

# Optimal statistical fault detection with nuisance parameters

Mitra Fouladirad and Igor Nikiforov

LM2S/UTT 12 rue Marie Curie BP 2060 - 10010 Troyes Cedex - France

fouladirad@utt.fr and nikiforov@utt.fr

## Abstract

Fault detection is addressed within a statistical framework. The corresponding inference problem is stated. Particular emphasis is put on dealing with nuisance parameters. A statistical tool for solving this inference problem is described. An example illustrates the proposed method.

## 1 Introduction

Monitoring complex structures and processes is necessary for security, condition-based maintenance, supervision. Faults can often be modeled as deviations, with respect to a nominal reference value, in the parameter vector of a stochastic system. A crucial issue is to state the significance of the observed changes with respect to noises, uncertainties (nuisance parameters), and changes in the environment of the monitored process [1, 3, 7].

The problem of nuisance parameters rejection is traditionally treated in the framework of the analytical redundancy approach. This approach is based on some natural geometric properties of static (dynamic) systems. The theory of analytical redundancy is purely deterministic. Random noises are only heuristically treated. For this reason the problem of optimal fault detection remains unsolved. On the other hand, there exist an invariant hypothesis testing approach (if the original problem is invariant) or an adaptive testing method (such as the generalized likelihood ratio test). Key features of these statistical methods are their ability to handle noises and uncertainties, to reject nuisance parameters, to decide between two hypotheses  $H_0$  (no faults) and  $H_1$  (there exists a fault).

The goal of this paper is to propose an optimal statistical tool to detect a fault in a linear stochastic (dynamical) system with nuisance parameters. It is supposed that the nuisance parameters are unknown but non random; practically, this means that the nuisance can be intentionally chosen to maximize its negative impact on the monitored system (for instance, to mask a fault).

The paper is organized as follows. We start with the problem statement in section 2. A simple linear model (without nuisance) is treated in section 3. Next, the nuisance parameters are introduced and the problem of optimal fault detection with nuisance is discussed in section 4. Section 5 is devoted to a dynamical model with a deterministic state equation governed by an unknown input. An example of

ground station based GPS integrity monitoring illustrating the relevance of the proposed tools is described in section 6.

## 2 Problem statement

Any detection procedure should perform a tradeoff between two incorrect decisions : false alarm (false rejection of the null hypothesis) and non detection (missed acceptance, or equivalently false rejection, of the alternative hypothesis). Two types of hypotheses have to be distinguished : *simple hypothesis*  $H_i$ ,  $i = 0, 1$ , is defined by a unique value of the parameter vector:  $H_i: \theta = \theta_i$ . A *composite hypothesis* refers to a set of parameters  $H_i: \theta \in \Theta_i$  with  $\Theta_i \subset \mathbb{R}^r$ . We assume that  $\Theta_0 \cap \Theta_1 = \emptyset$ . Composite hypotheses are more relevant than simple ones in practice, because of limited available amount of information, especially for the alternative (fault mode). The quality of a binary statistical test  $\delta$  is defined with the probability of false alarm:  $\alpha = \Pr_0(\delta \neq H_0)$ , where  $\Pr_i$  stands for observations  $Y_1, \dots, Y_k$  being generated by distribution  $P_i$ , and the *power function* :  $\beta_\delta(\theta) = \Pr_\theta(\delta = H_1)$ . In case of a vector parameter  $\theta$ , the crucial issue is to find an optimal solution over a set of alternatives which is rich enough. Unfortunately, uniformly most powerful (UMP) tests scarcely exist, except when parameter  $\theta$  is scalar, the family of distributions has a monotone likelihood ratio, and the test is one-sided [2, 4]. As we have mentioned in the introduction, another important issue is dealing with the nuisance parameters. To solve the composite hypotheses testing problem with nuisance parameters we use the theory developed by Wald in his paper [10]. We also adapt the Wald's theory to the case where the parameter of interest (fault)  $\theta$  and the nuisance  $X$  belong to different subspaces of the observation space. Therefore, the goal of this paper is twofold. First, we develop an optimal statistical tool to solve the problem of fault  $\theta$  detection in the following linear gaussian model :

$$Y = HX + M\theta + \xi, \quad (1)$$

where  $Y \in \mathbb{R}^n$  is the measured output,  $\theta \in \mathbb{R}^r$  is the parameter (fault) of model (1),  $M$  is a full column rank matrix of size  $(n \times r)$  with  $r < n$ ,  $X \in \mathbb{R}^m$  is an unknown and non random state vector (nuisance parameter),  $H$  is a matrix of size  $(n \times m)$  with  $m < n$ , and  $\xi$  is a zero mean gaussian noise  $\xi \sim N(0, \sigma^2 I_n)$  with the known variance  $\sigma^2 > 0$  and the  $(n \times n)$  identity matrix  $I_n$ . Second, the developed tool is used to examine several more practical questions, for in-

stance, how to assess the impact of nuisance dynamics, injected in the model via a state equation, or how to solve the problem of “optimal residual generation”.

