

STATISTICAL SIGNAL PROCESSING APPROACHES TO FAULT DETECTION

F. Gustafsson

*Department of Electrical Engineering, Linköpings universitet,
SE-581 83 Linköping, Sweden
fredrik@isy.liu.se*

Abstract: The parity space approach to fault detection and isolation (FDI) has been developed during the last twenty years, and the focus here is to describe its application to stochastic systems. A mixed model with both stochastic inputs and deterministic disturbances and faults is formulated over a sliding window. Algorithms for detecting and isolating faults on-line and analyzing the probability for correct and incorrect decisions off-line are provided. A major part of the paper is devoted to discussing properties of this model-based approach and generalizations to cases of incomplete model knowledge, and non-linear non-Gaussian models. For this purpose, a simulation example is used throughout the paper for numerical illustrations, and real-life applications for motivations. The final section discusses the reverse problem: fault detection approaches to statistical signal processing. It is motivated by three applications that a simple CUSUM detector in feedback loop with an adaptive filter can mitigate the inherent trade-off between estimation accuracy and tracking speed in linear filters.

Keywords: fault detection, diagnosis, Kalman filtering, adaptive filters, linear systems, principal component analysis, subspace identification

1. INTRODUCTION

The parity space approach to fault detection Basseville and Nikiforov (1993); Chow and Willsky (1984); Ding et al. (1999); Gertler (1997, 1998) is an elegant and general tool for additive faults in linear systems and is based on intuitively simple algebraic projections and geometry. Simply speaking, a residual r_t is a data projection

$$r_t = P^T Z_t, \quad Z_t = \begin{pmatrix} Y_t \\ U_t \end{pmatrix}, \quad (1)$$

where the data vector Z_t contains the measured input (U_t) and output (Y_t) over a certain time window. The parity space approach provides a tool to compute P that yields a residual vector that is zero when there is no fault and non-zero otherwise. A vector interpretation of the residual will be used for fault isolation, with a direct link to vector quantization in communication theory. Fault isolation is done by comparing the non-

zero residual to pre-computed fault vectors, and the closest one in a statistical meaning is chosen in an on-line algorithm. The vector approach provides an analytic approach to compute the *diagnosis probability matrix*. This contains all probabilities of giving an alarm of fault i given fault j has occurred, where also the non-faulty case is included ($i, j = 0$). Faults that are hard or even impossible to diagnose are discovered at this stage, and a system or algorithm re-design can be performed.

This stochastic theory comes as a consequence of the model assumptions. Section 4 motivates and generalizes the approach, using a simulated DC motor example and several applications:

- An example of how a small measurement error can severely perturb the standard parity space residual.
- A motivating application for the Gaussian assumption for the stochastic inputs.

- One important alternative approach is the Generalized Likelihood Ratio (GLR) test, and it is pointed out that this is actually equivalent to the stochastic parity space approach for the model assumptions at hand.
- Implications of the length of the time window in theory and for the example.
- Extensions and generalizations of the assumption of an additive single fault of constant magnitude.
- A successful application of the theory to a severely non-linear systems, by linearizing around a nominal state trajectory.
- The relation to data-driven approaches as Principal Component Analysis (PCA).
- The use and fusion of residuals from several independent models.

The major part of the paper treats statistical signal processing approaches to fault detection as outlined above. However, classical statistical signal processing problems as adaptive filtering can benefit from using fault detection ideas. Section 5 gives three concrete application examples where fault detection approaches to adaptive filtering are applied with very good results.

The model and basic parity space relations are given in Section 2. Section 3 provides the stochastic parity space theory, and Section 4 the motivations and extensions. The reverse problem of using fault detection for stochastic signal processing is exemplified in Section 5, and Section 6 concludes the paper.

2. PARITY SPACE NOTATION AND MODELS

The theory of parity space detection can be found in standard references as Basseville and Nikiforov (1993); Chow and Willsky (1984); Ding et al. (1999); Gertler (1997, 1998). The outline and notation below follows Chapter 11 in Gustafsson (2001).

2.1 Model

The linear system is here defined as the state space model

$$x_{t+1} = A_t x_t + B_{u,t} u_t + B_{d,t} d_t + B_{f,t} f_t + B_{v,t} v_t \quad (2a)$$

$$y_t = C_t x_t + D_{u,t} u_t + D_{d,t} d_t + D_{f,t} f_t + e_t. \quad (2b)$$

The matrices A_t, B_t, C_t, D_t depend on the system, while the signals belong to the following categories:

- Deterministic known input u_t , as is common in control applications.
- Deterministic unknown disturbance d_t , as is also common in control applications.
- Deterministic unknown fault input f_t , which is used in the fault detection literature. We here assume that f_t is either zero (no fault) or proportional to the unit vector $f_t = m_t e^i$, where e^i is all

zero except for element i which is one. Exactly which part of the system fault i affects is determined by the corresponding columns in $B_{f,t}$ and $D_{f,t}$. That is, we have $B_{f,t} f_t = B_{f,t}^i m_t$, where $B_{f,t}^i$ is column i in $B_{f,t}$ and similarly for $D_{f,t}^i$. This fault model covers offsets in actuators and sensors for instance. The fault magnitude m_t is assumed constant magnitude $m_t = m$, but the time-varying fault case will also be discussed.

- Stochastic unknown state disturbance v_t and measurement noise e_t , as used in a Kalman filter setting. There is an ambiguity of the interpretations of v_t and d_t . We might treat v_t as a deterministic disturbance, but in many cases this leads to an infeasible problem where no parity space exists. Both v_t and e_t are here assumed to be independent with zero mean and covariance matrices Q_t and R_t , respectively.
- The initial state is treated as an unknown variable, so no prior information is needed.

The dimension of any signal s_t is denoted as $n_s = \dim(s_t)$. Traditionally, either a stochastic ($d_t = 0$) or a deterministic ($v_t = 0, e_t = 0$) framework is used in the literature, but here we aim to mix them and combine the theories.

2.2 Parity space

There are many different derivations of the parity space. We here use one based on the discrete time state space model in (2) using data from a sliding window. In this case, the measurements over a sliding window of size L can be expressed explicitly in matrix form (without recursions) as:

$$Y_t = \mathcal{O} x_{t-L+1} + H_u U_t + H_d D_t + H_v V_t + H_f F_t + E_t. \quad (3)$$

where $Y_t = (y_{t-L+1}^T \dots y_t^T)^T$ and

$$H_s = \begin{pmatrix} D_s & 0 & \dots & 0 \\ C B_s & D_s & \dots & 0 \\ \vdots & & \ddots & \vdots \\ C A^{L-2} B_s & \dots & C B_s & D_s \end{pmatrix}, \quad \mathcal{O} = \begin{pmatrix} C \\ C A \\ \vdots \\ C A^{L-1} \end{pmatrix}. \quad (4)$$

Here H is defined for all signals $s \in \{u, d, f, v\}$. The covariance matrix of the measurement vector is $S = \text{Cov}(H_v V_t + E_t)$.

Without loss of generality, the residual generating matrix in (1) can be defined as

$$r_t = W^T (Y_t - H_u U_t) = \overbrace{W^T (I, -H_u)}^{P^T} Z_t \quad (5a)$$

$$= W^T (\mathcal{O} x_{t-L+1} + H_d D_t + H_f F_t + H_v V_t + E_t) \quad (5b)$$

The first equation is what can be computed on-line, and the second expression is what will be used in the analysis.

