STOCHASTIC FAULT DIAGNOSABILITY IN PARITY SPACES

F. Gustafsson

Department of Electrical Engineering, Linköpings universitet,
SE-581 83 Linköping, Sweden
Fax: +46-13-282622 Email: fredrik@isy.liu.se

Abstract: We here analyze the parity space approach to fault detection and isolation in a stochastic setting. Using a state space model with both deterministic and stochastic unmeasurable inputs we show a formal relationship between the Kalman filter and the parity space.

Based on a statistical fault detection and diagnosis algorithm, the probability for incorrect diagnosis is computed explicitly, given that only a single fault with known time profile has occurred. An example illustrates how the matrix of diagnosis probabilities can be used as a design tool for performance optimization with respect to, for instance, design variables and sensor placement and quality.

Keywords: fault detection, diagnosis, Kalman filtering, adaptive filters, linear systems

1. INTRODUCTION

The parity space approach to fault detection (Basseville and Nikiforov, 1993; Chow and Willsky, 1984; Ding *et al.*, 1999; Gertler, 1997; Gertler, 1998) is an elegant and general tool for additive faults in linear systems and is based on intuitively simple algebraic projections and geometry. It provides a tool to compute a residual vector that is zero when there is no fault in the system and reacts to different faults in different patterns, enabling a simple algorithm for diagnosis (deciding which fault actually occurred). Examples on simulated data often show very good results. Consider for instance Figure 1, where a DC motor is subject to first an offset in control input and then an offset in velocity sensor.

The upper plot shows how structured parity space residuals correctly points out which fault has occurred. A main drawback is that the approach does not take measurement errors and state noise into consideration as in the classical Kalman filter literature. The lower plot in Figure 1 illustrates the high sensitivity to even quite a small measurement noise.
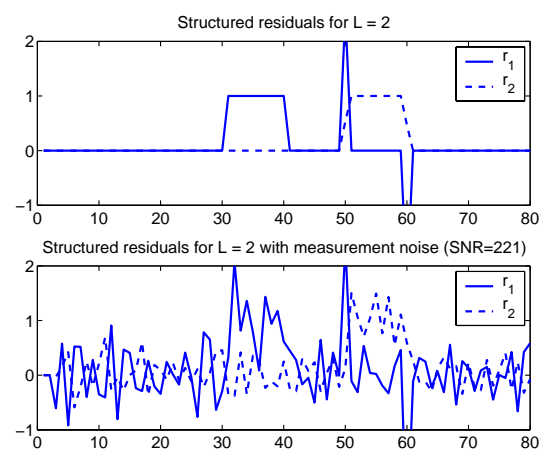

Fig. 1. Parity space residual for a DC motor, as described in Section 7, subject to first a input voltage offset and then a sensor offset. The two residuals are designed to be non-zero for only one fault each. The lower plot illustrates extremely high sensitivity in residuals to measurement noise (SNR=221).

We here mix the linear state space models used in fault detection and Kalman filtering, treating deter-

ministic and stochastic disturbances in different ways. Previous work in this direction include (Keller, 1999), (Basseville and Nikiforov, 1993) (Ch. 7) and (Gertler, 1998) (Ch. 11). Related ideas using principal component analysis are found in the chemical diagnosis literature as (Chiang *et al.*, 2001; Dunia and Qin, 1998) This work is a continuation of (Gustafsson, 2001), where an additive fault was included in an augmented state vector, and observability of the fault was used as the tool to assess diagnosability. In this paper, an explicit expression for $P^{i,j} = P(\text{diagnosis } j \mid \text{fault } i)$ is given for any parity space, and the parity space is optimally designed to minimize these probabilities in order to improve sensitivity issues in diagnosis. The approach relies on spatial *and* temporal whitening of the parity space residuals.

2. STOCHASTIC PARITY SPACES

The linear system is here defined as the state space model

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

The inputs are of four different 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 $f_t = m_t f^i$, where f^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 $D_{f,t}$.
- Stochastic unknown state disturbance v_t and measurement noise e_t, as are 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 and Gaussian (for the analysis) with zero mean and covariance matrices Q_t and R_t, respectively.
- The initial state is treated as an unknown variable, in contrast to the Kalman filter literature where it is assumed Gaussian.