### 3 Simple linear model

The goal of this section is to apply the Wald’s theory of binary hypotheses testing to a linear gaussian model, to adapt the optimality criterion to this case and to design the optimal algorithm that realizes this optimality criterion. We mainly use the technique developed by Wald [10] in this section.

#### 3.1 Design of the test and its optimality

Originally, Wald has proposed a solution to the following problem : the observation  $Y \in \mathbb{R}^n$  is generated by one of two gaussian distributions :  $N(0, \Sigma)$  and  $N(\theta \neq 0, \Sigma)$ , where  $\theta$  is the mean vector and  $\Sigma$  is a positive definite covariance matrix. The hypotheses testing problem consists in deciding between

$$H_0 : \{\theta = 0\} \text{ and } H_1 : \{\theta \neq 0\}. \quad (2)$$

The peculiarity of the above problem is that that a uniformly most powerful (UMP) test does not exist in case of a vector parameter  $\theta$ . To overcome this difficulty, Wald propose to impose an additional constraint on the class of considered tests, namely, a *constant power function* over a family of surfaces  $\mathcal{S}$  defined on the parameter space  $\Theta$ , in order to avoid the existence of UMP tests over a subspace  $\bar{\Theta}$  of  $\Theta$  which are very inefficient over  $\Theta \setminus \bar{\Theta}$ .

**Definition 1** A test  $\delta^* \in K_\alpha = \{\delta : \Pr_0(\delta \neq H_0) \leq \alpha\}$ , where  $\alpha$  is the prescribed probability of false alarm, is said to have uniformly best constant power (UBCP) on the family of surfaces  $\mathcal{S}$ , if the following conditions are fulfilled [10] :

1. for any pair of points  $\theta_1$  and  $\theta_2$  which lies on the same surface  $S_c \in \mathcal{S}$ ,  $\beta_{\delta^*}(\theta_1) = \beta_{\delta^*}(\theta_2)$ , where  $\beta_\delta(\theta) = \Pr_\theta(\delta = H_1)$  is the power function of the test  $\delta$ .
2. for another test  $\delta \in K_\alpha$ , which satisfies the previous condition, we have  $\beta_{\delta^*}(\theta) \geq \beta_\delta(\theta)$ .

To solve the hypotheses testing problem (2), Wald defines the following family of surfaces (ellipsoids) :

$$\mathcal{S} = \{S_c : \theta^T \Sigma^{-1} \theta = c^2\}. \quad (3)$$

The UBCP test defined on the family of surfaces (3) is given by

$$\delta^*(Y) = \begin{cases} H_0 & \text{if } \Lambda(Y) = Y^T \Sigma^{-1} Y < h(\alpha) \\ H_1 & \text{if } \Lambda(Y) = Y^T \Sigma^{-1} Y \geq h(\alpha) \end{cases}, \quad (4)$$

where the threshold  $h = h(\alpha)$  is chosen to satisfy the definition of the class  $K_\alpha$ . Let us assume the following gaussian linear model :

$$Y = M\theta + \xi, \quad (5)$$

where  $Y \in \mathbb{R}^n$  is the observation vector,  $\theta \in \mathbb{R}^r$  is the parameter of model (5), i.e. the vector of faults,  $M$  is a full column rank matrix of size  $(n \times r)$  with  $r < n$  and  $\xi$  is a zero mean gaussian noise  $\xi \sim N(0, \sigma^2 I_r)$ ,  $\sigma^2 > 0$ . As in the previous case, the problem consists in deciding between the hypotheses  $H_0 : \{\theta = 0\}$  and  $H_1 : \{\theta \neq 0\}$ . Let us apply the general Wald’s idea to design the UBCP test :

$$\delta^*(Y) = \begin{cases} H_0 & \text{if } \Lambda(Y) = \frac{1}{\sigma^2} Y^T M(M^T M)^{-1} M^T Y < h(\alpha) \\ H_1 & \text{if } \Lambda(Y) = \frac{1}{\sigma^2} Y^T M(M^T M)^{-1} M^T Y \geq h(\alpha) \end{cases}. \quad (6)$$

It is necessary to prove that the test  $\delta^*$ , given by equation (6), is UBCP over the following family of ellipsoids:

$$S = \left\{ S_c : \frac{1}{\sigma^2} \theta^T M^T M \theta = c^2 \right\}. \quad (7)$$

The main result of this section is established in the following Theorem 1.