The parity space is defined to be insensitive to the input (yielding the factorization in (5a)), the initial

state and deterministic disturbances, which implies that $r_t = 0$ for any initial state x_{t-L+1} and any disturbance sequence d_k , $k = t - L + 1, \dots, t$, provided that there is no stochastic term present ($e_k = 0$, $v_k = 0$). This is achieved if the projection matrix satisfies the following condition:

Parity space condition:

$$W^T (\mathcal{O} H_d) = 0 \Leftrightarrow W \in \mathcal{N}_{[\mathcal{O} H_d]}. \quad (6)$$

Here, \mathcal{N}_A denotes the null space for the matrix A . The maximal dimension of the residual vector is given by (see equation (11.12) in Gustafsson (2000)):

$$L(n_y - n_d) - n_x \leq n_r \leq Ln_y - n_x \quad (7)$$

The inequalities become an equality in case $n_d = 0$, that is, no disturbance. Equality with the lower bound holds if the matrix $(\mathcal{O} H_d)$ has full column rank. This shows that a parity space always exists ($n_r > 0$), if there are more observations than disturbances and if L is chosen large enough.

3. STOCHASTIC PARITY SPACE THEORY

3.1 Residual distribution

From (5.b) we get

$$\mathbf{E}(r_t) = W^T H_f F_t, \quad (8a)$$

$$\mathbf{Cov}(r_t) = W^T S W. \quad (8b)$$

For hypothesis testing and evaluation of detection probabilities, we need the distribution for the residual conditioned on that fault i has occurred. Assuming a constant fault magnitude m for a while, the fault term $f_t = m e^i$ in (2) can be written $H_f F_t = H_f^i \mathbf{1}_L m$ in (5). This notation will be used in the sequel.

Further, if the residual distribution is assumed Gaussian, then

$$(r_t | m, f^i) \in \mathbf{N}(\underbrace{m W^T H_f^i \mathbf{1}}_{\mu^i}, W^T S W). \quad (9)$$

The Gaussian distribution can be motivated in two ways:

- It is Gaussian if both V_t and E_t are Gaussian.
- It is approximately Gaussian by the central limit theorem when $\dim r_t \gg \dim V_t + \dim E_t$, which happens if the data window L is large enough. That is, asymptotically in L , it is Gaussian.

The parity space is unique up to a multiplication with a unitary matrix, so the projection matrix can be pre-multiplied with any full rank matrix. One solution is to take any $W^T \in \mathcal{N}_{[\mathcal{O} H_d]}$, but the main alternatives are the following ones:

Structured residual:

$$W^T = \mathcal{N}_{[\mathcal{O} H_d]} : W^T H_f^i \mathbf{1} = e^i, \quad (10)$$

Normalized residual:

$$W^T = \mathcal{N}_{[\mathcal{O} H_d]} : W^T S W = I. \quad (11)$$

Structured residuals are the suggested approach in many publications, as Keller (1999); White and Speyer (1987). The normalized residual can be derived from any parity space in the following way:

Normalized parity space:

$$\begin{aligned} \bar{r}_t &= \bar{W}^T (Y_t - H_u U_t), \\ \bar{W}^T &= (W^T S W)^{-1/2} W^T, \end{aligned} \quad (12)$$

for any parity space W^T , where $S = \mathbf{Cov}(E_t + H_v V_t)$.

We call $\|\bar{W}^T H_f^i \mathbf{1}_L\| = \|(W^T S W)^{-1/2} W^T H_f^i \mathbf{1}_L\|$ the *Fault to Noise Ratio* (FNR).

3.2 Illustrative example

As an illustrative example throughout the paper, we consider a DC motor subject to faults in input voltage (equivalent to a torque disturbance) or a velocity sensor offset:

$$\begin{aligned} x_{t+1} &= \begin{pmatrix} 1 & 0.3297 \\ 0 & 0.6703 \end{pmatrix} x_t + \begin{pmatrix} 0.0703 \\ 0.3297 \end{pmatrix} u_t \\ &\quad + \begin{pmatrix} 0.08 \\ 0.16 \end{pmatrix} v_t + \begin{pmatrix} 0.0703 & 0 \\ 0.3297 & 0 \end{pmatrix} f_t, \\ y_t &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x_t + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} f_t + e_t, \\ Q &= 0.01^2, \quad R = 0.1^2 \cdot I. \end{aligned}$$

With $L = 2$, the parity space residual becomes two-dimensional and easy to visualize graphically. There are also two possible faults to isolate. Figure 1 illustrates the distribution of the structured and normalized parity space residual, respectively, for the cases of no fault and fault.

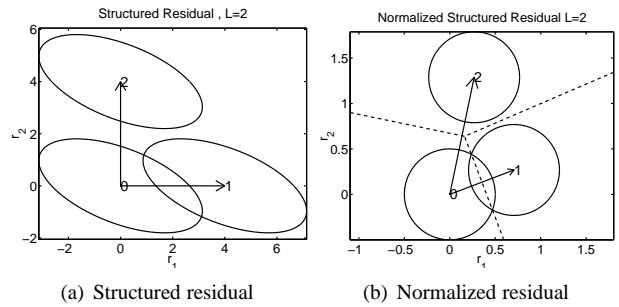


Fig. 1. Structured and normalized residual fault pattern with uncertainty ellipsoids for fault 1 and 2, respectively.

For the special case of known fault magnitude m , the decision boundary for the structured residual becomes a non-linear curve, while diagnosis for normalized residuals is based on the straight decision lines dashed in Figure 1(b). For the general case of unknown m , fault isolation is based on which fault vector has the smallest angle to the observed stochastic residual.

3.3 On-line algorithm

For normalized residuals $(\bar{r}_t|f=0) \in \mathbf{N}(0, I)$, we have $(\bar{r}_t^T \bar{r}_t|f=0) \in \chi^2(n_r)$. The χ^2 test provides a threshold h for detection, and fault isolation is performed by taking the closest fault vector in the sense of smallest angle difference (since the magnitude m of $\bar{\mu}$ is unknown in general).

Algorithm 1. On-line diagnosis

1. Compute a normalized parity space \overline{W} , e.g. (12).
2. Compute recursively:

$$\begin{aligned} \text{Residual: } \bar{r}_t &= \overline{W}^T (Y_t - H_U U_t) \\ \text{Detection: } \bar{r}_t^T \bar{r}_t &> h \\ \text{Isolation: } \hat{i} &= \arg \min_i \left\| \frac{\bar{r}_t}{\|\bar{r}_t\|} - \frac{\bar{\mu}^i}{\|\bar{\mu}^i\|} \right\|^2 \\ &= \arg \min_i \text{angle}(\bar{r}_t, \bar{\mu}^i) \end{aligned}$$

where $\bar{r}_t^T \bar{r}_t \in \chi^2(n_r)$ and $\text{angle}(\bar{r}_t, \bar{\mu}^i)$ denotes the angle between the two vectors \bar{r}_t and $\bar{\mu}^i$. A detection may be rejected if no suitable isolation is found ($\min_i \text{angle}(\bar{r}_t, \bar{\mu}^i)$ is too large) to improve false alarm rate.

For *diagnosability* of single faults, the only requirement is that all faults are mapped to different directions $\bar{\mu}^i$.