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.

We next formulate the diagnosis task as a recursive problem applied to a sliding window. Stack L signal values to define $Y_t = (y_{t-L+1}^T, \ldots, y_t^T)^T$ *etc.* for all

signals. Introducing appropriate Hankel matrices H_s for an arbitrary input signal s and an observability matrix $\mathcal{O}$, it is easily shown that (1) can be written as

$$
\begin{aligned}
Y_t - H_u U_t &= \\
&\mathcal{O} x_{t-L+1} + H_d D_t + H_f F_t + H_v V_t + E_t. \quad (2)
\end{aligned}
$$

where

$$
\mathcal{O} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{pmatrix}, \quad H = \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & & 0 \\ & \ddots & \ddots & 0 \\ CA^{L-2}B & \cdots & CB & D \end{pmatrix}. \quad (3)
$$

Define a residual to be used for detection and diagnosis as

$$
\begin{aligned}
r_t &= w^T (Y_t - H_u U_t) \quad &(4a) \\
&= w^T (\mathcal{O} x_{t-L+1} + H_d D_t + H_f F_t + H_v V_t + E_t) \\
&= w^T (H_f F_t + H_v V_t + E_t). \quad &(4b)
\end{aligned}
$$

The point here is that r_t is designed to belong to the parity space, defined by $w^T [\mathcal{O} \ H_d] = 0$. That is, the parity space is defined to be insensitive to 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, $k = t - L + 1, \ldots, t$, provided that there is no stochastic term present ($e_k = 0$, $v_k = 0$ for $k = t - L + 1, \ldots, t$) and no fault, $f_k = 0$, $k = t - L + 1, \ldots, t$. Any deviation from zero is either due to the noise or one of the possible faults, and the diagnosis task is to distinguish these causes.

The maximal dimension of the residual vector is given by

$$
L(n_y - n_d) - n_x \leq \max_w n_r \leq L n_y - n_x
$$

Here the size of any signal s_t is denoted as $n_s = \dim(s_t)$. 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 ($\max_w n_r > 0$) if there are more observations than disturbances, if L is chosen large enough.

Another approach, not pursued here, is to apply *fault decoupling*, where each residual is designed separately by the condition $w^T [\mathcal{O} \ H_d \ H_f F^{-i}] = 0$. Here F^{-i} is a fault vector that excites all faults except for fault i.

3. PROBLEM FORMULATION FOR A SPECIAL CASE

Let us consider the case of a scalar measurement and $Q = 0$, in which case we can write the observations as $Y = Y_t - H U_t$ where the components are independent stochastic variables with

$$
\mathbf{E}(y_k) = m_k, \quad \mathbf{Var}(y_k) = \sigma_k^2.
$$

Assume for the moment that the fault magnitude is constant $m_k = m$. With a Gaussian assumption (not

needed for this discussion indeed), we can write $y_k \in \mathbf{N}(m, \sigma_k^2)$. Let μ be a vector with all elements equal to m. A general linear estimator is given by

$$\hat{m} = w^T Y \in \mathbf{N}(w^T \mu, w^T \text{diag}(\sigma_k^2) w).$$

Unbiasedness is imposed by the constraint $\sum_k w_k = 1$. This expression can be normalized to

$$(w^T \text{diag}(\sigma_k^2) w)^{-1/2} w^T Y \triangleq \bar{w}^T Y \in \mathbf{N}(\bar{w}^T \mu, I).$$

This expression can be used as a test statistic to test if the mean is non-zero. This leads to the *most powerful test*, where w can be expressed as

$$w = \arg \max_w (w^T \text{diag}(\sigma_k^2) w)^{-1} w^T w, \quad (5)$$

subject to the unbiased condition $\sum_k w_k = 1$. A first try is to use the sample moving average (MA) by letting $w_k = 1/L$, which gives

$$\hat{m}^{MA} = (w^{MA})^T Y = \frac{1}{L} \sum y_k \in \mathbf{N}\left(m, \frac{\sum \sigma_k^2}{L^2}\right).$$

