Parity Functions as Universal Residual Generators and Tool for Fault Detectability Analysis

Mattias Nyberg
Vehicular Systems, ISY, Linköping University
S-581 83 Linköping, Sweden.
e-mail: matny@isy.liu.se

September 8, 1997

Abstract

An important issue in diagnosis research is design methods for residual generation. One method is the Chow-Willsky scheme. Here an extension to the Chow-Willsky scheme, called the ULPE scheme is presented. It is shown that previous extensions to the Chow-Willsky scheme can not generate all possible parity equations for some linear systems. This is the case when there are dynamics controllable from fault but not from the inputs or disturbances. The ULPE scheme is able to handle also this case since it is, for both discrete and continuous linear systems, shown to be a universal design method for perfectly decoupling residual generators. Also included are two new straightforward conditions on the process for fault detectability and strong fault detectability respectively. A general condition for strong fault detectability has not been presented elsewhere. It is shown that fault detectability and strong fault detectability can be seen as system properties rather than properties of the residual generator.

1 Introduction

In the seventies, research about using *analytical redundancy* for fault detection and diagnosis was intensified. One main area of interest was fault detection for aircrafts and especially their control and navigation systems. In a work within this field, Potter and Suman (1977) defined *parity equation* and *parity function* (and also *parity space* and *parity vector*). This was originally a concept for utilizing analytical redundancy in the form of linear *direct redundancy*. In 1984 the concept was generalized by Chow and Willsky (1984) to include also dynamic systems, i.e. to utilize *temporal redundancy*. However only discrete time parity equations were considered.

Included in their description is a design method to derive such parity equations. This method will in this work be referred to as the *Chow-Willsky scheme*. Based on this method, a number of extensions have been proposed. This class of design methods will in this paper be denoted *Chow-Willsky-like schemes*. One important extension, provided by Frank (1990), includes also decoupling of disturbances and non-monitored faults into the design. Another important extension, was made by Höfling (1993) who showed that Chow-Willsky-like schemes are valid also for continuous linear systems. In this paper the Universal Linear Parity Equation (ULPE) scheme, which belongs to the class of Chow-Willsky-like schemes, is presented. In addition to earlier extensions it has the property that for arbitrary linear system, all possible perfectly decoupling parity functions can be obtained. Among other extensions is for example the handling of the case when perfect decoupling is not possible (Lou *et al.*, 1986). Further Gertler (1991) has defined ARMA parity equations, which are equivalent to linear residual generators. Also non-linear parity equations have been discussed in literature. However no other uses of the term *parity equations*, than in accordance with the definitions made by Potter and Suman (1977) and later extended by Chow and Willsky (1984), have been widely accepted in the research community.

The objective of designing parity functions is to use them in residual generators. Many other design methods for linear residual generations exists; all resulting in similar or identical residual generators. Surveys of most well known methods can be found in (Gertler, 1991; Frank, 1993; Patton, 1994). However parity equations are attractive because they involve only simple mathematics. Also attractive is that the set of all possible parity functions and also all residual generators can, as shown in the ULPE scheme, be completely parameterized by a single vector.

Conditions for fault detectability have, in the literature, been treated in different contexts. Here new straightforward conditions for fault detectability and strong fault detectability are derived and presented. These conditions are formulated in the context of parity equations and answers the question whether there exists a residual generator in which the fault becomes detectable or strongly detectable. It is shown that both fault detectability and strong fault detectability can be seen as properties of the system. A condition for strong fault detectability, has to the author's knowledge, not been presented elsewhere. Both conditions are derived using the ULPE scheme.

In Section 2, the ULPE scheme is presented and it is shown that previous Chow-Willsky-like schemes are not able to generate all parity equations for some linear system. This is the case when there exists dynamics controllable only from the fault. The ULPE scheme is able to handle this case, and as shown, the ULPE scheme is able to generate all linear parity equations for arbitrary linear system. In Section 3, it is demonstrated how the ULPE scheme can be used to obtain any residual generator. Therefore the ULPE scheme is also a universal method for residual generator design for linear systems. Based on the concepts of the ULPE scheme, Section 4 provides the conditions for fault detectability and strong fault detectability analyses. To make the main text more readable, the proofs of some lemmas have been placed in an appendix.

2 Parity Equations

This section describes the ULPE (Universal Linear Parity Equation) scheme, and the relation to previous Chow-Willsky-like schemes. The purpose of parity equations is for use in residual generators. It is assumed that the principle of *structured residuals* is used. This means that the goal is to construct a residual that is sensitive to some faults, referred to as *monitored fault*, and not sensitive to other faults, i.e. *non-monitored faults*, or disturbances. We say that the non-monitored faults and disturbances are to be *decoupled*.

First parity equation (also called *parity relation*) and parity function are defined formally. These definitions are in accordance with the definitions of generalized parity equation and generalized parity function in (Chow and Willsky, 1984). To shorten the notation, the word "generalized" is here omitted.

Definition 1 ***[Parity Equation].*** *A parity equation is an equation that can, if all terms are moved to the right-hand side, be written as*

$$0 = \mathbf{A}(\sigma)\mathbf{y}(t) + \mathbf{B}(\sigma)\mathbf{u}(t)$$

where $\mathbf{A}(\sigma)$ and $\mathbf{B}(\sigma)$ are row vectors of polynomials in σ, $\mathbf{u}(t)$ and $\mathbf{y}(t)$ are the system input and output vectors, and σ denotes the differentiate operator p or the time-shift operator q. The equation is satisfied if no faults are present.

Definition 2 ***[Parity Function].*** *A parity function is a function $h(\mathbf{u}(t), \mathbf{y}(t))$ that can be written as*

$$h(\mathbf{u}(t), \mathbf{y}(t)) = \mathbf{A}(\sigma)\mathbf{y}(t) + \mathbf{B}(\sigma)\mathbf{u}(t)$$

where $\mathbf{A}(\sigma)$ and $\mathbf{B}(\sigma)$ are row vectors of polynomials in σ, $\mathbf{u}(t)$ and $\mathbf{y}(t)$ are the system input and output vectors, and σ denotes the differentiate operator p or the time-shift operator q. The value of the function is zero if no faults are present.

Chow and Willsky (1984) also defined the *order* of the parity equation (and function) as the highest degree α of σ^{α}, that is present in the parity equation.

Parity equations or parity functions for linear systems can conveniently be designed with Chow-Willsky-like schemes. When there exists dynamics that is controllable only from faults, the previous Chow-Willsky-like schemes are not universal as will be shown in Example 1.

2.1 The ULPE Scheme

Following is a description of an extension of the Chow-Willsky scheme, called the ULPE scheme. In addition to previous Chow-Willsky-like schemes, the ULPE scheme has the important property that it is universal in the sense that for an arbitrary linear system, continuous or discrete, all parity equations can be obtained. The description is formulated in a general framework valid for both the continuous and discrete case. The notation σ is used to denote the differentiate operator p and time-shift operator q for the continuous and discrete case respectively.

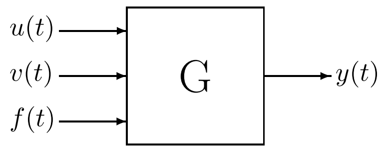

Figure 1: The system with inputs u (known or measurable), v (disturbances), f (the fault), and output y.

Consider the linear system illustrated in Figure 1. The system has an m-dimensional output $y(t)$ and three kinds of inputs: known or measurable inputs collected in the k-dimensional vector $u(t)$, disturbances in the k_d-dimensional vector $v(t)$, and the monitored fault $f(t)$. For simplicity reasons, we assume that only one fault affects the system, i.e. f is scalar. The extension to more than one fault is straightforward. To achieve isolation, it is desirable that non-monitored faults do not affect the residual, i.e. decoupling. Such faults are included in v.