In a two-dimensional normalized residual space, as in the example in Figure 1(b), the probability for false alarm, P_{FA} , (incorrect detection) can be computed explicitly as

$$\begin{aligned} P_{FA} &= \int_{r_t^T r_t > h} \frac{1}{2\pi} e^{-\frac{r_t^T r_t}{2}} dr_t \\ &= \int_0^{2\pi} \int_h^\infty \frac{x}{2\pi} e^{-\frac{x^2}{2}} dx d\phi = e^{-\frac{h^2}{2}}. \end{aligned}$$

This means that the detection circle in Figure 1(b) has a radius h that can be directly computed from the desired false alarm rate. A more precise analysis is given below.

3.4 Off-line analysis

We can interpret the fault isolation step as a classification problem, and compare it to modulation in digital communication. Performance depends on the SNR, which here corresponds to FNR $m\|\bar{\mu}^i\|$. In vector coding in communication theory, using an additive Gaussian error assumption, it is straightforward to compute the risk for incorrect symbol detection. We will here extend these expressions from regular 2D (complex plane) patterns to general vectors in \mathcal{R}^{n_r} .

The risk of incorrect diagnosis can be computed exactly in the case of only two faults as follows. It relies on the symmetric distribution of \bar{r}_t , where the decision region becomes a line, as illustrated by the dashed lines in Figure 1(b). The first step is a change

of coordinates to one where one axis is perpendicular to the decision plane. Because of the normalization, the Jacobian of this transformation equals one. The second step is to marginalize all dimensions except the one perpendicular to the decision plane. All these marginals integrate to one. The third step is to evaluate the Gaussian error function. Here we use the definition

$$\text{erfc}(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

The result in \mathcal{R}^2 (cf. Figure 1) can be written

$$\begin{aligned} \text{P}(\text{diagnosis } i | \text{fault } m f^j) &= \\ \text{erfc} \left(m \|\bar{\mu}^j\| \sin\left(\frac{\alpha_i - \alpha_j}{2}\right) \right). \end{aligned} \quad (13)$$

In the general case, the decision line becomes a plane, and the line perpendicular to it is given by the projection distance to the intermediate line $\bar{\mu}^1 + \bar{\mu}^2$ as

$$m \left(\bar{\mu}^1 - \frac{(\bar{\mu}^1, \bar{\mu}^1 + \bar{\mu}^2)}{(\bar{\mu}^1 + \bar{\mu}^2, \bar{\mu}^1 + \bar{\mu}^2)} (\bar{\mu}^1 + \bar{\mu}^2) \right), \quad (14)$$

where $(a, b) = a^T b$ denotes a scalar product, and we get the following algorithm:

Algorithm 2. Off-line diagnosis analysis

1. Compute a normalized parity space W , e.g. (12).
2. Compute the normalized fault vectors $\bar{\mu}^i$ in the parity space.
3. The probability of incorrect diagnosis is approximately

$$\begin{aligned} \text{P}(\text{diagnosis } i | \text{fault } m f^j) &= \\ \text{erfc} \left(m \left\| \bar{\mu}^j - \frac{(\bar{\mu}^j, \bar{\mu}^j + \bar{\mu}^i)}{(\bar{\mu}^j + \bar{\mu}^i, \bar{\mu}^j + \bar{\mu}^i)} (\bar{\mu}^j + \bar{\mu}^i) \right\| \right) \end{aligned} \quad (15)$$

Here m denotes the magnitude of the fault. If this is not constant, replace $\bar{\mu}^i = \overline{W}^T H_f^i \mathbf{1}_L m$ with $\bar{\mu}^i = \overline{W}^T H_f^i M_t$.

For more than two faults, this expression is an approximation but, just as in the case of vector coding, generally quite a good one. The approximation becomes worse when there are several conflicting faults, which means that there are three or more fault vectors in about the same direction.

We can now define the diagnosability matrix P as

$$\begin{aligned} P^{(i,j)} &= \text{P}(\text{diagnosis } i | \text{fault } f^j), i \neq j \\ P^{(j,j)} &= 1 - \sum_{i \neq j} P^{(i,j)}. \end{aligned} \quad (16)$$

It tells us everything about fault association probabilities for normalized faults $m = 1$, and the off-diagonal elements are monotonically decreasing functions of the fault magnitude m .

Furthermore, in the classification we should allow the non-faulty class (0), where $f = 0$, to decrease the false alarm rate by neglecting residual vectors, though having large amplitude, being far from the known fault vectors. Consider for instance the residual $r_t =$

$(-1, -1)^T$ in Figure 1(b). This would most likely be caused by noise, not a fault. The missed detection probabilities are computed in a similar way as

$$P(\text{diagnosis } 0 | \text{fault } f^j) = \frac{1}{2} \operatorname{erfc} \left(\frac{m \|\bar{\mu}^j\|}{2} \right) \quad (17a)$$

$$P^{(0,0)} = 1 - \sum_j P^{(0,j)} < P_{FA}. \quad (17b)$$

For the DC motor residuals with residuals distributed according to Figure 1, the diagnosis probability matrix is

$$P^{(1:2,1:2)} = \begin{pmatrix} 0.995 & 0.005 \\ 0.005 & 0.995 \end{pmatrix}. \quad (18)$$

4. STOCHASTIC PARITY SPACE FAQ

We now have the basic tools for fault detection and isolation, and it remains to conclude if this approach is useful in practice. There are many natural questions, and the following subsections contain partial answers to some of them.

4.1 Why bother about noise?

If the fault is large enough, any diagnosis approach will work. The critical case is when the fault to noise ratio is rather small. Figure 2 shows how a two-dimensional parity space residual reacts to the two simulated faults in the DC motor. The upper plot in

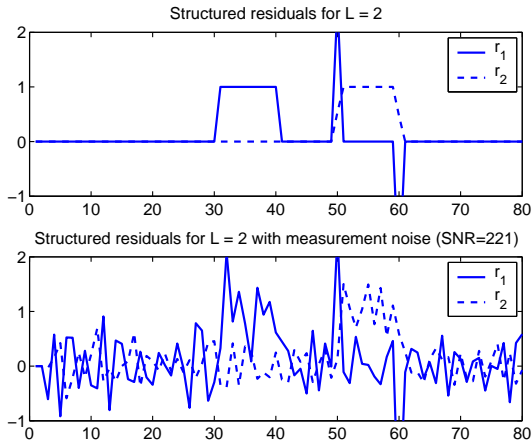


Fig. 2. Parity space residual for a DC motor for the residual distribution in Figure 1. The motor is subject to first an input voltage offset and then a sensor offset. The two residuals are designed to be non-zero for only one fault each. The lower plot illustrates extremely high sensitivity in residuals to measurement noise (SNR=221).

2 shows the case with no measurement noise $e_i(t) = 0$. The transient in the residual filter can be avoided with a slightly different design, otherwise the residual response is perfect, and the correct diagnosis can be made.

The lower plot in Figure 2, however, reveals that the design is rather sensitive to noise. Diagnosis based on

these residuals becomes more or less random. There are two approaches to robustify FDI:

- Pre-filter the system inputs and outputs, or post-filter the residuals. The post-filter can be a linear low-pass filter, non-linear threshold counter or CUSUM test.
- Make a stochastic analysis of the problem and an appropriate design.

The former approach is probably the dominating one in practice. The second approach is the soundest one if the model assumptions are plausible. More on these assumption follows below.

4.2 Does Gaussian noise exist in practice?

An often heard criticism to stochastic approaches to control and signal processing is that Gaussian noise is a theoretical concept which is seldom useful in practice. While there certainly are many diagnosis problems where a Gaussian model is not plausible, there are others where this a very well founded assumption.