However, the optimal solution to (5) is to use the minimum variance (MV) estimate

$$\hat{m}^{MV} = (w^{MV})^T Y = \frac{\sum y_k / \sigma_k}{\sum 1/\sigma_k} \quad (6a)$$

$$\in \mathbf{N}\left(m, \frac{L}{(\sum 1/\sigma_k)^2}\right). \quad (6b)$$

One can show that $\frac{L}{(\sum 1/\sigma_k)^2} \leq \frac{\sum \sigma_k^2}{L^2}$ with equality if and only if $\sigma_k = \sigma$ are all constant, regardless of the choice of L.

The corresponding residual r, subject to the constraint $\mathbf{Var}(r) = 1$, for non-zero mean detection is thus $r = \bar{w}^T Y$, where

$$\bar{w}_k = \frac{1}{\sqrt{L}\sigma_k}.$$

This minimum variance principle is exactly what will be used for the parity space design. The literature suggests to further low-pass filter the residuals using IIR filters (yielding what is sometimes referred to as exponential window moving average (EWMA)). However, this basically corresponds to a w with a larger time window L and improves computational complexity rather than performance. As a conclusion, a minimum variance unbiased estimator corresponds to the optimal test statistic for non-zero mean detection.

4. STOCHASTIC ANALYSIS OF RESIDUAL

To simplify the notation, assume time invariant covariance matrices $\mathbf{Cov}(E_t) = I_L \otimes R$ and $\mathbf{Cov}(V_t) = I_L \otimes Q$, respectively ($\otimes$ denotes the Kronecker product). A unity fault f^i ($\|f^i\| = 1$) with constant magnitude m gives a fault vector $F_t = mF^i$ in (2) and we have

$$(r_t | mf^i) = w^T(H_v V_t + E_t + mH_f F^i)$$

$$\in \mathbf{N}(m \underbrace{w^T H_f F^i}_{\mu^i}, w^T S w)$$

where

$$S = H_v(I_L \otimes Q)H_v^T + I_L \otimes R. \quad (7)$$

That is, each fault is mapped onto a vector μ^i with a covariance matrix $w^T S w$. The case of non-constant fault magnitude is commented on in Algorithm 2. We can normalize the residual distribution to obtain a minimum variance residual, which will enable probability calculations in Section 6, as

$$\bar{r}_t = (w^T S w)^{-1/2} r_t = \underbrace{(w^T S w)^{-1/2} w^T}_{\bar{w}^T}(Y_t - H_u U_t)$$

$$(8)$$

in which case we get

$$(\bar{r}_t | mf^i) = \bar{w}^T(H_v V_t + E_t + mH_f F^i) \quad (9a)$$

$$\in \mathbf{N}(m \underbrace{\bar{w}^T H_f F^i}_{\bar{\mu}^i}, I) = \mathbf{N}(m\bar{\mu}^i, I) \quad (9b)$$

The uncertainty in the residual is now symmetric. More specifically, the covariance ellipsoid around μ^i now becomes a circle around $\bar{\mu}^i$. See Figure 2 for an illustration. We can now formulate the design task as:
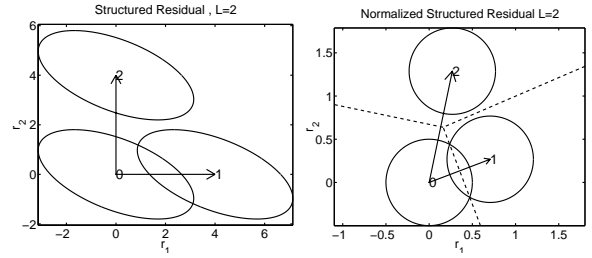

Fig. 2. Original and normalized structured residual fault pattern with uncertainty ellipsoids for fault 1 and 2, respectively. Solid line is for unnormalized residuals, and dashed line after normalization. The dashed line is the optimal decision region.

