}_t^{(1)}\|_2^2 \right) \\ &= -\frac{1}{2} \left(\bar{r}_t^{(1)T} (I - P_{S^{-1/2}H_f})^T (I - P_{S^{-1/2}H_f}) \bar{r}_t^{(1)} - \bar{r}_t^{(1)T} \bar{r}_t^{(1)} \right) \\ &= \frac{1}{2} \left(\bar{r}_t^{(1)T} P_{S^{-1/2}H_f} \bar{r}_t^{(1)} \right) = \\ &= \frac{1}{2} \|\bar{r}_t^{(1)}\|_{P_{S^{-1/2}H_f}}^2, \quad (15)\end{aligned}$$

where P_A denotes the orthogonal projection onto the row space of A .

3.1.2. Robust GLR Test The hypotheses in the robust test questions if there is no fault or if the fault resides in the subspace orthogonal to the signal space. Thus, the hypotheses in (3) are

$$\begin{aligned}H_0 &: \bar{r}_t^{(1)} \in \mathbf{N}(0, I) \\ H_1 &: \bar{r}_t^{(1)} \in \mathbf{N}(S^{-1/2}(I-P_{\mathcal{O}})H_f F, I)\end{aligned}$$

The log-likelihood ratio becomes

$$\lambda_{1,r} = \max_F \log \frac{e^{-\frac{1}{2}\|\bar{r}_t^{(1)} - S^{-1/2}(I-P_{\mathcal{O}})H_f F\|_2^2}}{e^{-\frac{1}{2}\|\bar{r}_t^{(1)}\|_2^2}}. \quad (16)$$

This ratio is maximized for $F = (S^{-1/2}(I - P_{\mathcal{O}})H_f)^\dagger \bar{r}_t^{(1)}$, this gives

$$\begin{aligned} \lambda_{1,r} &= \frac{1}{2} \bar{r}_t^{(1)T} P_{S^{-1/2}(I-P_{\mathcal{O}})H_f} \bar{r}_t^{(1)} = \\ &= \frac{1}{2} \|\bar{r}_t^{(1)}\|_{P_{S^{-1/2}(I-P_{\mathcal{O}})H_f}}^2. \end{aligned} \quad (17)$$

3.1.3. Comments on the Tests In both the detector for the conventional and robust test we have a quadratic expression in terms of the normalized residual. The only difference between the two cases is in the projection matrix, compare (15) and (17).

3.2 Case 2: parity space approach

In this case the residual, $Y - H_u U$, is multiplied with the orthogonal complement of \mathcal{O} , $I - P_{\mathcal{O}}$, to eliminate the dependence on x . Then the prediction error becomes

$$\begin{aligned} r_t^{(2)} &= \underbrace{(I - P_{\mathcal{O}})}_{W_2^T} (Y - H_u U) \\ &\in \mathbf{N}(W_2^T H_f F, W_2^T S W_2). \end{aligned} \quad (18)$$

In order to get unit variance, the residual can be normalized as

$$\begin{aligned} \bar{r}_t^{(2)} &= (W_2^T S W_2)^{-1/2} r_t^{(2)} \\ &= \underbrace{(W_2^T S W_2)^{-1/2} W_2^T}_{\bar{W}_2^T} (Y - H_u U) \\ &\in \mathbf{N}(\bar{W}_2^T H_f F, I). \end{aligned} \quad (19)$$

3.2.1. Conventional GLR Test The log-likelihood ratio for fault/no fault is

$$\begin{aligned} \lambda_{2,c} &= \max_F \log \frac{e^{-\frac{1}{2} \|\bar{r}_t^{(2)} - \bar{W}_2^T H_f F\|_2^2}}{e^{-\frac{1}{2} \|\bar{r}_t^{(2)}\|_2^2}} = \\ &= \max_F -\frac{1}{2} \left(\|\bar{r}_t^{(2)} - \bar{W}_2^T H_f F\|_2^2 - \|\bar{r}_t^{(2)}\|_2^2 \right) \end{aligned} \quad (20)$$