Here, we illustrate the Gaussian assumptions with measurements from the IMU in Figure 3 as presented in Törnqvist (2006). The sensor and application are described in Chandaria et al. (2006).

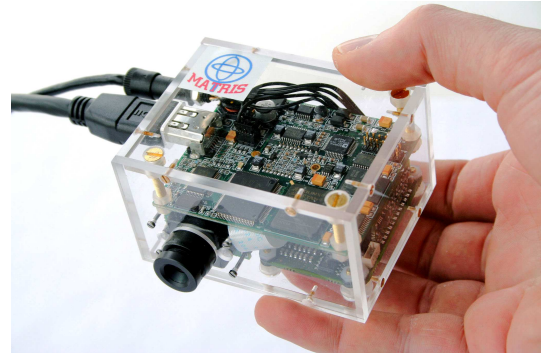


Fig. 3. Miniature Inertial Measurement Unit (IMU) by xsens.com with 3D measurements of acceleration, angular speeds and magnetic field as well as a built in camera.

Figure 4 shows the histogram of the nine available measurements from this unit. Clearly, the Gaussian bell fits well to the observations. The correlation function plots presented in Törnqvist (2006) confirm independence in time. However, as shown in Figure 5, the measurements from the magnetometer is not spatially white, so the covariance matrix is full. This is also true for accelerometers and gyroscope measurements.

4.3 Which are the design parameters?

The principal design parameter is the size L of the sliding window. When L is increased, more information is used in the diagnosis and a better result is expected. Figure 6 illustrates how the confidence

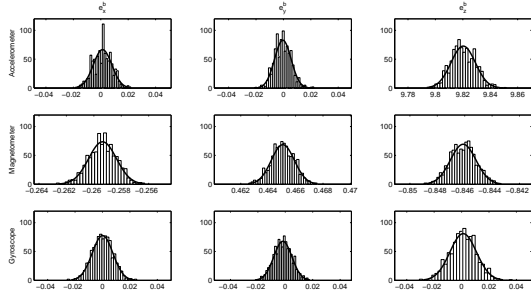


Fig. 4. Histograms of measurement errors from the IMU in Figure 3.

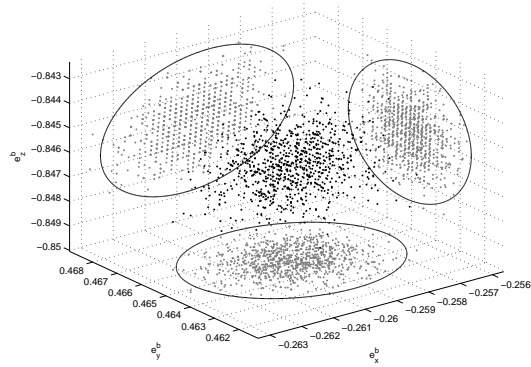


Fig. 5. Scatter plot of magnetometer measurements when the IMU is lying on a table. The projections using equal axis scaling indicates spatial correlation.

circles of the residuals become more separated when L is increased. A larger L means that it takes a longer time to get a complete window with faulty data, so the delay for detection should increase with L . This is the basic performance trade-off.

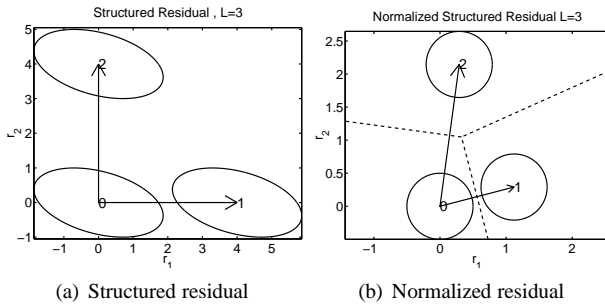


Fig. 6. Similar to Fig. 1, but with L increased from 2 to 3. The circles are now more separated, decreasing the risk of incorrect decisions.

Figure 7 shows a systematic evaluation of the design parameter L . The miss-classification probabilities decrease quickly in L . For the DC motor, we have $\lim_{L \rightarrow \infty} P = I$.

The diagnosis probability matrix can also be used for sensor choice and system design in general.

4.4 Why not using the GLR test?

Once the distribution of the stochastic inputs is known, Gaussian or not, the Generalized Likelihood Ratio

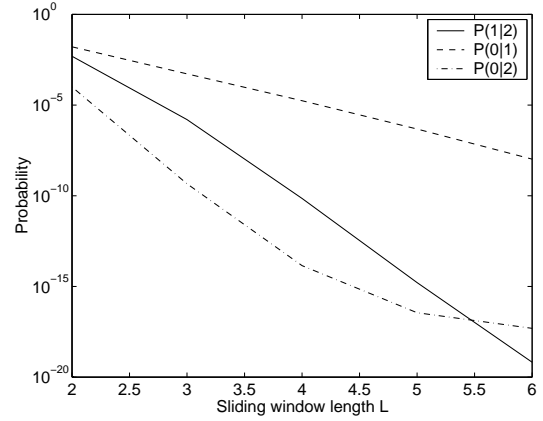


Fig. 7. Miss-classification probabilities in diagnosis as a function of sliding window length, where 18 gives the full matrix for $L = 2$.

(GLR) test is the natural approach, since it has many appealing properties. For instance, it is the Uniformly Most Powerful (UMP) test under certain conditions Kay (1998). The GLR test statistic for fault i for data in the sliding window considered here is given by

$$g_t^i = \frac{\max_{x,m} p(Y_t|x_{t-L+1}, m, \text{fault } i)}{\max_x p(Y_t|x_{t-L+1}, \text{no fault})} \quad (19)$$

for fault i .

However, it is shown in Törnqvist (2006) that the GLR test and stochastic parity space tests are equivalent. That is, which approach is used does not matter in the end, since they both lead to the same algorithms.

4.5 How many faults can be detected?

In the case that *one or more* faults may be present, a necessary condition for diagnosability is $n_r \geq n_f$. That is, there must be at least as many residuals as faults. A sufficient condition is that $W^T[H_f^1 \mathbf{1}, \dots, H_f^{n_f} \mathbf{1}]$ has full column rank.

For *single faults*, these conditions can be relaxed. Similar to vector coding techniques, there can be arbitrarily many faults that are diagnosable if $n_r \geq 2$. To illustrate this, a more complex system is used (F16 pitch dynamics with five states, three inputs, three outputs and six faults). See Hagenblad et al. (2003) for details. Figure 8 illustrates a four dimensional residual ($L = 2$).

The residual indicates that fault four is not detectable, and faults two and five are hard to distinguish. This is also revealed in the diagnosis probability matrix from Algorithm 2:

$$P = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & 0 & 0.0000 & 0.0000 \\ 0.0000 & 0.5980 & 0.0000 & 0 & 0.4020 & 0.0001 \\ 0.0000 & 0.0000 & 0.9999 & 0 & 0.0001 & 0.0000 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0000 & 0.4020 & 0.0001 & 0 & 0.5415 & 0.0564 \\ 0.0000 & 0.0001 & 0.0000 & 0 & 0.0564 & 0.9436 \end{pmatrix} \quad (20)$$

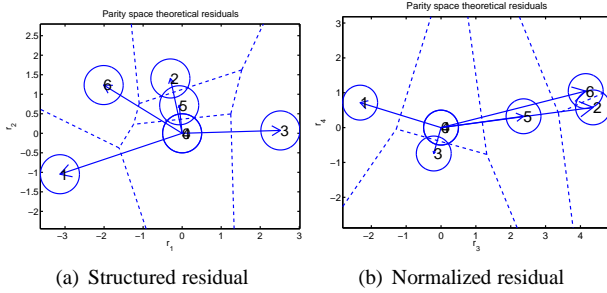


Fig. 8. Illustration of the residuals from parity space for no fault (0) and fault 1–6, respectively, but here in another basis. The decision lines for fault isolation for known m are indicated.