Definition 1. The normalized parity space is defined as

$$\bar{r}_t = \bar{w}^T(Y_t - H_u U_t), \ \bar{w}^T = (w^T S w)^{-1/2} w^T,$$

$$(10)$$

for any parity space w^T, where S is defined in (7). The parity space is unique up to a multiplication with a unitary matrix. We call $\bar{w}^T H_f = (w^T S w)^{-1/2} w^T H_f$ the *Fault to Noise Ratio* (FNR), since it explicitly tells us how much each fault is amplified relative to other faults.

One interpretation of this definition is that the parity space residual is whitened spatially and temporally. We stress that a transformation of the residual space affects how the fault vectors look like, but not the ability to make diagnosis. The point to keep in mind is that there are many obtainable parity spaces, the sliding window size L affects its dimension n_r and the weighting matrix w its stochastic properties. The latter requires stochastic analysis, and here we get help from the Kalman filter theory.

5. RELATION TO THE KALMAN FILTER

One interpretation of the Kalman filter is that it computes the minimum variance estimate of the state vector, given the measurements available. One characterization of this is that the *innovations* (residuals in our setting) are independently distributed Gaussian variables. Let us reformulate the Kalman filter and pose the following problem: What is the minimum variance estimate of the state x_k, $k = t - L + 1, t - L + 2, \ldots t$, given the observations y_k, $k = t - L + 1, t - L + 2, \ldots t$. The solution can be implemented by the *Kalman smoother*, initialized at time $t - L + 1$, using $P_{t-L+1} = \infty \times I$, or formally more correct, using the information filter with $P_{t-L+1}^{-1} = 0$. Another more direct approach is to use the signal model (2).

The *sliding window Kalman filter* in (Gustafsson, 2001) is here re-phrased using the following notation: $A^\dagger$ denotes the pseudo-inverse $(A^T A)^{-1} A^T$ with the property $A^\dagger A = I$. The orthogonal projection matrix $A^\perp$ is defined as $A^\perp = I - A(A^T A)^{-1} A^T = I - A A^\dagger$ with the properties $A^\perp A^\perp = A^\perp$ (projection) and $A^\perp A = 0$. We define a row basis for $A^\perp$ (square matrix) as the null space of A and denote it by $\mathcal{N}_A$ (thick matrix). Note that the parity space now can be written $w^T = \mathcal{N}_{[\mathcal{O} H_d]}$.

Generally, a linear state estimator can be written

$$\hat{x}_{t-L+1} = K(Y_t - H_u U_t) \in \mathbf{N}(x_{t-L+1}, KSK^T).$$

It generates residuals as

$$\begin{aligned}
\varepsilon_t &= Y_t - \hat{Y} = Y_t - \mathcal{O}\hat{x}_{t-L+1} - H_u U_t \\
&= (I - \mathcal{O}K)(Y_t - H_u U_t) \\
&= (I - \mathcal{O}K)(H_d D_t + H_v V_t + E_t + H_f m F^i) \\
&\in \mathbf{N}((I - \mathcal{O}K)(H_d D_t + H_f m F^i), \\
&\qquad (I - \mathcal{O}K)S(I - \mathcal{O}K)^T).
\end{aligned}$$

The link to the normalized minimum variance parity space follows in three steps:
1. The minimum variance state estimator (the Kalman filter) is derived for a sliding window.
2. The size of the prediction error vector ε_t is decreased, where purely deterministic parts are removed. This yields the minimum variance parity space residual $r_t = w^T(Y_t - H_u U_t)$. (It is a parity space, since $r_t = 0$ when d_t, f_t, e_t, v_t are all zero.)
3. The residual is whitened spatially and temporally, which gives $\bar{r}_t$.