This system can be described by the following realization:

$$\begin{aligned} \begin{bmatrix} \sigma x \\ \sigma z \end{bmatrix} &= \begin{bmatrix} A_x & A_{12} \\ 0 & A_z \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \\ & \quad \begin{bmatrix} E \\ 0 \end{bmatrix} v + \begin{bmatrix} K_x \\ K_z \end{bmatrix} f && \text{(1a)} \\ y &= [C_x C_z] \begin{bmatrix} x \\ z \end{bmatrix} + Du + Jv + Lf && \text{(1b)} \end{aligned}$$

where $[x\,z]^T$ is the $n = n_x + n_z$ -dimensional state. It is assumed that the realization has the property that the state x is controllable from $[u\,v]^T$ and the state z is controllable from the fault f. It is assured from Kalman's decomposition theorem that such a realization always exists. Finally it is assumed that the state z is asymptotically stable, which is the same as saying that the whole system is stabilizable.

Before going into more details, we define the following matrices, that will be used to simplify the notation of the system description (1):

$$A = \begin{bmatrix} A_x & A_{12} \\ 0 & A_z \end{bmatrix} \quad E = \begin{bmatrix} E_x \\ 0 \end{bmatrix} \quad C = [C_x \ \ C_z] \quad K = \begin{bmatrix} K_x \\ K_z \end{bmatrix} \tag{2}$$

Now by substituting (1a) into (1b), we can obtain σy as

$$\begin{aligned} \sigma y &= C_x \sigma x + C_z \sigma z + D\sigma u + J\sigma v + L\sigma f = \\ &= C_x A_x x + C_x A_{12} z + C_z A_z z + C_x B u + D\sigma u + C_x E v + \\ &\quad + C_x K_x f + C_z K_z f \end{aligned}$$

By continuing in this fashion for $\sigma^2 y \ \ldots \ \sigma^\rho y$, the following equation can be obtained:

$$Y(t) = R_x x(t) + R_z z(t) + QU(t) + HV(t) + PF(t) \tag{3}$$

where

$$Y(t) = \begin{bmatrix} y(t) \\ \sigma y(t) \\ \vdots \\ \sigma^\rho y(t) \end{bmatrix} \qquad R_x = \begin{bmatrix} C_x \\ C_x A_x \\ \vdots \\ C_x A_x^\rho \end{bmatrix}$$

$$R_z = \begin{bmatrix} C_z & 0 & 0 & \ldots \\ C_x A_{12} & C_z & 0 & \ldots \\ \vdots & & \ddots & \\ C_x A_x^{\rho-1} A_{12} & \ldots & C_x A_{12} & C_z \end{bmatrix} \begin{bmatrix} I \\ A_z \\ \vdots \\ A_z^\rho \end{bmatrix}$$

$$Q = \begin{bmatrix} D & 0 & 0 & \ldots \\ C_x B & D & 0 & \ldots \\ \vdots & & \ddots & \\ C_x A_x^{\rho-1} B & \ldots & C_x B & D \end{bmatrix} \qquad U(t) = \begin{bmatrix} u(t) \\ \sigma u(t) \\ \vdots \\ \sigma^\rho u(t) \end{bmatrix}$$

$$H = \begin{bmatrix} J & 0 & 0 & \ldots \\ C_x E_x & J & 0 & \ldots \\ \vdots & & \ddots & \\ C_x A_x^{\rho-1} E_x & \ldots & C_x E_x & J \end{bmatrix} \qquad V(t) = \begin{bmatrix} v(t) \\ \sigma v(t) \\ \vdots \\ \sigma^\rho v(t) \end{bmatrix}$$

$$P = \begin{bmatrix} L & 0 & 0 & \ldots \\ CK & L & 0 & \ldots \\ \vdots & & \ddots & \\ CA^{\rho-1} K & \ldots & CK & L \end{bmatrix} \qquad F(t) = \begin{bmatrix} f(t) \\ \sigma f(t) \\ \vdots \\ \sigma^\rho f(t) \end{bmatrix}$$

The size of Y is $(\rho+1)m \times 1$, R_x is $(\rho+1)m \times n_x$, R_z is $(\rho+1)m \times n_z$, Q is $(\rho+1)m \times (\rho+1)k$, U is $(\rho+1)k \times 1$, H is $(\rho+1)m \times (\rho+1)k_d$, F is $(\rho+1) \times 1$, P is $(\rho+1)m \times (\rho+1)$, and V is $(\rho+1)k_d \times 1$. The constant ρ determines the maximum order of the parity equation. This can be seen by studying the definitions of the vectors Y and U. The choice of ρ is discussed in Section 4.

Now, with a column vector w of length $(\rho+1)m$, a parity function $h(y,u)$ can be formed as

$$h(y,u) = w^T(Y - QU) \tag{4}$$

From Equation (3) it follows that the value h_v of the parity function also can be written

$$h_v = w^T(R_x x + R_z z + HV + PF) \tag{5}$$

Since the parity function must be zero in the fault free case and the disturbances must be decoupled, Equation (5) implies that w must satisfy

$$w^T [R_x\, H] = 0 \tag{6}$$

For use in fault detection, it is also required that the parity function is non-zero in the case of faults. This is assured by letting

$$w^T [R_z\, P] \neq 0 \tag{7}$$

In conclusion, the ULPE scheme is a method for designing parity functions useful for fault detection. A parity function is constructed by first setting up all the matrices in (3) and then finding a w such that (6) and (7) are fulfilled.

An algorithm in accordance with previous Chow-Willsky-like schemes, is obtained by replacing Equation (6) and (7) with $w^T [R\,H] = 0$ and $w^T P \neq 0$ respectively, where the matrix R is defined as $R = [R_x\, R_z]$.

2.2 The ULPE Scheme is Universal

The presented scheme has the property that all parity functions for a linear system can be designed by different choices of ρ and w. This is addressed in the following lemma:

Theorem 1. *Any parity equation satisfying a model can be obtained from the ULPE scheme.*

Proof. Any parity equation that satisfies a model can be written

$$M \begin{bmatrix} Y \\ U \end{bmatrix} = 0 \tag{8}$$

where M is a row vector of length $(\rho+1)(m+k)$, m the number of outputs, and k the number of inputs. Let M be partitioned as $[M_1\, M_2]$ and assume that there are no faults, which implies that z is zero. Then by using (3), (8) can be rewritten as

$$
\begin{aligned}
[M_1\, M_2] \begin{bmatrix} Y \\ U \end{bmatrix} = [M_1\, M_2] \begin{bmatrix} R_x x + QU + HV \\ U \end{bmatrix} = \\
= M_1(R_x x + QU + HV) + M_2 U = \\
= M_1 R_x x + M_1 HV + (M_1 Q + M_2)U = 0
\end{aligned}
$$

Here all matrices Y, Q, U, H, V, and R_x are defined using $\rho = \alpha$. For a parity equation that satisfies the model, this equation must hold for all x, all U, and all V, which implies $M_1 R_x = 0$, $M_1 H = 0$, and $M_1 Q + M_2 = 0$. Remember that x is controllable from inputs and disturbances. A parity equation obtained from the ULPE scheme has the form

$$
w^T(Y - QU) = w^T \, [I \quad -Q] \begin{bmatrix} Y \\ U \end{bmatrix} = 0 \tag{9}
$$

where w is constrained by $w^T \, [R_x \quad H] = 0$.

We are to show that for any choice of M in (8), there exists a w such that Equation (9) becomes identical with Equation (8). It is obvious that this is the case if and only if

$$
w^T \, [I \quad -Q] = M \tag{10}
$$

Now choose w as $w^T = M_1$, which is clearly a possible choice since we know that $M_1 \, [R_x \quad H] = 0$. This together with the fact $M_2 = -M_1 Q = -w^T Q$, implies that (10) is fulfilled. All M-vectors, and therefore all parity equations satisfying (8) can therefore be obtained from the ULPE scheme. $\square$

Remarks

The ULPE scheme implies that all possible parity equations are parameterized as follows. Let $N_{R_x H}$ denote a matrix of dimension $(\rho+1) \times \eta$, and its columns are a basis for the η-dimensional left null-space of the matrix $[R_x \; H]$. Then all parity functions up to order ρ can be obtained by in (4) selecting w as $w = N_{R_x H}\gamma$, where γ is a column vector of dimension η. Thus γ is a complete parameterization of all perfectly decoupling parity functions up to order ρ.