4.6 What happens if the fault magnitude is time-varying

In the analysis, it was assumed that the fault is mf^i , where m is the magnitude and f^i the direction of the fault. The original model (5) allows for general fault profiles in the fault term $H_f F_t$. However, letting F_t be arbitrary gives too much freedom and any noise influence can be interpreted as a fault, since H_f generally spans the whole measurement space for Y_t . A natural model-based approach to model incipient faults that grows gradually in time or fluctuates slowly is to use a dedicated model for the fault magnitude m_t . For instance, a low order polynomial model is used in Hendeby and Gustafsson (2006); Törnqvist (2006). This will reduce the span of H_f to a low dimensional space, where noise and fault can be better separated.

4.7 What happens if the model is non-linear?

Possible approaches for non-linear stochastic systems include:

- Linearize the model in the same spirit as the extended Kalman filter. Here, the linearized model is used over a time horizon of size L , which puts a further requirement on this design variable. It cannot be chosen arbitrarily large, since in that case the linearized model becomes poorer.
- Use non-linear filters as the particle filter, see Azimi-Sadjadi and Krishnaprasad (2002); Kadiramanathan et al. (2002); Vaswani (2004); Hendeby and Gustafsson (2006).

The first approach is suitable for the stochastic parity space approach. The IMU in the previous section is here re-visited. The intended task is to estimate orientation based on the IMU in Figure 3. This is done by dead-reckoning gyroscope measurements. To avoid drift in time, supporting information with respect to the earth frame is needed. The magnetometer gives two degrees of freedom, and when the IMU is still, the accelerometer vector can be used as an inclinometer, once the gravity field is subtracted. This is state of the art, as described in for instance Roetenberg et al. (2005).

The orientation dynamics as used in Törnqvist (2006) is based on quaternions as state vector x_t . The process model $x_{t+1} = f(x_t, u_t, v_t)$ is quite non-linear, and so are the measurement equations $y_t = h(x_t, e_t)$. The dynamics is quite fast, so a fixed linearization state cannot be used. The orientation may change too much during the sliding window. Instead, the model is linearized around the nominal trajectory in the sliding window, which is obtained by integrating the gyroscope signals (here considered as inputs u_t). This leads to an error state model, with much slower dynamics. The size of the residual computed from the accelerometers and magnetometers, respectively, is 26 ($L = 10, n_y = 3, n_x = 4, n_d = 0$ in (7)), so the test statistic should be distributed as $\chi^2(26)$ in theory.

Figure 9 shows how the two test statistics vary in time. There is a perfect match between theory and experiment, which validates the involved assumptions:

- The model is correct.
- The linearization error is negligible.
- The Gaussian assumption, and all other stochastic assumptions, hold.

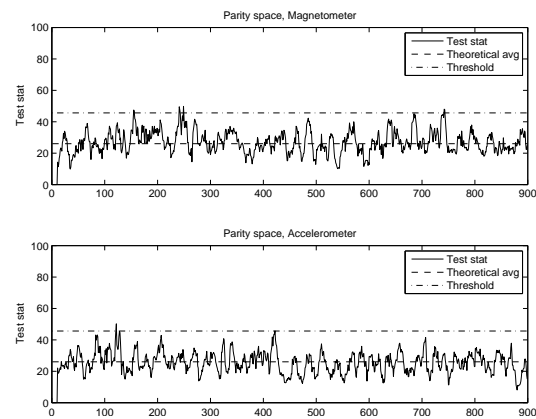


Fig. 9. Sum of squared residuals $r_t^T r_t$ in case of no fault.

Just to illustrate detection performance, the test statistic in Figure 10 shows what happens during an incipient magnetic field disturbance.

4.8 What happens if the fault is not additive?

Multiplicative faults as changes in system parameters are generally considered to be harder to detect than additive faults. To apply the stochastic parity space approach, the linearization technique around a nominal state trajectory in the previous section can be applied. The use of linearized error models is for instance standard in aircraft navigation applications. The success of diagnosis depends on the linearization error, and, of course, on the fault to noise ratio.

A further alternative is to include system parameters in the state vector. This can sometimes be done for sub-models on the ARX form, leading to a linear model with additive changes in the system parameters.

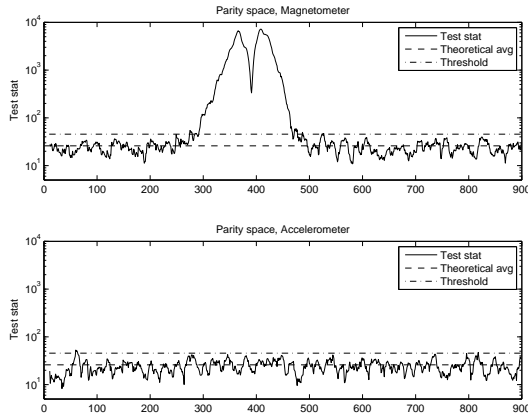


Fig. 10. Sum of squared residuals $r_t^T r_t$ in case of fault disturbance in magnetic field caused by a metallic object passing by. The axes are set equal to illustrate the spatial correlation.

4.9 What happens if the noise is non-Gaussian?

The stochastic parity space approach is model-based, and its result depends heavily on the accuracy of the model. If the model is incorrect, the performance of the diagnosis can be expected to degrade. This is true in general, but not for the Gaussian assumption on noise.

If the dynamic part of the model is correct, and the first and second order moments of the noise are also correct, the central limit theorem indicates that the residual r_t is still Gaussian distribution when the sliding window size L is large enough. That is, even if the model of the system is incorrect, the model of the residual satisfies all assumptions and Algorithm 1 applies.

However, one can sometimes do much better! As described in Hendeby and Gustafsson (2006) in an information theoretic setting, the Gaussian distribution is worst case. That is, of all true systems with the same linear dynamics and first and second order moments of process noise and measurement noise, respectively, the Gaussian one is the one that gives the worst upper bound on performance. This upper bound is provided by the asymptotic GLR test, which is the UMP (uniformly most powerful) test.

It turns out that this upper bound depends on something called the intrinsic accuracy of the probability density function of the noise process, see Kay and Sengupta (1993); Hendeby (2005). The more non-Gaussian noise in terms of intrinsic accuracy, the higher performance bound.

A practical question is whether there is a feasible algorithm to compute the GLR test, and how far from the asymptotic assumption the sensor information is. The upper bound indicates a potential benefit for more sophisticated algorithms, but gives no promises. One has to try out on a case to case basis.

Consider again the DC motor example. Figure 11 is taken from Hendeby and Gustafsson (2006). It shows the ROC (receiver operator characteristics) curve

where the trade-off between false alarms and detection performance is illustrated. It shows four curves:

- The upper bound for a Gaussian system.
- The upper bound for the simulated system (which has a heavy-tail distribution).
- The performance of an approximate GLR test, computed by plugging in the best linear unbiased (KF) state estimate into (19).
- The performance of the GLR test, computed by plugging in the minimum variance state estimate, approximated by the PF, into (19).