The *Kalman filter* state estimate is by definition the minimum variance estimate and requires pre-whitening of data, so first we normalize (2),

$$S^{-1/2}(Y_t - H_u U_t) =$$
$$S^{-1/2}(\mathcal{O}x_{t-L+1} + H_d D_t + H_v V_t + E_t + H_f m F^i),$$

and we get

$$K^{KF} = (S^{-1/2}\mathcal{O})^\dagger = (\mathcal{O}^T S^{-1} \mathcal{O})^{-1})\mathcal{O}^T S^{-1/2}.$$

We can here note that the *observer approach*, where $K^{obs} = \mathcal{O}^\dagger = (\mathcal{O}^T \mathcal{O})^{-1})\mathcal{O}^T$, and all other choices

of K will give a larger covariance matrix for $\hat{x}_{t-L+1}$, and thus produce less efficient residuals. The Kalman filter prediction errors are distributed as

$$\begin{aligned}
\varepsilon_t^{KF} \in \mathbf{N}(&(I - \mathcal{O}(\mathcal{O}^T S^{-1}\mathcal{O})^{-1}\mathcal{O}^T S^{-1}) \\
&\cdot (H_d D_t + H_f m F^i), \\
&S - \mathcal{O}(\mathcal{O}^T S^{-1}\mathcal{O})^{-1}\mathcal{O}^T).
\end{aligned}$$

Now, note that since the data projection matrix $W_{KF}^T \triangleq I - \mathcal{O}(\mathcal{O}^T S^{-1}\mathcal{O})^{-1}\mathcal{O}^T S^{-1}$ is singular, the covariance matrix of the prediction errors is singular, and there are many linear combinations of ε_t^{KF} that are always zero, independent of the data. By introducing a basis w_{KF}^T for the row space of W_{KF}^T, we get a residual generator $r_t = w_{KF}^T(Y_t - H_u U_t)$ that is low-dimensional but still contains all information relevant for diagnosis.

The parity space and Kalman filter are related as follows:

- The parity space design in (4) is unconstrained, while it in this section is of the form of a row basis for $I - \mathcal{O}K$, where K defines an unbiased state estimator.
- The observer (w_{obs}) and Kalman filter (w_{KF}) automatically compute a residual space for the case of no disturbance $D_t = 0$, where the latter gives minimum variance residuals.
- Since $r_t = w_{KF}^T(Y_t - H_u U_t)$ has the same size as the parity space residual defined in (4) (namely $Ln_y - n_x$) and it does not depend on the initial state, it is by definition a parity space residual. Since it is also minimum variance, the whitened version must coincide with the $\bar{r}_t$ in (9b) up to a unitary transformation.
- The Kalman filter innovation can be transformed to a parity space where also the disturbance is decoupled (besides the initial state), by another projection $\bar{\bar{r}}_t = (w_{KF}^T H_d)^\perp \bar{r}_t$.

6. DIAGNOSIS ALGORITHM

Since $(\bar{r}_t | f = 0) \in \mathbf{N}(0, I)$ 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. The following algorithm is well-known in the statistical literature:

Algorithm 1. **On-line diagnosis**
1. Compute a normalized parity space $\bar{w}$, *e.g.* (8).
2. Compute recursively:

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

where $\text{angle}(\bar{r}_t, \bar{\mu}^i)$ 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 the two-dimensional residual space, as the considered DC motor example, the probability for false alarm, P_{FA}, (incorrect detection) can be computed explicitly as $P_{FA} = e^{-h^2}$, which means that the threshold design is to choose P_{FA} and then letting $h = \sqrt{-2\log(P_{FA})}$. Note that the true false alarm rate may be lower if we reject alarms where $\min_i \text{angle}(\bar{r}_t, \bar{\mu}^i)$ is too large. A more precise analysis is given below.

We can interpret the diagnosis step as a classification problem, and compare it to modulation in digital communication. Performance depends on the SNR, which in this context often is denoted $\frac{E_b}{N_0}$. SNR here corresponds to FNR $m\|\bar{\mu}^i\|$. In modulation 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, using the Gaussian noise assumption. 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 2. The first step is a change of coordinates to one where one axis is perpendicular to the decision plane. 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 (Matlab) definition

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

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

$$\mathbf{Prob}(\text{diagnosis } i|\text{fault } mf^j) =$$
$$\frac{1}{2}\text{erfc}\left(m\|\bar{\mu}^j\|\sin(\frac{\alpha_i - \alpha_j}{2})\right).$$

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

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

and we get the following algorithm:

Algorithm 2. **Off-line diagnosis analysis**
1. Compute a normalized parity space w, *e.g.* (8).
2. Compute the normalized fault vectors $\bar{\mu}^i$ in the parity space as in (9b).