The vector w lies in the left null-space of $[R_x \, H]$. Let the columns of a matrix $N_{R_x H}$ be a basis for this null space. If the null-space has dimension η, then there exists η linearly independent vectors $w_1, \ldots w_\eta$, which fulfills (6). Then η different parity functions $h_1, \ldots h_\eta$, can be formed in accordance with (4). The vector $[h_1(t) \ldots h_\eta(t)]$ is called *generalized parity vector*, which is zero in the fault free case. The generalized parity vector will lie in the η dimensional *generalized parity space*, spanned by the rows of $N_{R_x H}$. In other words, the column-space of $N_{R_x H}$ is the left null-space of $[R_x \, H]$, and the row-space is the generalized parity space. This is in accordance with the definitions made by Chow and Willsky (1984), which is a generalization of the *parity vector* and *parity space* originally defined by Potter and Suman (1977).

2.3 Previous Chow-Willsky-like Schemes are not Universal

Following is an example showing that if the system has dynamics controllable only from the fault, none of the previous Chow-Willsky-like schemes can generate all possible parity equations.

Example 1
Consider a system described by the transfer functions

$$\begin{aligned} y_1 &= \frac{1}{s-1}u + \frac{1}{s+1}f \\ y_2 &= \frac{1}{s-1}u + \frac{s+3}{s+1}f \end{aligned}$$

and the realization

$$\begin{aligned} \dot{\phi} &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\phi + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u + \begin{bmatrix} 0 \\ 1 \end{bmatrix}f \\ y &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\phi + \begin{bmatrix} 0 \\ 1 \end{bmatrix}f \end{aligned}$$

Also consider the function

$$h = (1 - s + s^2)y_1 - s^2 y_2 + u \tag{11}$$

If y_1 and y_2 in (11) are substituted with their transfer functions we get

$$\begin{aligned} h &= \frac{1}{s-1}\big((1-s+s^2) - s^2 + (s-1)\big)u + \\ &\quad + \frac{1}{s+1}\big((1-s+s^2) - s^2(s+3)\big)f = \\ &= \frac{-s^3 - 2s^2 - s + 1}{s+1}f \end{aligned}$$

We see that h is zero in the fault free case and becomes non-zero when the fault occurs. Therefore the function (11) is, according to Definition 2, a parity function. With the matrices used in Equation (3), the parity function (11) can be written as

$$h = \begin{bmatrix} 1 & 0 & -1 & 0 & 1 & -1 \end{bmatrix}(Y - QU) = w^T(Y - QU)$$

in which w is uniquely defined. With the realization above, the matrix R is

$$R = [R_x \; R_z] = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & -1 \\ 1 & -2 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$$

The first column of R, i.e. R_x, is orthogonal to w but not the second. This means that the parity function (11) can not be obtained from any of the previous Chow-Willsky-like schemes. Therefore they are not universal. However in the ULPE scheme, the parity function (11) can be obtained because the requirement that w must be orthogonal to the second column of R, is relaxed.

3 Forming the Residual Generator

In this section, the relation between a parity function and a general linear residual generator is discussed. First a residual generator is defined:

Definition 3 *[Residual Generator]*. *A residual generator is a system that takes process inputs and outputs as inputs and generates a signal called* residual, *which is equal to zero when no monitored faults occur and becomes non zero when a monitored fault occurs.*

Many design methods for linear residual generators exists. All result in a filter for which the computational form, i.e. the residual expressed in y_i:s and u_i:s, can be expressed as

$$r = \frac{A_1(\sigma)y_1 + \ldots + A_m(\sigma)y_m + B_1(\sigma)u_1 + \ldots + B_k(\sigma)u_k}{C(\sigma)} \tag{12}$$

where $A_i(\sigma)$, $B_i(\sigma)$, and $C(\sigma)$ are *polynomials* in σ. This includes for example the case when the residual generator is based on observers formulated in state space. According to Definition 3, the objective of residual generation is to create a signal that is affected by monitored faults but not by any other signals. This is equivalent to finding a filter which fulfills the following two requirements: the transfer functions from the monitored faults to the residual must be non-zero, and the transfer functions from all other signals to the residual must be zero, i.e. *decoupling*. These two requirements introduces a constraint on the numerator polynomial of (12) only. The constraint equals the definition of parity function and therefore the numerator polynomial must always be a parity function. There are no constraints on the denominator polynomial $C(\sigma)$ which therefore can be chosen freely.

We will now illustrate how a residual generator can be formed from the parity function (4). For the discrete case, the resulting parity function designed with a Chow-Willsky-like scheme is

$$h = A_1(q)y_1 + \ldots + A_m(q)y_m + B_1(q)u_1 + \ldots + B_k(q)y_k$$

This expression can not be implemented as it is because it is a non-causal transfer function. A common method to obtain a casual transfer function is to introduce $\rho - 1$ units delay. Then the transfer function from system outputs and inputs becomes

$$G_r(q) = [\frac{A'_1(q)}{q^{\rho-1}} \ldots \frac{A'_m(q)}{q^{\rho-1}}\, \frac{B'_1(q)}{q^{\rho-1}} \ldots \frac{B'_k(q)}{q^{\rho-1}}]$$

This is a FIR-filter (or dead-beat observer) with its poles in the origin. However, there is no reason to constrain the poles to the origin only because a Chow-Willsky-like scheme is used when designing the residual generator. Instead, the poles can be placed arbitrarily within the unit circle to obtain stability. Often there is a need for LP-filtering so these poles can be made to function like such a filter. In contrast to observers used in control theory, there is no reason to follow the rule of thumb that says that observer poles should be faster than process poles. If $C(q)$ is the resulting denominator polynomial, the transfer function becomes

$$G_r(q) = [\frac{A_1'(q)}{C(q)} \dots \frac{A_m'(q)}{C(q)} \frac{B_1'(q)}{C(q)} \dots \frac{B_k'(q)}{C(q)}]$$

To get a causal filter, the degree of $C(q)$ must be greater or equal to the maximum degree of the polynomials $A_i(q)$ and $B_i(q)$.

For the continuous case, the resulting parity function designed with a Chow-Willsky-like scheme is

$$h = A_1(s)y_1 + \ldots + A_m(s)y_m + B_1(s)u_1 + \ldots + B_k(s)y_k$$

In general this expression can not be used as a residual generator because the difficulty to measure the derivative of signals. Therefore, poles must be added, but as for the discrete case, these poles can naturally work as for example an LP-filter. The resulting transfer function of the residual generator is

$$G_r(s) = [\frac{A_1(s)}{C(s)} \dots \frac{A_m(s)}{C(s)} \frac{B_1(s)}{C(s)} \dots \frac{B_k(s)}{C(s)}]$$

As seen, there is no need for an explicit *state variable filter*, which is used in (Höfling, 1993) to construct a residual generator from the continuous parity function.

Note the relation to diagnostic observer design, e.g. eigenstructure or the unknown input observer, in which poles also are placed arbitrarily.

Now we know from Theorem 1 that all parity functions can be obtained with the ULPE. Also we know that for any linear residual generator, the numerator polynomial is a parity function and the denominator polynomial can be chosen freely. Therefore the following result is obtained:

Corollary 1. *When discrete or continuous linear systems are considered, the ULPE is a universal residual generator design method for achieving perfect decoupling.*

4 Detectability Analysis

In this section it is investigated whether it is possible to construct a residual generator with given decoupling properties, for the system (1). If this is the case, we say that the fault that is to be monitored, is *detectable*. The analysis of detectability is here approached in the context of parity equations and the ULPE scheme. Criterions for fault detectability has been studied also in other contexts: unknown input observer (Wünnenberg, 1990), detection filter (White and Speyer, 1987), frequency domain (Frank and Ding, 1994), and statistical approach (Basseville and Nikiforov, 1993). However fault detectability has, to the author's knowledge, not been studied in the context of parity equations.

In (Chen and Patton, 1994), *fault detectability* and *strong fault detectability* for a given residual generator, are defined as follows:

Definition 4 ***[Fault Detectability].*** *A fault f is detectable in residual r if the transfer function from the fault to the residual $G_{rf}(\sigma)$ is nonzero:*

$$G_{rf}(\sigma) \neq 0$$

Definition 5 ***[Strong Fault Detectability].*** *A fault f is strongly detectable in residual r if*

$$G_{rf}(0) \neq 0 \qquad \text{(continuous case)}$$

$$G_{rf}(1) \neq 0 \qquad \text{(discrete case)}$$

If a fault is detectable but not strongly detectable, the term *weak detectability* will be used.