The natural but ad-hoc approximation of the GLR test, neglecting the distribution of the noises, lead to ambiguous results. The GLR test gives a substantial performance gain. That is, in this case it pays off a lot to use all available information in the fault detector.

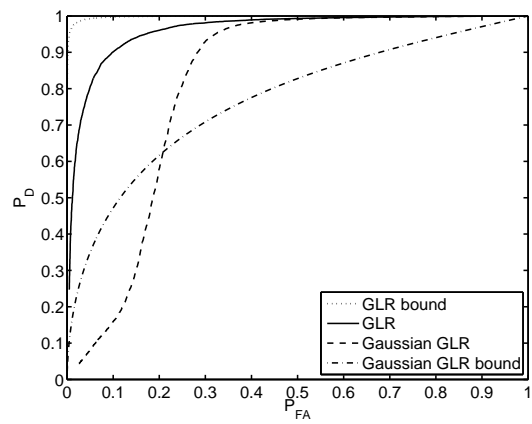


Fig. 11. Plot of probability of detection versus false alarm rate when the detection threshold is varied for GLR test on the DC motor.

4.10 What happens if there is no model available?

Approaches to compute the projection P in (1) include the following options, starting with the one based on a physical model and ending with a completely data-driven approach:

- (1) The model-based parity space, where $P(A, B, C, D)$ depends on the known state space model, described by the quadruple (A, B, C, D) .
- (2) System identification gives $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$, from which the parity space can be approximated as $P(\hat{A}, \hat{B}, \hat{C}, \hat{D})$. One here needs to know the structure of the state space model.
- (3) Certain subspace approaches to system identification, as the one in Verhaegen (1994), provide a way to directly compute \hat{P} Zhang and Ding (2005). Again, one needs to know the structure of the state space model.
- (4) The principal component analysis (PCA) approach Dunia et al. (1996); Qin and Li (1999), where one directly estimates \hat{P} from data. Compared to above, one needs to know the state order, but not how the data Z_t is split into inputs and outputs. That is, causality is no concern in

the PCA approach. This is one main reason for its wide spread in chemical engineering, where sometimes thousands of variables are measured, see Chiang et al. (2001).

All cases can be unified in the following algorithm:

- (1) Determine the projection matrix P as outlined above.
- (2) Estimate the residual covariance matrix from training data from a fault-free system, and normalize the residuals:

$$\hat{S} = \frac{1}{N-L} \sum_{t=L+1}^N r_t r_t^T \quad (21a)$$

$$\bar{r}_t = \hat{S}^{-1/2} r_t. \quad (21b)$$

- (3) Get learning data from each faulty mode and compute the normalized fault vector as

$$\bar{\mu}_i = \frac{1}{N^i - L} \sum_{t=L+1}^{N^i} \bar{r}_t^i. \quad (22)$$

Figure 12 shows how the probability of detection for each fault in the DC motor example depends on the fault magnitude using Monte Carlo simulations. The same data sets are used for all three approaches.

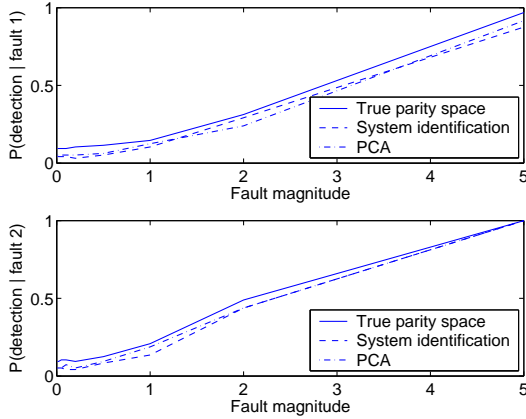


Fig. 12. Comparison of residuals from parity space on known model, parity space using estimated model, and PCA residuals. The identified model and PCA analysis were performed on the same data set, thus using the same information.

The difference in performance is not very significant, despite the large difference in prior information. In an attempt to understand the relation between PCA and parity space, consider the following split of our model:

$$\begin{aligned} Z_t &= \begin{pmatrix} \mathcal{O} \\ 0 \end{pmatrix} x_{t-L+1} + \begin{pmatrix} H_f & H_u & H_v & I \\ 0 & I & 0 & 0 \end{pmatrix} \begin{pmatrix} F \\ U \\ V \\ E \end{pmatrix} \\ &= P_x x_{t-L+1} + P_r r_t. \end{aligned} \quad (23)$$

PCA splits the covariance matrix of data Z_t based on its eigenvalues into two parts: the model and the residuals. This results in an expression of the same form as (23). We conclude the following relations:

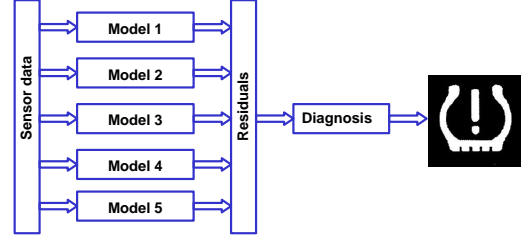


Fig. 13. Structure of residual generation in a tire pressure monitoring system (TPMS) based on multiple physical models.

- The split of eigenvalues should give $\text{rank}(P_x) = n_x$.
- The inputs in the data are revealed by zero rows in P_x , so causality is cleared out.
- The residual part must also explain dynamics in the input data, and changes in input dynamics can be mixed up with system changes.
- It cannot be guaranteed that the eigenvalues of the system are larger than the other ones, so the PCA split based on sorted eigenvalues can be dubious.

Despite the two last points, the example demonstrates excellent performance, though these points should be kept in mind.

4.11 What happens if there are many models?

In some applications, there might be many models that relate to the fault. One example is road-friction estimation in Section 5.1 where a change in friction can be detected from sound, visual information, longitudinal and lateral dynamic behaviour of the vehicle, tire vibrations, or even from the driver behaviour, see Müller et al. (2003); Gustafsson (1997).

Another currently hot topic in automotive safety is Tire Pressure Monitoring Systems (TPMS), which is proposed to be mandatory on the US market by NHTSA. State of the art for software systems that do not use sensors mounted in the tire is that a warning can be given if 1-3 wheels have low tire pressure, but not isolate the fault. Thus, they cannot detect the case of diffusion, when all four tires loose the same amount of pressure over time.

To overcome these limitations, a system based on many different models is proposed in Gustafsson et al. (2001), see Figure 13. Each model delivers one residual that can be used for detection but only partial isolation of 15 different faults (each tire can be faulty or non-faulty, giving 16 different combinations). For instance, model 1 may provide the residual $r = p_1 - p_2$, which is the difference of pressure in two wheels, based on vehicle dynamics. A non-zero residual indicates a fault in tire 1 or 2. A positive residual gives a vote for isolating a fault in tire 2 (the pressure cannot increase), and so on. Another residual may be $r = p_3 - p_4$, so two residuals would suffice to isolate all

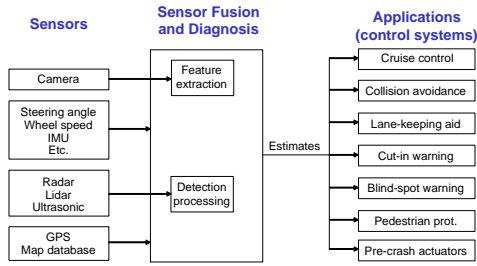


Fig. 14. Automotive sensor fusion and diagnosis as a generalization of the approach in Figure 13.

single faults. Even more models are needed to isolate all combinations in a robust way.