This ratio is maximized when $F = (\bar{W}_2^T H_f)^\dagger \bar{r}_t^{(2)}$, then the expression becomes

$$\begin{aligned} \lambda_{2,c} &= -\frac{1}{2} \left(\|\bar{r}_t^{(2)} - \underbrace{\bar{W}_2^T H_f (\bar{W}_2^T H_f)^\dagger}_{P_{\bar{W}_2^T H_f}} \bar{r}_t^{(2)}\|_2^2 - \|\bar{r}_t^{(2)}\|_2^2 \right) = \\ &= \frac{1}{2} \left(\bar{r}_t^{(2)T} P_{\bar{W}_2^T H_f} \bar{r}_t^{(2)} \right) = \frac{1}{2} \|\bar{r}_t^{(2)}\|_{P_{\bar{W}_2^T H_f}}^2, \end{aligned} \quad (21)$$

where P_A denotes the orthogonal projection onto the row space of A .

3.2.2. Robust GLR Test Decompose the term from the fault in two parts, the projection onto the signal space and the space orthogonal to it.

$$H_f F = P_{\mathcal{O}} H_f F + (I - P_{\mathcal{O}}) H_f F \quad (22)$$

Then the hypotheses for the robust test are if the fault is in the signal space or the space orthogonal to it.

$$\begin{aligned} H_0 : & Y = \mathcal{O}x + H_u U + P_{\mathcal{O}} H_f F + E \\ H_1 : & Y = \mathcal{O}x + H_u U + (I - P_{\mathcal{O}}) H_f F + E \end{aligned}$$

The residuals for the two cases becomes (with $W_2 = \mathcal{N}_{\mathcal{O}}$)

$$\begin{aligned} H_0 : & r_t^{(2)} = W_2^T (\mathcal{O}x + P_{\mathcal{O}} H_f F + E) = W_2^T E \\ H_1 : & r_t^{(2)} = W_2^T (\mathcal{O}x + (I - P_{\mathcal{O}}) H_f F + E) \\ &= W_2^T (H_f F + E) \end{aligned}$$

This is the same residuals that we are testing with the conventional test. So when estimating x by projection, the robust and conventional tests coincide.

3.3 Case 3: state estimation in preceding and sliding window

In this case x is estimated by a minimum variance estimator $\hat{x}^{(2)}$ with covariance $P^{(2)}$ from data in the sliding window as in section 3.1, but also with a Kalman filter from old data, providing $\hat{x}^{(1)}$ with covariance matrix $P^{(1)}$. This appears to be a logical approach to detect faults in the signal space, since the difference in state estimates should be due to estimation errors and faults in the signal space only. For instance, we have

$$\begin{aligned} r_t^{(3)} &= \mathcal{O}(\hat{x}^{(2)} - \hat{x}^{(1)}) \\ &\in \mathbf{N}(P_{\mathcal{O}} H_f F_t, \mathcal{O}(P^{(1)} + P^{(2)})\mathcal{O}^T). \end{aligned} \quad (23)$$

This is a residual for testing faults in the signal space. For faults orthogonal to the signal space, we proceed by forming the joint state estimate over all data by the standard sensor fusion formula as outlined below. The estimate from the Kalman filter is Gaussian distributed

$$\hat{x}^{(1)} \in \mathbf{N}(x, P^{(1)}). \quad (24)$$

The minimum variance estimate of x from data in the window is given by

$$\begin{aligned} \hat{x}^{(2)} &= (S^{-1/2} \mathcal{O})^\dagger S^{-1/2} (Y - H_u U) \\ &\in \mathbf{N}(x + (S^{-1/2} \mathcal{O})^\dagger S^{-1/2} H_f F, \underbrace{(\mathcal{O}^T S^{-1} \mathcal{O})^{-1}}_{P^{(2)}}). \end{aligned} \quad (25)$$