4.1 Detectability as a System Property

As will be shown in Theorem 2 and 3, detectability is a system property in the sense that it is the system that limits the possibilities of constructing a residual that is fault detectable and strongly fault detectable respectively. This leads to the following redefinitions of fault detectability and strong fault detectability:

Definition 6 ***[Fault Detectability].*** *A fault is detectable in a system if and only if there exists a residual in which the fault is detectable according to Definition 4.*

Definition 7 ***[Strong Fault Detectability].*** *A fault is strongly detectable in a system if and only if there exists a residual in which the fault is strongly detectable according to Definition 5.*

Next are two theorems to be used for the analysis of fault detectability and strong fault detectability. In the following, the notation $(\dots)_{\rho=n}$ is used to denote that the condition within the parenthesis considers matrices and vectors Y, R_x, R_z, Q, U, H, V, P, and F with $\rho = n$ according to Equation 3. The notation N_X is used to denote a basis for the left null-space of the matrix X.

Theorem 2. *A fault is detectable if and only if*

$$\left(N_{R_xH}^T P \neq 0\right)_{\rho=n}$$

where N_{R_xH} is a basis for the left null-space of $[R_x\ H]$.

For the proof of this theorem we need the following two lemmas:

Lemma 1. *Consider the system (1) with given properties and the corresponding matrices. For all $\rho \geq n$, it holds that $N_{R_xH}^T P = 0$ implies $N_{R_xH}^T R_z = 0$.*

Lemma 2. *If $\left(N_{RH}^T P = 0\right)_{\rho=n}$, then $\forall \rho \geq n\ \ \{N_{RH}^T P = 0\}$.*

The proofs of Lemma 2 and Lemma 1 are given in the appendix. Following is the proof of Theorem 2:

Proof. A parity function, and also a residual, derived with the ULPE scheme is according to Equation 5 sensitive to a fault, that is the fault is detectable, if and only if

$$\exists f(t)\ \{w^T(PF + R_z z) \neq 0\} \tag{13}$$

where $f(t)$ is a fault signal. This condition is equivalent to Definition 4. Since z is controllable from $f(t)$, it holds that at any time point t_1, $z(t_1)$ and $F(t_1)$ can take arbitrary values independently from each other. This means that (13) is equivalent to

$$\exists f(t)\ \{w^T PF \neq 0 \lor w^T R_z z \neq 0\} \tag{14}$$

Such a parity function exists if and only if

$$\exists \rho, w\ \{w^T [R_x\, H] = 0\ \land\ w^T [P\ R_z] \neq 0\}$$

which is equivalent to

$$\exists \rho\ \{N_{R_xH}^T [P\ R_z] \neq 0\} \tag{15}$$

This condition holds if and only if

$$\exists \rho \geq n\ \ \{N_{R_xH}^T [P\ R_z] \neq 0\} \tag{16}$$

because if the ρ in (15) is $\geq n$, then (16) follows directly and if the ρ in (15) is $< n$, then it is always possible to find a larger ρ because the extra terms that appear in (4) and (5) can be canceled by zeros in w.

From Lemma 1, we know that if $N_{R_x H}^T P = 0$ then also $N_{R_x H}^T R_z = 0$. This means that the R_z in condition (16) can be neglected, which results in

$$\exists \rho \geq n \ \ \{N_{R_x H}^T P \neq 0\}$$

According to Lemma 2 it is sufficient to investigate the case $\rho = n$, that is

$$\left(N_{R_x H}^T P \neq 0\right)_{\rho=n} \tag{17}$$

Now since we know from Corollary 1 that the ULPE scheme is universal, (17) is a necessary and sufficient condition for fault detectability. □

The next theorem deals with strong detectability. To the author's knowledge, a general criterion for strong detectability has not been presented elsewhere. The criterion presented here answers the question if there exists a residual generator in which the fault becomes strongly detectable. In (Chen and Patton, 1994), this is reported to be an unsolved research problem.

Strong detectability deals with the stationary residual response when a constant fault is present. A constant fault can be written $f(t) \equiv c$ where c is the constant level of the fault. By studying the definitions of $F(t)$, in Equation (3), for the discrete and continuous case respectively, it is seen that $F(t) \equiv vc$ where $v = [1 \ldots 1]^T$ in the case of a discrete system and $v = [1\,0 \ldots 0]^T$ in the case of a continuous system.

Theorem 3. *A fault is strongly detectable if and only if*

$$\left(N_{R_x H}^T (Pv - R_z A_z^{-1} K_z) \neq 0\right)_{\rho=n} \qquad \textit{(continuous case)}$$

$$\left(N_{R_x H}^T (Pv + R_z (I - A_z)^{-1} K_z) \neq 0\right)_{\rho=n} \qquad \textit{(discrete case)}$$

where $N_{R_x H}$ is a basis for the left null space of $[R_x\ H]$ and $v = [1\,0 \ldots 0]^T$ in the continuous case and $v = [1 \ldots 1]^T$ in the discrete.

For the proof of this theorem we need the following two lemmas:

Lemma 3. *If $\left(N_{RH}^T (Pv - R_z A_z^{-1} K_z) = 0\right)_{\rho=n}$, then $\forall \rho \geq n \quad \{N_{RH}^T (Pv - R_z A_z^{-1} K_z) = 0\}$, where $v = [1\,0 \ldots 0]^T$.*

Lemma 4. *If $\left(N_{RH}^T (Pv + R_z (I - A_z)^{-1} K_z) = 0\right)_{\rho=n}$, then $\forall \rho \geq n \ \ \{N_{RH}^T (Pv + R_z (I - A_z)^{-1} K_z) = 0\}$, where $v = [1\,1 \ldots 1]^T$.*

The proofs of Lemma 3 and Lemma 4 are given in the appendix. Following is the proof of Theorem 3:

Proof. The proof is presented only for the continuous case. The discrete case is treated similarly.

Consider the case when a constant fault is present. We know that the state z will reach steady state because, according to the preconditions described in Section 2.1, the state z is asymptotically stable. This also guarantees that the inverse of A_z exists. If the constant fault is of size c, the stationary value of the parity function becomes

$$w^T(R_z z_{stat} + Pvc) = w^T(-R_z A_z^{-1} K_z + Pv)c \qquad (18)$$

For a residual, also the poles affects the stationary value. However if the residual is derived according to the description in Section 3, the stationary value differs only by a non-zero factor compared to (18).

Now since we know from Corollary 1 that the ULPE scheme is universal, a necessary and sufficient condition for fault detectability is

$$\exists \rho, w \quad \left\{ w^T(Pv - R_z A_z^{-1} K_z) \neq 0 \right\}$$

This is equivalent to

$$\exists \rho \quad \left\{ N_{R_x H}^T (Pv - R_z A_z^{-1} K_z) \neq 0 \right\} \qquad (19)$$

This condition holds if and only if

$$\exists \rho \geq n \quad \left\{ N_{R_x H}^T (Pv - R_z A_z^{-1} K_z) \neq 0 \right\} \qquad (20)$$

because if the ρ in (19) is $\geq n$, then (20) follows directly and if the ρ in (19) is $< n$, then it is always possible to find a larger ρ because the extra terms that appear in (4) and (5) can be canceled by zeros in w.

Now Lemma 3 shows that it is sufficient to consider the case $\rho = n$, that is

$$\left(N_{R_x H}^T (Pv - R_z A_z^{-1} K_z) \neq 0 \right)_{\rho = n}$$

□

Remarks

If one thinks of the P matrix as a description of how the fault propagates through the system, the conditions for fault detectability and strong fault detectability are intuitive. For example the condition for fault detectability says that the fault must not affect the system in the same way as the state or the disturbances.

As seen in Lemma 2, 3, and 4, it is sufficient to chose ρ as $\rho = n$, if fault detectability or strong fault detectability is considered. This means that a residual generator that is able to (strongly) detect a fault, never needs to be designed using a parity function of order larger than n. There may however be other reasons to chose a ρ larger than n.