Similar multi-model based diagnosis units utilizing *residual fusion* are natural to introduce in vehicles, as the number of sensors for driver assistance and safety systems increases, and the sensor fusion software becomes more integrated over the different sub-systems, see Figure 14 and Gustafsson (2005).

5. DETECTION FOR ADAPTIVE FILTERING

The aim of diagnosis is usually either to warn an operator, or to feedback the fault message to a controller that adapts itself to the new faulty conditions to make the best of the situation. Another well-known principle is to feedback the alarm to adaptive filters, as illustrated in Figure 15. This is a structure with good practical potential in applications. It enables a method to design non-linear filters in a systematic way to overcome the inherent trade-off between tracking speed and noise suppression in linear adaptive filters. These include algorithms as Recursive Least Squares (RLS), Least Mean Square (LMS) and its normalized version as well as the Kalman filter for state estimation. We here show three completely different applications based on essentially the same principle.

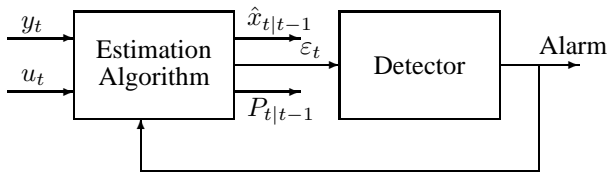


Fig. 15. The estimation algorithm delivers residuals, which are used in the detector to decide whether a change has occurred or not. If a change is detected this information is fed back for use in the estimation algorithm.

5.1 Tire-road friction estimation

Road-friction estimation can be based on comparing the wheel slip s_t (how much faster a driven wheel rotates relative to its absolute speed) and normalized tractions force μ_t . A phenomenon noticeable in practice Dieckmann (1992) is that the linear relation between these computable quantities depend on friction, see

Figure 16. This would suggest a linear model with an offset,

$$s_t = \theta_1 \mu_t + \theta_2 + e_t. \quad (24)$$

This model cannot be found in the tire literature, but is still a model that gives promising results for friction estimation Gustafsson (1997, 1998).

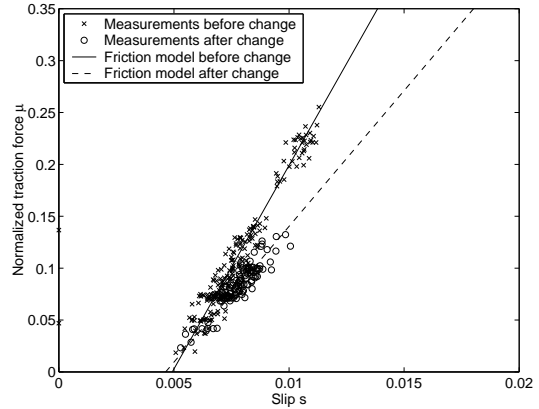
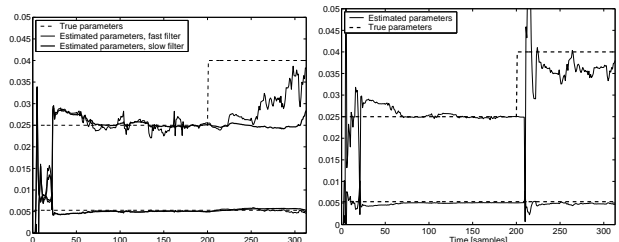


Fig. 16. Principle for slip-based tire-road friction estimation: The scatter plot of normalized tire force versus wheel slip indicates the linear relation (24).

The excitation in this model can be very poor during normal driving, since the tire force to overcome air drag is relatively constant. Obviously, it is not easy to estimate a straight line to a cluster of data in Figure 17. That means that a linear adaptive filter must be tuned to be quite slow. Figure 17(a) shows one slow adaptive filter that gives sufficient accuracy, and one faster filter that is too noisy to base driver alerts on, but still not quick enough to warn in time. The CUSUM supported Kalman filter proposed in Gustafsson (1997, 1998) solves this problem as illustrated in Figure 17(b).

The system has been extensively tested, and used by several road authorities to monitor road conditions. The current status is that further residuals are needed to make the system robust to (1) all tire and road combinations (2) without the need for special calibration procedures and (3) to be used by ordinary drivers not educated on the system.



(a) Linear filters

(b) The non-linear filter in Figure 15

Fig. 17. Comparison of two linear adaptive filters and the adaptive filter based on the change detection feedback loop of Figure 15.

5.2 Combined road and vehicle tracking

In collision avoidance algorithms, it is important to predict both the road ahead and other vehicles relative position and lane assignments. Road prediction is in current systems solved by computer vision systems inside the camera, while vehicle tracking is an algorithm that takes input from both radar, lidar and camera Gustafsson (2005). One approach based on joint tracking and road prediction is suggested in Schön et al. (2006). The state vector contains lateral deviation from the own lane center of all cars and two road parameters: curvature and clothoid (derivative of curvature). These states are central for emergency lane assist systems warning the driver that the host car is leaving its lane, and collision avoidance systems, respectively. The idea of joint estimation is to (1) utilize the lateral movements of leading vehicles for road tracking and (2) to use road prediction to detect lane changes of leading vehicles. The exchange of information between these two estimation tasks is thus crucial, and it is important to know if the observed lateral movements in Figure 18 depend on a change in curvature or lane.

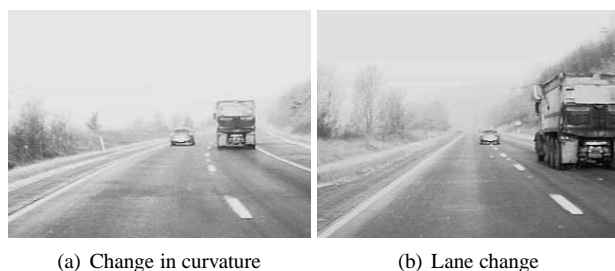


Fig. 18. Two different changes to detect and isolate.

Figure 19 shows curvature tracking, when the leading vehicle initiates the lane change in Figure 18(b) at time 4272 s. The curvature estimate for two Kalman filters is shown in Figure 19(a). The fast KF is not accurate enough, and the slow filter has a very long recovery time after a lane change. The CUSUM boosted adaptive filter combines the good features of fast and accurate tracking. Further, the plots shows a filter allowing for a smoothing delay, where the estimated change point is used to re-process measurements after the change to eliminate the radar measurements from the Kalman filter.

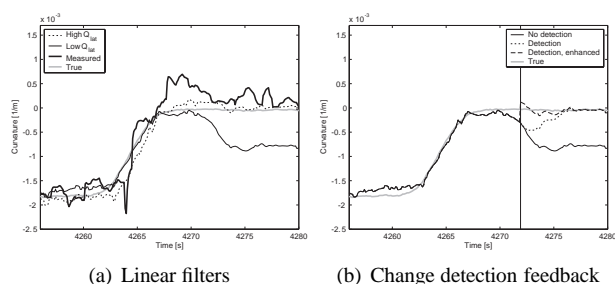


Fig. 19. The curvature tracking of different adaptive filters.



Fig. 20. Acoustic guitar feedback simulation.