Then,

$$\hat{x} = P^{-1} (P^{(1)-1} \hat{x}^{(1)} + P^{(2)-1} \hat{x}^{(2)}), \quad (26)$$

where

$$P^{-1} = (P^{(1)-1} + P^{(2)-1})^{-1}. \quad (27)$$

The residual is then formed as

$$\begin{aligned}
r_t^{(3)} &= Y - H_u U - \mathcal{O}\hat{x} \\
&= \underbrace{(I - \mathcal{O}P^{-1}P^{(2)})^{-1}(S^{-1/2}\mathcal{O})^\dagger S^{-1/2}}_{W_3^T} \times \\
&\quad \times (Y - H_u U) - \mathcal{O}P^{-1}P^{(1)-1}\hat{x}^{(1)} \\
&= W_3^T(\mathcal{O}x + H_f F + E) - \mathcal{O}P^{-1}P^{(1)-1}(x + \tilde{x}^{(1)}) \\
&= \mathcal{O}x - \underbrace{(\mathcal{O}P^{-1}P^{(2)})^{-1}x + \mathcal{O}P^{-1}P^{(1)-1}x}_{\mathcal{O}x} \\
&\quad + W_3^T(H_f F + H_v V + E) - P^{-1}P^{(1)-1}\tilde{x}^{(1)} \\
&= W_3^T(H_f F + E) - P^{-1}P^{(1)-1}\tilde{x}^{(1)} \\
&\quad \in \mathbf{N}(W_3^T H_f F, \text{Cov}(r_t^{(3)}))
\end{aligned}$$

where

$$\tilde{x}^{(1)} = \hat{x}^{(1)} - x \in \mathbf{N}(0, P^{(1)}) \quad (28)$$

$$\text{Cov}(r_t^{(3)}) = W_3^T S W_3 + \mathcal{O}P^{-1}P^{(1)-1}P^{-1}\mathcal{O}^T. \quad (29)$$

The normalized residual is then formed as

$$\bar{r}_t^{(3)} = (\text{Cov}(r_t^{(3)}))^{-1/2}(Y - H_u U - \mathcal{O}\hat{x}) \quad (30)$$

$$\in \mathbf{N}\left(\underbrace{(\text{Cov}(r_t^{(3)}))^{-1/2}W_3^T H_f F, I}_{\bar{W}_3^T}\right). \quad (31)$$

3.3.1. Conventional Likelihood Test The hypotheses for the conventional test are

$$H_0 : \bar{r}_t^{(3)} \in \mathbf{N}(0, I)$$

$$H_1 : \bar{r}_t^{(3)} \in \mathbf{N}(\bar{W}_3^T H_f F, I).$$

This yields the log-likelihood ratio

$$\begin{aligned}
\lambda_{3,c} &= \max_F \log \frac{e^{-\frac{1}{2}\|\bar{r}_t^{(3)} - \bar{W}_3^T H_f F\|_2^2}}{e^{-\frac{1}{2}\|\bar{r}_t^{(3)}\|_2^2}} = \\
&= \max_F -\frac{1}{2} \left(\|\bar{r}_t^{(3)} - \bar{W}_3^T H_f F\|_2^2 - \|\bar{r}_t^{(3)}\|_2^2 \right). \quad (32)
\end{aligned}$$

Which is maximized for $F = (\bar{W}_3^T H_f)^\dagger \bar{r}_t^{(3)}$, then

$$\begin{aligned}
\lambda_{3,c} &= \\
&= -\frac{1}{2} \left(\|(I - \underbrace{\bar{W}_3^T H_f (\bar{W}_3^T H_f)^\dagger}_{P_{\bar{W}_3^T H_f}})\bar{r}_t^{(3)}\|_2^2 - \|\bar{r}_t^{(3)}\|_2^2 \right) \\
&= \frac{1}{2} \bar{r}_t^{(3)T} P_{\bar{W}_3^T H_f} \bar{r}_t^{(3)} = \frac{1}{2} \|\bar{r}_t^{(3)}\|_{P_{\bar{W}_3^T H_f}}^2 \quad (33)
\end{aligned}$$