4.2 Examples

In an inverted pendulum example in (Chen and Patton, 1994), an observer based residual generator was used. It was shown that no residual generator with this specific structure could strongly detect a fault in sensor 1. It was posed as an open question if any residual generator, in which this fault is strongly detectable, exists and in that case how to find it. In the following example, this problem is re-investigated by means of Theorem 3. Also included is a demonstration of Theorem 2.

Example 2
The system description represents a continuous model of an inverted pendulum. It has one input and three outputs:

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1.93 & -1.99 & 0.009 \\ 0 & 36.9 & 6.26 & -0.174 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & -0.3205 & -1.009 \end{bmatrix}^T$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad D = 0_{3\times 1}$$

The faults considered are sensor faults. There are no disturbances and also, there are no states controllable only from faults. This means that there is no R_z matrix or H matrix. For the detectability analysis, we calculate the N_R matrix and form $N_R^T P_1 \neq 0$, $N_R^T P_2 \neq 0$, and $N_R^T P_3 \neq 0$ for the three faults respectively. Then from Theorem 2 it can be concluded that all sensor faults are detectable, i.e. for each sensor fault, it is possible to construct residual generators for which the fault is detectable. To check strong detectability we form the vectors

$$N_R^T P_1 v = [0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0]^T$$

$$N_R^T P_2 v = [-1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0]^T$$

$$N_R^T P_3 v = [* * * * * * * * * * *]^T$$

where $*$ represents nonzero elements. By using Theorem 3 it can be concluded that the second and third sensor faults are strongly detectable, i.e. for each of these faults a residual generator can be found for which the fault is strongly detectable. Also concluded is that the first sensor fault is only weakly detectable, i.e. it is not possible to construct a residual generator in which the fault in sensor 1 is strongly detectable.

As is seen in Equation (5), the fault affects the parity function through both R_z and P. One may note that in the condition of Theorem 2 it is sufficient to consider the matrix P while in Theorem 3 both R_z and P must be considered. The following example shows that this is really the case.

Example 3
The system is continuous and has one structured disturbance and two outputs:

$$A = \begin{bmatrix} -2 & -3 \\ 0 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad E = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \quad K = \begin{bmatrix} -6 \\ -6 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 4 \\ 2 & 4 \end{bmatrix} \quad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad J = \begin{bmatrix} 6 \\ 5 \end{bmatrix} \quad L = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

For this system, $N^T_{R_x H}(Pv - R_z A_z^{-1} K_z) = 0$. This means that the fault is not strongly detectable. However it also holds that $N^T_{R_x H} Pv \neq 0$ which shows that the influence of the fault via R_z must be considered in the condition of strong fault detectability.

---

5 Conclusions

The Universal Linear Parity Equation (ULPE) scheme has been presented. This is an extension to the well known Chow-Willsky scheme. It is shown that none of the previous extensions to the Chow-Willsky scheme are able to generate all parity equations in the case where there are dynamics controllable only from faults. The ULPE scheme is able to handle also this case since it is *universal* in the sense that for any linear, continuous or discrete system, all parity equations can be generated.

It is demonstrated how any perfectly decoupling linear residual generator can be constructed by the help of the ULPE scheme. Therefore the ULPE scheme is also a universal design method for linear residual generation.

Two new conditions for fault detectability and strong fault detectability, formulated in the context of the ULPE scheme, are provided. A general condition for strong fault detectability has not been presented elsewhere.

It is shown that if fault detectability or strong fault detectability are considered, it is sufficient to have $\rho = n$ when designing the parity functions. This means that a parity function, to be used in the design of a residual generator, do not need to have an order larger than n.

6 Acknowledgment

This research was supported by NUTEK (Swedish National Board for Industrial and Technical Development).

References

Basseville, M. and I.V. Nikiforov (1993). *Detection of Abrupt Changes.* PTR Prentice-Hall, Inc.

Chen, J. and R.J. Patton (1994). A re-examination of fault detectability and isolability in linear dynamic systems. Fault Detection, Supervision and Safety for Technical Processes, pp. 567–573. IFAC, Espoo, Finland.

Chow, E.Y. and A.S. Willsky (1984). Analytical redundancy and the design of robust failure detection systems. *IEEE Trans. on Automatic Control*, **29**(7), 603–614.

Frank, P.M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - a survey and some new results. *Automatica*, **26**(3), 459–474.

Frank, P.M. (1993). Advances in observer-based fault diagnosis. Proc. TOOLDIAG'93, pp. 817–836. CERT, Toulouse, France.

Frank, P.M. and X. Ding (1994). Frequency domain approach to optimally robust residual generation and evaluation for model-based fault diagnosis. *Automatica*, **30**(5), 789–804.

Gertler, J. (1991). Analytical redundancy methods in fault detection and isolation; survey and synthesis. IFAC Fault Detection, Supervision and Safety for Technical Processes, pp. 9–21. Baden-Baden, Germany.

Höfling, T. (1993). Detection of parameter variations by continuous-time parity equations. IFAC World Congress, pp. 513–518. Sydney, Australia.

Lou, X.C., A.S. Willsky, and G.C. Verghese (1986). Optimally robust redundancy relations for failure detection in uncertain systems. *Automatica*, **22**(3), 333–344.

Patton, R.J. (1994). Robust model-based fault diagnosis: The state of the art. IFAC Fault Detection, Supervision and Safety for Technical Processes, pp. 1–24. Espoo, Finland.

Potter, J.E. and M.C. Suman (1977). Threshold redundancy management with arrays of skewed instruments. *Integrity Electron. Flight Contr. Syst.*, 15–11 to 15–25.

White, J.E. and J.L. Speyer (1987). Detection filter design: Spectral theory and algorithms. *IEEE Trans. Automatic Control*, **AC-32**(7), 593–603.

Wünnenberg, Jürgen (1990). *Observer-Based Fault Detection in Dynamic Systems.* Ph. D. thesis, University of Duisburg.

7 Appendix

The appendix contains Lemma 5, Lemma 6, Lemma 3, Lemma 4, Lemma 2, and Lemma 1, all with proofs included. If the system (1) does not contain any disturbances, then most proofs in the appendix are simplified and especially Lemma 5 follows easily from Cayley-Hamilton's Theorem. To clarify the relations between all theorems and lemmas, contained in this paper, Figure 2 has been included. The arrows represent implications. Theorems and lemmas given in the main text are represented by boxes with thick lines, and lemmas given in the appendix are represented by boxes with thin lines.

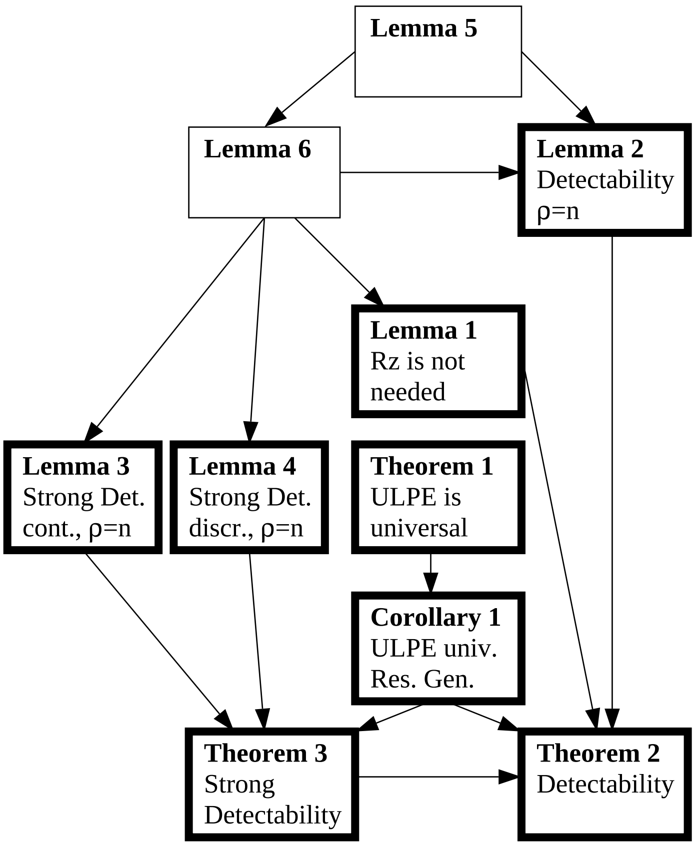

Figure 2: Roadmap of theorems and lemmas.