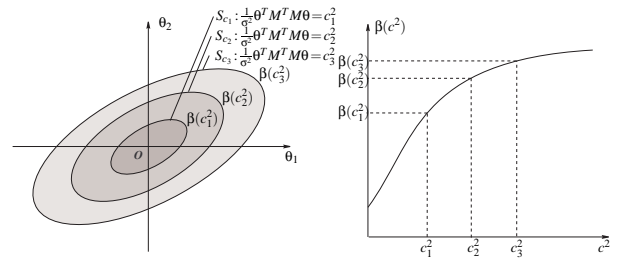
**Theorem 1** Let us consider the regression model (5). The test  $\delta^*(Y) \in K_\alpha$ , given by equation (6), is UBCP for deciding between the hypotheses  $H_0 : \{\theta = 0\}$  and  $H_1 : \{\theta \neq 0\}$  over the family of ellipsoids (7).

#### 3.2 Discussion of the UBCP test

Let us discuss now some issues in deciding between  $H_0 : \{\theta = 0\}$  and  $H_1 : \{\theta \neq 0\}$ . First of all, the test  $\delta^*(Y)$  coincides with the generalized likelihood ratio (GLR) test. After a simple algebra we obtain the following expression for the GLR

$$\hat{\Lambda}(Y) = 2 \log \frac{\sup_{\theta} f_{\theta}(Y)}{f_0(Y)} = \frac{1}{\sigma^2} Y^T M(M^T M)^{-1} M^T Y = \Lambda(Y),$$

where  $\Lambda(Y)$  is given by equation (6). A typical family of



**Figure 1:** Family of ellipsoids and the power function  $\beta(c^2)$ .

ellipsoids and power function  $\beta_{\delta^*}$  of the test  $\delta^*$  are depicted in figure 1. Because the power is constant on the same surface  $S_c : \frac{1}{\sigma^2} \theta^T M^T M \theta = c^2$ , it is reasonable to present the power as a function of  $c^2 : c^2 \mapsto \beta_{\delta^*}(c^2)$ . The statistics  $\Lambda(Y) = \frac{1}{\sigma^2} Y^T M(M^T M)^{-1} M^T Y$  is distributed according to the  $\chi^2$  law with  $m$  degrees of freedom. The law  $\chi^2$  is central under  $H_0$  and noncentral under  $H_1$ . It is easy to see that  $c^2$  is the parameter of non centrality of the  $\chi^2$  law under  $H_1$ . Hence, the power function is given by

$$\beta_{\delta^*}(c^2) = \Pr_{c^2}(\delta^* = H_1) = \int_h^\infty f_{c^2}(y) dy, \quad (8)$$

where  $f_\lambda(y) = f_0(y)e^{-\frac{\lambda}{2}}G\left(\frac{r}{2}, \frac{\lambda y}{4}\right)$  is the density of the non central  $\chi_{r,\lambda}^2$  with  $r$  degrees of freedom,  $f_0(y) = \frac{y^{\frac{r}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{r}{2}}\Gamma(\frac{r}{2})}$ ,  $\lambda$  is the noncentrality parameter,  $\Gamma(x) = \int_0^\infty u^{x-1}e^{-u}du$  is the gamma function, and  $G(\kappa, x) = 1 + \frac{x}{\kappa} + \frac{x^2}{\kappa(\kappa+1)2!} + \dots + \frac{x^p}{\kappa(\kappa+1)\dots(\kappa+p-1)p!} + \dots$  is the hypergeometric function. The threshold  $h = h(\alpha)$  is defined by solving the following equation :

$$\alpha = \Pr_0(\delta^* \neq H_0) = \int_h^\infty f_0(y)dy, \quad (9)$$

where  $\alpha$  is the prescribed probability of false alarm.

## 4 Linear model with nuisance parameters

The goal of this section is to apply the Wald's theory of binary hypotheses testing to a linear gaussian model with a nuisance parameter. As in the previous section, the Wald's idea to define a family of surfaces around the point  $\theta = 0$  on the parameter space will be extensively used now to adapt the optimality criterion to this case and to design a decision rule that realizes the optimality criterion.

### 4.1 Nuisance parameters rejection

Let us recall the regression model (1) with the nuisance parameter  $X$ . The new hypotheses testing problem consists in deciding between

$$H_0 : \{\theta = 0, X \in \mathbb{R}^m\} \text{ and } H_1 : \{\theta \neq 0, X \in \mathbb{R}^m\}, \quad (10)$$