3.3.2. Robust Likelihood Test The hypotheses for the robust test are

$$H_0 : \bar{r}_t^{(3)} \in \mathbf{N}(0, I)$$

$$H_1 : \bar{r}_t^{(3)} \in \mathbf{N}(\bar{W}_3^T (I - P_{\mathcal{O}}) H_f F, I).$$

Similar calculations as for the conventional test yields the robust log-likelihood ratio

$$\begin{aligned}
\lambda_{3,r} &= \frac{1}{2} \bar{r}_t^{(3)T} P_{\bar{W}_3^T (I - P_{\mathcal{O}}) H_f} \bar{r}_t^{(3)} = \\
&= \frac{1}{2} \|\bar{r}_t^{(3)}\|_{P_{\bar{W}_3^T (I - P_{\mathcal{O}}) H_f}}^2. \quad (34)
\end{aligned}$$


3.3.3. Comments on the Tests These tests are closely related to case 1 where x is determined only from data in the window. If $P^{(2)} \rightarrow \infty I$, then the normalized residual in (30) is identical to the one in case 1 and therefore also the detector.

4. CONCLUSIONS

The original generalized likelihood ratio (GLR) test for fault detection where derived using a Kalman filter approach, and thus all passed data influence the test statistic. On the other hand, the parity space approach is defined over a sliding window. In order to compare these standard approaches, we have derived the GLR test statistic for fault detection in linear Gaussian systems based on data over a sliding window. This is done in a systematic way for different assumption on the fault range space (robust/non-robust) and residual generators (parity space/state estimation). It turns out that all these cases correspond to different projections of the basic model residual.

REFERENCES

- Basseville, M. and I. V. Nikiforov (1993). *Detection of Abrupt Changes, Theory and Application*. Information and system sciences series. Prentice-Hall, Enlewood 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.
- Desai, Makund N. and Rami S. Mangoubi (2003). Robust gaussian and non-gaussian matched subspace detection. *IEEE Transactions on Signal Processing* **51**(12), 3115–3127.
- Ding, X., L. Guo and T. Jeansch (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.
- Gustafsson, Fredrik (2001). *Adaptive filtering and change detection*. 2 ed.. John Wiley & Sons, Inc.
- Gustafsson, Fredrik (2002). Stochastic fault diagnosability in parity spaces. In: *Proceedings of the 15th IFAC World Congress*. Barcelona, Spain.
- Willsky, A.S. and H.L. 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.

	Avdelning, Institution Division, Department Division of Automatic Control Department of Electrical Engineering	Datum Date 2004-10-20
	Språk Language <input type="checkbox"/> Svenska/Swedish <input checked="" type="checkbox"/> Engelska/English <input type="checkbox"/> _____	Rapporttyp Report category <input type="checkbox"/> Licentiatavhandling <input type="checkbox"/> Examensarbete <input type="checkbox"/> C-uppsats <input type="checkbox"/> D-uppsats <input checked="" type="checkbox"/> Övrig rapport <input type="checkbox"/> _____
URL för elektronisk version http://www.control.isy.liu.se		LiTH-ISY-R-2635
Titel GLR tests for fault detection over sliding data windows Title		
Författare David Törnqvist, Fredrik Gustafsson, Inger Klein Author		
Sammanfattning Abstract <p>The Generalized Likelihood Ratio (GLR) test for fault detection as derived by Willsky and Jones is a recursive method to detect additive changes in linear systems in a Kalman filter framework. Here, we evaluate the GLR test on a sliding window and compare it to stochastic parity space approaches. Robust fault detection defined as being insensitive to faults in the signal space is also studied in the GLR framework.</p>		
Nyckelord Keywords fault detection, statistical signal processing, robust estimation, parity space		