7.1 Lemma 5

This first lemma is the basis for all other lemmas in the appendix. The matrices C, A, E, J, and K are defined in (1) and (2).

Lemma 5. *If there exists two vectors ψ and $\mathbf{t} = [t_1 \ldots t_{n+1}]^T$ such that*

$$CA^{j-1}\psi + CA^{j-1}Et_1 + \ldots + CEt_j + Jt_{j+1} = CA^{j-1}K$$

for $j = 1 \ldots n$, then for all $\rho \geq n$, there exists a $\mathbf{t}' = \left[t_1 \ t'_2 \ldots t'_{\rho+1}\right]^T$ such that this equation is satisfied for $j = 1 \ldots \rho$.

Proof. Given are the equations

$$\begin{array}{rcl} C\psi + CEt_1 + Jt_2 & = & CK \\ & \vdots & \\ CA^{n-1}\psi + CA^{n-1}Et_1 + \ \ldots \ + CEt_n + Jt_{n+1} & = & CA^{n-1}K \end{array} \tag{21}$$

and the goal is to show that there exists t'_i:s, $i \geq 2$, such that

$$CA^{j-1}\psi + CA^{j-1}Et_1 + CA^{j-2}Et'_2 + \ \ldots \ + CEt'_j + Jt'_{j+1} = CA^{j-1}K \tag{22}$$

for all $j = 1 \ldots \rho$. The equations (21) and (22) specify a condition on the variables t'_i, which are to be found. To be able to carry out the proof we first need to derive a new, more tractable, set of equations which specify an equivalent condition on these variables.

Define the matrices J_1 and J_Λ such that

$$J = [C \quad \Lambda] \begin{bmatrix} J_1 \\ J_\Lambda \end{bmatrix}$$

where Λ has all its columns orthogonal to C. From the equations (21), it is clear that

$$\begin{aligned} Jt_i = CJ_1t_i + \Lambda J_\Lambda t_i = \\ = C(A^{i-2}K - A^{i-2}\psi - A^{i-2}Et_1 - \ \ldots \ - Et_{i-1}) \end{aligned}$$

for $i = 2 \ldots n + 1$. Because of the second equality, it must hold that $\Lambda J_\Lambda t_i = 0$ and therefore $Jt_i = CJ_1t_i$. In the equations (21), Jt_i can now be replaced by CJ_1t_i, which results in

$$\begin{array}{rcl} C\psi + CEt_1 + CJ_1t_2 & = & CK \\ & \vdots & \\ CA^{n-1}\psi + CA^{n-1}Et_1 + \cdots + CEt_n + CJ_1t_{n+1} & = & CA^{n-1}K \end{array}$$

What these equations says is that

$$A^{j-1}\psi + A^{j-1}Et_1 + A^{j-2}Et_2 + \ldots + Et_j + J_1 t_{j+1} - A^{j-1}K$$

for $j = 1 \ldots n$, lies in the right null-space of C. An alternative way of saying this is that there exists g_i:s such that

$$\begin{aligned} Ng_2 + \psi + Et_1 + J_1 t_2 &= K \\ &\vdots \\ Ng_{n+1} + A^{n-1}\psi + A^{n-1}Et_1 + \ldots + Et_n + J_1 t_{n+1} &= A^{n-1}K \end{aligned}$$

where the columns of N are a basis for the right null-space of C. Multiplying the i:th equation from the top with A from the left and then subtracting the $i+1$:th equation results in

$$ANg_i + AJ_1 t_i = Ng_{i+1} + Et_i + J_1 t_{i+1} \tag{23}$$

for $i = 2 \ldots n$. Recall that the goal of the proof is to, for all $\rho \geq n$, find a new set of variables $t'_2 \ldots t'_{\rho+1}$ that fulfills Equation (22). We will do this by finding t'_i:s such that Equation (23) is satisfied for all $i \geq 2$, and then showing that these t'_i:s also satisfies (22).

If $[N \;\; J_1]$ has rank n it is always possible to find t'_{i+1} and g'_{i+1} such that Equation (23) is satisfied also for $i > n$. Otherwise, introduce matrices J_2 and y such that $[N \;\; J_1] = [N \;\; J_2]\, y$, where $[N \;\; J_2]$ has full column rank $\leq n-1$. Now study

$$\left[y \begin{pmatrix} g_2 \\ t_2 \end{pmatrix} \ldots y \begin{pmatrix} g_{n+1} \\ t_{n+1} \end{pmatrix} \right] \tag{24}$$

If the first column in this matrix is 0, then

$$[N \;\; J_2]\, y \begin{pmatrix} g_2 \\ t_2 \end{pmatrix} = Ng_2 + J_1 t_2 = 0$$

Now select a new $g'_2 = 0$ and a new $t'_2 = 0$, and Equation (23) for $i = 2$ becomes

$$0 = Ng_3 + J_1 t_3$$

By continuing selecting new $g'_l = 0$ and $t'_l = 0$ for all $l \geq 2$, then Equation (23) will be satisfied for all $i \geq 2$.

If $y \left[g_2^T \quad t_2^T \right]^T \neq 0$, then from the fact that the matrix (24) has n columns and less than n rows, we know that there exists an $l > 2$ and a vector x such that

$$y \begin{pmatrix} g_l \\ t_l \end{pmatrix} = \left[y \begin{pmatrix} g_2 \\ t_2 \end{pmatrix} \ldots y \begin{pmatrix} g_{l-1} \\ t_{l-1} \end{pmatrix} \right] x$$

Select a new $g'_l = [g_2 \, \ldots \, g_{l-1}] \, x$ and a new $t'_l = [t_2 \, \ldots \, t_{l-1}] \, x$. This choice ensures that Equation (23) for $i = 1 \ldots l$, will be satisfied because the condition

$$y \begin{pmatrix} g'_l \\ t'_l \end{pmatrix} = y \begin{pmatrix} g_l \\ t_l \end{pmatrix}$$

is fulfilled.

Next, select a new $g'_{l+1} = [g_3 \, \ldots \, g_{l-1} \, g'_l] \, x$ and $t'_{l+1} = [t_3 \, \ldots \, t_{l-1} \, t'_l] \, x$. This implies that Equation (23) for $i = l + 1$ is satisfied because

$$\begin{aligned} ANg'_l + AJ_1 t'_l &= AN \, [g_2 \, \ldots \, g_{l-1}] \, x + AJ_1 \, [t_2 \, \ldots \, t_{l-1}] \, x = \\ &= N \, [g_3 \, \ldots \, g_{l-1} \, g'_l] \, x + E \, [t_2 \, \ldots \, t_{l-1}] \, x + J_1 \, [t_3 \, \ldots \, t_{l-1} \, t'_l] \, x = \\ &= Ng'_{l+1} + Et'_l + J_1 t'_{l+1} \end{aligned}$$

The second equality is a consequence of Equation (23). By continuing selecting new g'_{l+2} and t'_{l+2} in the same way and so on, it can be shown that Equation (23) will be satisfied for all $i \geq 2$.

Going back to the original problem, we have now shown that for each ρ there exists a $\mathbf{t}' = \left[t_1 \; t'_2 \ldots t'_{\rho+1}\right]^T$ such that the equation

$$CA^{j-1}\psi + CA^{j-1}Et_1 + CA^{j-2}Et'_2 + \cdots + CEt_j + CJ_1 t'_{j+1} = CA^{j-1}K \quad (25)$$

is satisfied for $j = 1 \ldots \rho$. This equation equals Equation (22) except for the last term of the left side. For all $j \geq 2$, there exists a ϕ such that $t'_j = [t_2 \, \ldots \, t_{n+1}] \, \phi$. Therefore it must hold that $CJ_1 t'_j = Jt'_j$ for all $j \geq 2$. This implies that Equation (25) is equivalent to Equation (22) which ends the proof. □

7.2 Lemma 6

The matrices R_x, R_z, H, and P are defined in Equation (3) and the matrices C, A, E, J, and K are defined in (1) and (2).

Lemma 6. *If* $\left(N_{RH}^T P v = 0\right)_{\rho=n}$, *then* $\forall \rho \geq n \quad \{N_{RH}^T P v = 0\}$, *where* $v = [1\,0 \ldots 0]^T$.