5.3 Simulated guitar feedback

Acoustic guitar feedback is a roaring phenomenon that some people like and others not. However, guitarists who want to practice playing with feedback must play with loud volume, which might be disturbing. The idea illustrated in Figure 20 is to simulate this phenomenon in software, so the guitarist can use headphones and still practice, or apply post-processing feedback effects in studio recordings.

The system in Gustafsson and Kilberg (2005) uses the structure in Figure 15. The adaptive filter contains a frequency tracker for which string is hit, and its local variations around the nominal frequency caused intentionally by the guitarist. A rather advanced model decides which harmonics that would have survived in the feedback path. However, to be useful at all, the latency (time-delay) must be very small. The change in tone should appear almost instantaneously after the guitarist hits a new string or blocks the strings with his hand to turn off feedback as he is used to. A feedback delay more than 10ms (corresponding to 3 m feedback path, or 441 samples with 44.1 kHz sampling, or one period of the tone C) is an upper bound in practice. This specification is easily achieved with the CUSUM feedback to the frequency tracker.

6. CONCLUSIONS

The paper has demonstrated how statistical signal processing theory can bring insight and contribute to fault detection and isolation problems, and *vice versa* how fault detection algorithms can improve statistical signal processing algorithms. An example was used to explain the intuition of the stochastic parity space and the involved model assumptions and algorithms. Several applications were used to motivate how the model assumptions can be verified in practice and how these assumptions can be relaxed to get useful algorithms for non-linear non-Gaussian models. Three applications were used to demonstrate how a simple feedback mechanism from a CUSUM detector can boost adaptive filters when needed to overcome the inherent trade-off between tracking speed and estimation accuracy in linear adaptive filters.

REFERENCES

B. Azimi-Sadjadi and P.S. Krishnaprasad. Change detection for nonlinear systems: a particle filtering

- approach. In *Proceedings of the American Control Conference*, Anchorage, AK, May 2002.
- M. Basseville and I.V. Nikiforov. *Detection of abrupt changes: theory and application*. Information and system science series. Prentice Hall, Englewood Cliffs, NJ., 1993.
- J. Chandaria, G. Thomas, B. Bartczak, K. Koeser, R. Koch, M. Becker, G. Bleser, D. Stricker, C. Wohlleber, M. Felsberg, F. Gustafsson, J. Hol, T.B. Schön, J. Skoglund, P.J. Slycke, and S. Smeitz. Real-time camera tracking in the matrix project. In *Proceedings of International Broadcasting Convention*, Amsterdam, September 2006. IEE.
- L.H. Chiang, E.L. Russell, and R.D. Braatz. *Fault Detection and Diagnosis in Industrial Systems*. Springer, 2001.
- E.Y. Chow and A.S. Willsky. Analytical redundancy and the design of robust failure detection systems. *IEEE Transactions on Automatic Control*, 29(7):603–614, 1984.
- T. Dieckmann. Assessment of road grip by way of measured wheel variables. In *Proceedings of FISITA*, London, June 1992.
- X. Ding, L. Guo, and T. Jeansch. A characterization of parity space and its application to robust fault detection. *IEEE Transactions on Automatic Control*, 44(2):337–343, 1999.
- R. Dunia, S.J. Qin, T.F. Edgar, and T.J. McAvoy. Use of principal component analysis for sensor fault identification. *Computers & Chemical Engineering*, 20(971):S713–S718, May 1996.
- J. Gertler. *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, Inc, 1998.
- J. Gertler. Fault detection and isolation using parity relations. *Control Engineering Practice*, 5(5):653–661, 1997.
- F. Gustafsson. Challenges in signal processing for automotive safety systems (plenary paper). In *IEEE Statistical Signal Processing Workshop*, Bordeaux, 2005. IEEE.
- F. Gustafsson. Slip-based estimation of tire – road friction. *Automatica*, 33(6):1087–1099, 1997.
- F. Gustafsson. Estimation and change detection of tire – road friction using the wheel slip. *IEEE Control System Magazine*, 18(4):42–49, 1998.
- F. Gustafsson. *Adaptive filtering and change detection*. John Wiley & Sons, Ltd, 2000.
- F. Gustafsson. *Adaptive filtering and change detection*. John Wiley & Sons, Ltd, 2001.
- F. Gustafsson and U. Kilberg. A system and a method for simulating acoustic feedback. Patent application PCT (filed 2005-11-17 by softube.se), 2005.
- F. Gustafsson, M. Drevö, U. Forssell, M. Löfgren, N. Persson, and H. Quicklund. Virtual sensors of tire pressure and road friction. In *Society of Automotive Engineers World Congress*, number SAE 2001-01-0796, Detroit, 2001.
- A. Hagenblad, F. Gustafsson, and I. Klein. A comparison of two methods for stochastic fault detection: the parity space approach and principal component analysis. In *IFAC Symposium on System Identification*, Rotterdam, NL, 2003.
- G. Hendeby. Fundamental estimation and detection limits in linear non-Gaussian systems. LIU-TEK-LIC-2003:1199, Dept. of Electrical Engineering, Linköping University, Sweden, 2005.
- Gustaf Hendeby and Fredrik Gustafsson. Detection limits for linear non-Gaussian state-space models. In *6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, Beijing, P. R. China, 2006.
- V. Kadirkamanathan, P. Li, M.H. Jaward, and S.G. Fabri. Particle filtering based multiple-model approach to fault diagnosis in nonlinear stochastic systems. *International Journal of Systems Science*, 33:259–265, 2002.
- S.M. Kay. *Fundamentals of signal processing – detection theory*. Prentice Hall, 1998.
- S.M. Kay and D. Sengupta. Detection in incompletely characterized colored non-gaussian noise via parametric modeling. *IEEE Transactions on Signal Processing*, 41(10):3066–3070, 1993.
- J.-Y. Keller. Fault isolation filter design for linear stochastic systems. *Automatica*, 35:1701–1706, 1999.
- S. Müller, M. Uchanski, and K. Hedrick. Estimation of the maximum tire-road friction coefficient. *Journal of Dynamic Systems, Measurement, and Control*, 125:607–617, 2003.
- S.J. Qin and W. Li. Detection, identification and reconstruction of faulty sensors with maximized sensitivity. *AIChE Journal*, 45:1963–1976, 1999.
- Daniel Roetenberg, Henk J. Luinge, Chris T. M. Baten, and Peter H. Veltink. Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(3):395–405, 2005.
- T. Schön, A. Eidehall, and F. Gustafsson. Lane departure detection for improved road geometry estimation. In *IEEE Conference on Intelligent Vehicles (IV)*, Tokyo, 2006. IEEE.
- D. Törnqvist. Statistical fault detection with applications to imu disturbances. LIU-TEK-LIC-2003:1258, Dept. of Electrical Engineering, Linköping University, Sweden, 2006.
- N. Vaswani. Bound on errors in particle filtering with incorrect model assumptions and its implication for change detection. In *Proceedings of Acoustics, Speech, and Signal Processing Conference (ICASSP)*, 2004.
- M. Verhaegen. Identification of the deterministic part of mimo state space models given in innovations form from input-output data. *Automatica*, 30:61–74, 1994.
- J.E. White and J.L. Speyer. Detection filter design: Spectral theory and algorithms. *IEEE Transactions on Automatic Control*, AC-32(7):593–603, 1987.
- P. Zhang and S.X. Ding. A model-free approach to fault detection of periodic systems. In *Proceedings of the 2005 IEEE International Symposium on Intelligent Control*, Limassol, Cyprus, June 2005.