3. The probability of incorrect diagnosis is approximately

$$\mathbf{Prob}(\text{diagnosis } i|\text{fault } mf^j)$$
$$= \frac{1}{2}\text{erfc}\left(m\left\|\bar{\mu}^j - \frac{(\bar{\mu}^j, \bar{\mu}^j + \bar{\mu}^i)}{(\bar{\mu}^j + \bar{\mu}^i, \bar{\mu}^j + \bar{\mu}^i)}(\bar{\mu}^j + \bar{\mu}^i)\right\|\right)$$

Here m denotes the magnitude of the fault. If this is not constant, we replace $\bar{\mu}^i = \bar{w}^T H_f F^i$ in (9b) with $\bar{\mu}^i = \bar{w}^T H_f \text{diag}(m_{t-L+1}, \ldots, m_t)F^i$ and let $m = 1$.

For more than two faults, this expression is an approximation but as in modulation theory generally quite a good one. The approximation becomes worse when there are several conflicting faults, which means that there are two or more fault vectors in about the same direction. Note that the Gaussian assumption using erfc is justified by the central limit theorem for large L, due to the averaging inherent in $w^T(E_t + H_v V_t)$.

We can now define the diagnosability matrix P as

$$P^{(i,j)} = \mathbf{Prob}(\text{diagnosis } i|\text{fault } f^j), i \neq j$$
$$P^{(j,j)} = 1 - \sum_{i \neq j} P^{(i,j)}. \tag{11}$$

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. The miss-classification probabilities are computed in a similar way as

$$P(\text{diagnosis } 0|\text{fault } f^j) = \frac{1}{2}\text{erfc}\left(\frac{m\|\bar{\mu}^j\|}{2}\right) \tag{12a}$$
$$P^{(0,0)} = 1 - \sum_j P^{(0,j)} < P_{FA}. \tag{12b}$$

7. EXAMPLE

Consider a sampled state space model of a DC motor with continuous time transfer function

$$G(s) = \frac{1}{s(s+1)} = \frac{1}{s^2 + s}.$$

The state variables are angle (x^1) and angular velocity (x^2) of the motor. The derivation of the corresponding state space model is straightforward, and can be found in any textbook in control theory. Sampling with sample interval $T_s = 0.4$ s gives

$$A = \begin{pmatrix} 1 & 0.3297 \\ 0 & 0.6703 \end{pmatrix}, B_u = \begin{pmatrix} 0.0703 \\ 0.3297 \end{pmatrix}, B_v = \begin{pmatrix} 0.08 \\ 0.16 \end{pmatrix}$$

$$B_d = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, B_f = \begin{pmatrix} 0.0703 & 0 \\ 0.3297 & 0 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$D_u = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, D_d = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, D_f = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is assumed that both x_1 and x_2 are measured. Here we have assumed that the fault is either an input voltage disturbance (or torque disturbance) or a velocity sensor offset. The matrices in the sliding window model become for $L = 2$:

$$\mathcal{O} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0.3297 \\ 0 & 0.6703 \end{pmatrix}, \ H_u = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0.0703 & 0 \\ 0.3297 & 0 \end{pmatrix},$$

$$H_f = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0.0703 & 0 & 0 & 0 \\ 0.3297 & 0 & 0 & 1 \end{pmatrix},$$

and

$$\bar{w}^T = \mathcal{N}_{[\mathcal{O} \ H_d]} =$$
$$\begin{pmatrix} -0.6930 & -0.1901 & 0.6930 & -0.0572 \\ 0.0405 & -0.5466 & -0.0405 & 0.8354 \end{pmatrix}. \quad (13)$$

The residual space with structured residuals, as shown in Figure 2, is

$$w^T = \begin{pmatrix} -1 & -0.3297 & 1 & 0 \\ 0 & -0.6703 & 0 & 1 \end{pmatrix}. \quad (14)$$

The difference of the parity spaces generated by (13) and (14), respectively, is illustrated in Figure 2. The faults in the normalized parity space are not orthogonal, but on the other hand the decision region is particularly simple.