Proof. If $\left(N_{RH}^T P v = 0\right)_{\rho=n}$ then Pv can be written as a linear combination of the columns in R and H. This means that there exists two vectors s and $\mathbf{t} = [t_1 \ldots t_{n+1}]$ such that

$$Pv = Rs + H\mathbf{t}$$

This equation can be written as

$$\begin{aligned} Cs + Jt_1 &= L \\ CAs + CEt_1 + Jt_2 &= CK \\ &\vdots \\ CA^n s + CA^{n-1}Et_1 + \ldots + CEt_n + Jt_{n+1} &= CA^{n-1}K \end{aligned}$$

By defining $\psi = As$ and then applying Lemma 5 to all equations except the first one, it can be concluded that $\forall \rho \geq n \ \{N_{RH}^T P v = 0\}$. $\square$

7.3 Proof of Lemma 2

The following lemma was given without proof in Section 4.1. It is generally applicable to all Chow-Willsky-like schemes. It says that if no parity function, derived with a Chow-Willsky-like scheme with $\rho = n$, can detect the fault, then also no parity functions derived with $\rho > n$ can detect the fault. For use with the ULPE scheme, R is to be changed to R_x. The matrices R_x, R_z, H, and P are defined in Equation (3) and the matrices C, A, E, J, and K are defined in (1) and (2).

Lemma 2. *If* $\left(N_{RH}^T P = 0\right)_{\rho=n}$, *then* $\forall \rho \geq n \ \{N_{RH}^T P = 0\}$.

Proof. If $N_{RH}^T P = 0$ and $\rho = n$, then all columns of P can be written as linear combinations of the columns in R and H. This means that for each column P_i there exists vectors s_i and $\mathbf{t}_i$ such that

$$P_i = Rs_i + H\mathbf{t}_i \qquad i = 1 \dots n+1$$

With $\mathbf{t}_i = [t_{i,1} \dots t_{i,\rho+1}]^T$, these systems of equations can be rewritten as

$$\begin{array}{rcl} Cs_i + Jt_{i,1} &=& b_{i,1} \\ CAs_i + CEt_{i,1} + Jt_{i,2} &=& b_{i,2} \\ &\vdots& \\ CA^n s_i + CA^{n-1}Et_{i,1} + \cdots + CEt_{i,n} + Jt_{i,n+1} &=& b_{i,n+1} \end{array}$$

where

$$b_{i,j} = \begin{cases} 0 & \text{if } j < i \\ L & \text{if } j = i \\ CA^{j-i-1}K & \text{if } j > i \end{cases}$$

and $i = 1 \dots n+1$. If P_i for $i > n+1$ is defined to be the zero vector, i.e. $P_i = 0$, then the system of equations for all $i \geq 1$ are satisfied. Let $\Psi_{i,j}(s_i, \mathbf{t}_i)$ denote the j:th equation in the i:th equation system, i.e.

$$\Psi_{i,j}(s_i, \mathbf{t}_i) \triangleq \quad CA^{j-1}s_i + CA^{j-2}Et_{i,1} + \cdots + CEt_{i,j-1} + Jt_{i,j} = b_{i,j}$$

The goal is to show that when s_i and $\mathbf{t}_i$ satisfies equations $\Psi_{i,j}(s_i, \mathbf{t}_i)$, $\forall i \geq 1$, $\rho = n$, $j = 1 \dots \rho+1$, then for all $\rho \geq n$ there exists $\mathbf{t}'_i$ such that also $\Psi_{i,j}(s_i, \mathbf{t}'_i)$, are satisfied. This is done by induction on i.

Base Step

From Lemma 6, this is already assured for $i = 1$.

Induction Step

The induction hypothesis is that for an arbitrary $\rho \geq n$, the equations $\Psi_{i_f,j}(s_{i_f}, \mathbf{t}'_{i_f})$, $j = 1 \dots \rho + 1$, are satisfied for a certain i_f. Then we want to show that these equations are also satisfied for $i_f + 1$.

To do this we need a new set of equations $\Theta_{i,j}(\bar{s}_i, \bar{\mathbf{t}}_i)$ obtained by subtracting $\Psi_{i-1,j}(s_{i-1}, \mathbf{t}_{i-1})$ from $\Psi_{i,j+1}(s_i, \mathbf{t}_i)$. This is denoted as

$$\Theta_{i,j}(\bar{s}_i, \bar{\mathbf{t}}_i) \simeq \Psi_{i,j+1}(s_i, \mathbf{t}_i) - \Psi_{i-1,j}(s_{i-1}, \mathbf{t}_{i-1}) \tag{26}$$

for all $i > 1$, $\rho = n$, $j = 1, \dots, \rho + 1$, where $\bar{s}_i = As_i - s_{i-1}$ and

$$\bar{\mathbf{t}}_i = [t_{i,1} \quad (t_{i,2} - t_{i-1,1}) \, \dots \, (t_{i,\rho+1} - t_{i-1,\rho})]^T$$

The right side of all these equations is 0, because $b_{i,j+1} = b_{i-1,j}$. By applying Lemma 5 with $\psi = \bar{s}_i$ and $K = 0$, to all equation systems

$$\begin{pmatrix} \Theta_{i,1}(\bar{s}_i, \bar{\mathbf{t}}_i) \\ \vdots \\ \Theta_{i,n}(\bar{s}_i, \bar{\mathbf{t}}_i) \end{pmatrix}$$

where $i > 1$, it can be concluded that for each $\rho \geq n$, there exists a $\bar{\mathbf{t}}'_i$ such that

$$\Theta_{i,j}(\bar{s}_i, \bar{\mathbf{t}}'_i) \tag{27}$$

where $i > 1$, are satisfied for $j = 1 \dots \rho$.

From (26) it follows that

$$\Psi_{i_f+1,j+1}(s_{i_f+1}, \mathbf{t}'_{i_f+1}) \simeq \Theta_{i_f+1,j}(\bar{s}_{i_f+1}, \bar{\mathbf{t}}'_{i_f+1}) + \Psi_{i_f,j}(s_{i_f}, \mathbf{t}'_{i_f})$$

For all $\rho \geq n$ it follows from Equations (27) and the induction hypothesis, that the equation $\Psi_{i_f+1,j+1}(s_{i_f+1}, \mathbf{t}'_{i_f+1})$ is satisfied for $j = 1 \dots \rho$. The first element of $\mathbf{t}_i$ equals the first element of $\mathbf{t}'_i$ which means that the equation $\Psi_{i_f+1,1}(s_{i_f+1}, \mathbf{t}'_{i_f+1})$ equals the equation $\Psi_{i_f+1,1}(s_{i_f+1}, \mathbf{t}_{i_f+1})$, which is satisfied as a consequence of the preconditions of the theorem. Therefore for all $\rho \geq n$, the equation $\Psi_{i_f+1,j}(s_{i_f+1}, \mathbf{t}'_{i_f+1})$ is satisfied for $j = 1 \dots \rho + 1$. This ends the induction and the proof. □

7.4 Proofs of Lemma 3 and 4

The following two lemmas were given without proof in Section 4.1. The first lemma is for the continuous case and the second lemma for the discrete case. Both are generally applicable to all Chow-Willsky-like schemes. They say that if no parity function, derived with a Chow-Willsky-like scheme with $\rho = n$, can strongly detect the fault, then also no parity functions derived with $\rho > n$ can strongly detect the fault. For use with the ULPE scheme, R is to be changed to R_x. The matrices R_x, R_z, H, and P are defined in Equation (3) and the matrices C, A, E, J, and K are defined in (1) and (2).

7.4.1 Continuous Case: Lemma 3

Lemma 3. *If* $\left(N_{RH}^T(Pv - R_z A_z^{-1} K_z) = 0\right)_{\rho=n}$, *then* $\forall \rho \geq n \quad \{N_{RH}^T(Pv - R_z A_z^{-1} K_z) = 0\}$, *where* $v = [1\,0\ldots 0]^T$.