The probability matrix (11) is here

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

Note that this is independent of the choice of original parity space (13), (14) or the Kalman filter. By increasing the length of the sliding window to $L = 3$, we get a much better performance with a probability matrix that is very close to diagonal and a very small missed detection probability. The confidence circles of the structured residuals in Figure 3 are more separated than the ones in Figure 2.
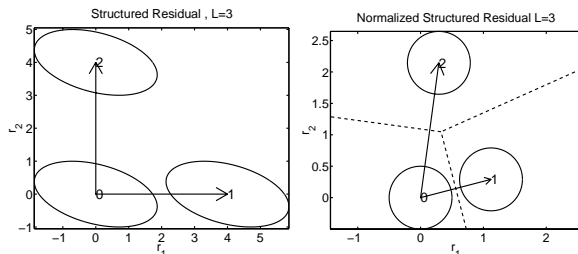

Fig. 3. Similar to Fig. 2, but with L increased from 2 to 3. The circles are now more separated, decreasing the risk of incorrect decisions.

Figure 4 shows a systematic evaluation of the design parameter L. 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. On the other hand, the miss-classification probabilities decreases quickly in L.
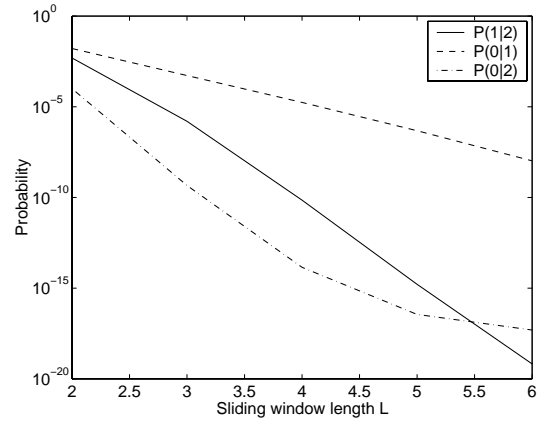

Fig. 4. Miss-classification probabilities in diagnosis as a function of sliding window length.

As a final illustration, one can investigate how much we lose in performance using a cheaper velocity sensor with variance 10 instead of 1, and the result is

$$P^{(1:2,1;2)} = \begin{pmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{pmatrix}.$$

The ten times larger miss-classification probabilities can be compensated for by sacrificing a short delay for detection and using a longer sliding window.

8. CONCLUSIONS

We have here introduced the normalized parity residual space for additive faults in linear stochastic systems. It was shown how this can be derived in a Kalman filter framework. We have derived explicit formulas for miss-classification probabilities as a definition of diagnosability, and these depend critically on the fault to noise ratio. An example illustrated how the diagnosability matrix can be used as a design tool with respect to sensor quality and design parameters.

REFERENCES

Basseville, M. and I.V. Nikiforov (1993). *Detection of abrupt changes: theory and application*. Information and system science series. Prentice Hall. Englewood Cliffs, NJ.

Chiang, L.H., E.L.Russell and R.D. Braatz (2001). *Fault Detection and Diagnosis in Industrial Systems*. Springer.

Chow, E.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. Jeinsch (1999). A characterization of parity space and its application to robust fault detection. *IEEE Transactions on Automatic Control* **44**(2), 337–343.

Dunia, R. and S.J. Qin (1998). Joint diagnosis of process and sensor faults using principal component analysis. *Control Engineering Practice* **6**, 457–469.

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, F. (2001). Stochastic observability and fault diagnosis of additive changes in state space models. In: *ICASSP'01*.

Keller, J.Y. (1999). Fault isolation filter design for linear stochastic systems. *Automatica* **35**(10), 1701–1706.