Proof. The proof is based on using Lemma 6. To be able to do so, we define L' as

$$L' = L - C_z A_z^{-1} K_z$$

and define K' as

$$K' = K - A \begin{bmatrix} 0_{n_x \times n_z} \\ I_{n_z} \end{bmatrix} A_z^{-1} K_z$$

Then $Pv - R_z A_z^{-1} K_z$ can be written as

$$Pv - R_z A_z^{-1} K_z = \begin{bmatrix} L' \\ CK' \\ CAK' \\ \vdots \\ CA^{n-1} K' \end{bmatrix} \tag{28}$$

It is seen that the right part of (28) has the same structure as Pv in Lemma 6. Therefore we can use Lemma 6 and conclude that

$$\forall \rho \geq n \;\; \{N_{RH}^T(Pv - R_z A_z^{-1} K_z) = 0\}$$

□

7.4.2 Discrete Case: Lemma 4

Lemma 4. *If* $\left(N_{RH}^T(Pv + R_z(I - A_z)^{-1} K_z) = 0\right)_{\rho=n}$, *then* $\forall \rho \geq n \;\; \{N_{RH}^T(Pv + R_z(I - A_z)^{-1} K_z) = 0\}$, *where* $v = [1\,1\ldots 1]^T$.

Proof. The proof is carried out by induction on ρ.

Base Step

For $\rho = n$, we know from the preconditions of the theorem that $N_{RH}^T(Pv + R_z(I - A_z)^{-1}K_z) = 0$ is already satisfied.

Induction Step

The induction hypothesis is that $N_{RH}^T(Pv + R_z(I - A_z)^{-1}K_z) = 0$ for $\rho = \alpha$. This is equivalent to that $Pv + R_z(I - A_z)^{-1}K_z$ is a linear combination of the columns in R and H. That is there exists vectors s and $\mathbf{t}$ such that the equation

$$Pv - R_z(I - A_z)^{-1}K_z = Rs + H\mathbf{t} \tag{29}$$

is satisfied for $\rho = \alpha$. Then the goal is to show that (29) is also satisfied for $\rho = \alpha + 1$.

By defining

$$\beta = \begin{bmatrix} 0_{n_x \times n_z} \\ I_{n_z} \end{bmatrix} (I - A_z)^{-1}K_z$$

the equation (29) can be rewritten as

$$\begin{aligned} Cs + Jt_1 &= L + C\beta \\ CAs + CEt_1 + Jt_2 &= CK + L + CA\beta \\ &\vdots \\ CA^\alpha s + CA^{\alpha-1}Et_1 + \ldots + CEt_\alpha + Jt_{\alpha+1} &= CA^{\alpha-1}K + \ldots + CK + L + \\ &\quad +CA^{\alpha+1}\beta \end{aligned}$$

For these equations introduce the notation $\Phi_i(s, \mathbf{t})$, denoting the $i+1$:th equation from the top, i.e. $i = 0 \ldots \alpha$. Now calculate a new set of equations $\Theta_i(\bar{s}, \bar{\mathbf{t}})$ by subtracting equation $\Phi_{i-1}(s, \mathbf{t})$ from equation $\Phi_i(s, \mathbf{t})$. This is denoted

$$\Theta_i(\bar{s}, \bar{\mathbf{t}}) \simeq \Phi_i(s, \mathbf{t}) - \Phi_{i-1}(s, \mathbf{t})$$

for $i = 1 \ldots \alpha$, where $\bar{s} = As - s$ and $\bar{\mathbf{t}} = [t_1 \;\; (t_2 - t_1) \; \ldots \; (t_{\alpha+1} - t_\alpha)]^T$. This means that the equations $\Theta_1(\bar{s}, \bar{\mathbf{t}})$ to $\Theta_\alpha(\bar{s}, \bar{\mathbf{t}})$ are

$$\begin{aligned} C\bar{s} + CEt_1 + J\bar{t}_2 &= C\big(K + C(A - I)\beta\big) \\ &\vdots \\ CA^{\alpha-1}\bar{s} + CA^{\alpha-1}Et_1 + CA^{\alpha-2}E\bar{t}_2 + \ldots & \\ \ldots + CE\bar{t}_\alpha + J\bar{t}_{\alpha+1} &= CA^{\alpha-1}\big(K + C(A - I)\beta\big) \end{aligned}$$

By using Lemma 5, with $\psi = \bar{s}$ and $K' = (K + C(A - I)\beta)$, it follows that for $\rho = \alpha+1$, there exists a $\bar{\mathbf{t}}' = \big[t_1 \;\; \bar{t}'_2 \ldots \bar{t}'_{\alpha+2}\big]^T$ such that the equations $\Theta_i(\bar{s}, \bar{\mathbf{t}}')$, $i = 1 \ldots \alpha + 1$ are fulfilled.

The original form of equations can now be obtained by calculating

$$\Phi_i(s,\bar{\mathbf{t}}') \simeq \Theta_i(\bar{s},\bar{\mathbf{t}}') + \Phi_{i-1}(s,\mathbf{t})$$

for $i = 1 \ldots \alpha + 1$. Together with $\Phi_0(s,\mathbf{t})$ we have then shown that $\Phi_i(s,\mathbf{t}')$ is fulfilled for $i = 0 \ldots \alpha + 1$. This is the same as saying that equations (29) are also satisfied for $\rho = \alpha + 1$. This ends the induction and the proof. □

7.5 Proof of Lemma 1

The following lemma was given without proof in Section 4.1. The lemma can be interpreted, by means of Equation (3), as follows. If the fault vector F, acting through the matrix P, cannot make the parity function become non-zero, then neither can the state z, acting through R_z. The matrices R_x, R_z, H, and P are defined in Equation (3) and the matrices C, A, E, J, and K are defined in (1) and (2).

Lemma 1. *Consider the system (1) with given properties and the corresponding matrices. For all $\rho \geq n$, it holds that $N_{R_xH}^T P = 0$ implies $N_{R_xH}^T R_z = 0$.*

Proof. If $N_{R_xH}^T P = 0$ then also $N_{R_xH}^T Pv = 0$ where $v = [1\,0 \ldots 0]^T$. If it holds that $N_{R_xH}^T Pv = 0$ for a $\rho = \alpha \geq n$, then Pv can be written as a linear combination of the columns in R_x and H. This means that there exists two vectors s_x and $\mathbf{t} = [t_1 \ldots t_{n+1}]$ such that

$$Pv = R_x s_x + H\mathbf{t}$$

where $\rho = \alpha$. This equation also holds for $\rho = n$ so by defining $s = [s_x^T \quad 0_{1\times n_z}]^T$ and using Lemma 6, the equation can be extended with $(\alpha - n + n_z)m$ number of rows. The equation obtained is denoted as

$$P'v' = R_x' s_x + H'\mathbf{t}' \tag{30}$$

where P', v', R_x', and H' are the matrices we get if a $\rho = \alpha + n_z$ is used instead of $\rho = n$. By studying the definitions of $P'v'$, R_x', and H', and rearranging Equation (30), we can obtain the equation

$$R_z\mathcal{C} = R_x\mathcal{T}_1 + H\mathcal{T}_2 \tag{31}$$

where $\mathcal{C} = \begin{bmatrix} K_z & A_zK_z \ldots A_z^{n_z-1}K_z \end{bmatrix}$, i.e. the controllability matrix of the state z,

$$\mathcal{T}_1 = [\tau_1 \ldots \tau_{n_z}]$$

$$\tau_i = A_x^i s_x + \begin{bmatrix} A_x^{i-1}E_x & \ldots & E_x \end{bmatrix} \begin{bmatrix} t_1' \\ \vdots \\ t_i' \end{bmatrix} - [I_{n_x} \quad 0]\, A^{i-1}K$$

and

$$\mathcal{T}_2 = \begin{bmatrix} t'_2 & \cdots & t'_{n_z+1} \\ \vdots & & \vdots \\ t'_{\alpha+2} & \cdots & t'_{\alpha+1+n_z} \end{bmatrix}$$

Since we know z is controllable from the single fault, $\mathcal{C}$ is invertible. Therefore, for any w in the left null space of $[R_x\, H]$ it holds that

$$w^T R_z = w^T R_x \mathcal{T}_1 \mathcal{C}^{-1} + w^T H \mathcal{T}_2 \mathcal{C}^{-1} = 0$$

This is equivalent to

$$\left(N^T_{R_x H} R_z = 0\right)_{\rho=\alpha}$$

which ends the proof.

$\square$