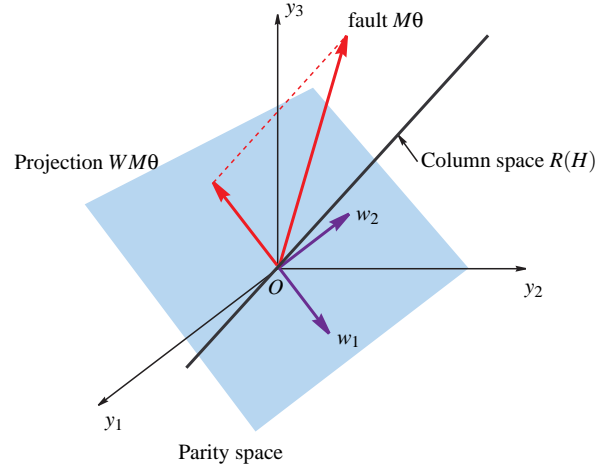
where  $\theta$  is the informative parameter and  $X$  is the nuisance parameter, while considering  $X$  as an *unknown* vector. Because the nuisance parameter  $X$  is non random and its current values are not bounded ( $X \in \mathbb{R}^m$ ), the only solution is to eliminate the impact of  $X$  on the decision function  $\Lambda$ .

First of all, let us note that the family of distributions  $Y \sim N(HX + M\theta, \sigma^2 I_n)$  remains invariant (see [4] for details and definitions) under the group of translations  $G = \{g : g(Y) = Y + HC\}$ ,  $C \in \mathbb{R}^m$ , which in the parameter space induces the group  $\bar{G} = \{\bar{g} : \bar{g}(Y) = Y + HB\}$ ,  $B \in \mathbb{R}^m$ , that preserves both  $\Omega_0 = \{\mathbb{E}(Y) = HX, X \in \mathbb{R}^m\}$  and  $\Omega_1 = \{\mathbb{E}(Y) = HX + M\theta, X \in \mathbb{R}^m, \theta \neq 0\}$ , i.e.  $\bar{g}\Omega_0 = \Omega_0$  and  $\bar{g}\Omega_1 = \Omega_1$ . Hence, the hypotheses testing problem (10) remains invariant under  $G$  [4]. As it is mentioned in [4], the optimal invariant tests are based on the maximal invariants (principle of invariance). Let us recall the notion of *maximal invariant* [4] :

**Definition 2** A function  $M = M(Y)$  is said to be maximal invariant to a group of transformation  $G$ , if the following two conditions are satisfied: 1)  $M(g(Y)) = M(Y)$  for all  $g \in G$ ; 2)  $M(Y_1) = M(Y_2)$  implies  $Y_2 = g(Y_1)$  for some  $g \in G$ .

To apply the principle of invariance, let us define the column space  $R(H)$  of the matrix  $H$ . The standard solution is the projection of  $Y$  on the orthogonal complement  $R(H)^\perp$  of the column space  $R(H)$ . The space  $R(H)^\perp$  is well-known

under the name "parity space" in the analytical redundancy literature [3]. This situation is depicted in figure 2 where



**Figure 2:** Column space and its orthogonal complement (parity space).

$Y \in \mathbb{R}^3, X \in \mathbb{R}$ . If the matrix  $H$  is full column rank, the parity vector  $Z = WY$  is the transformation of the measured output  $Y$  into a set of  $n - m$  linearly independent variables by projection onto the left null space of the matrix  $H$ . The matrix  $W^T = (w_1, \dots, w_{n-m})$  of size  $n \times (n - m)$  is composed of the eigenvectors  $w_1, \dots, w_{n-m}$  of the projection matrix  $P_H = I_n - H(H^T H)^{-1}H^T$  corresponding to eigenvalue 1. The matrix  $W$  satisfies the following conditions :  $WH = 0$ ,  $W^T W = P_H$ ,  $W W^T = I_{n-m}$ . If the rank of  $H$  is equal to  $q : 1 \leq q < m$ , then  $P_H = I_n - H(H^T H)^- H^T$ , where  $A^-$  is a generalized inverse of  $A$ , and the matrix  $W$  is of size  $n \times (n - q)$ . It follows from the first of the above conditions that the transformation by  $W$  completely removes the interference of the nuisance parameter  $X$ . It can be shown that the function  $M(Y) = Z = WY$  is maximal invariant to the group of translations  $G = \{g : g(Y) = Y + HX\}$ . For this reason all invariant tests should depend on  $Y$  only via the vector  $Z = WY$ . The measurement model (1) can be rewritten by the following manner :

$$Z = WY = WM\theta + W\xi = WM\theta + \zeta, \quad (11)$$

where  $\zeta \sim N(0, \sigma^2 I_{n-m})$ ,  $\sigma^2 > 0$ .

### 4.2 Design of the UBCP invariant test

Let us assume that the matrix  $WM$  is full column rank of size  $((n - m) \times r)$  with  $r < n - m$ . Hence, the results of section 3 (Theorem 1) can be directly applied to the model given by equation (11) for deciding between  $H_0 : \{\theta = 0\}$  and  $H_1 : \{\theta \neq 0\}$ . Putting together equations (6) and (11) we get the invariant test :

$$\delta^*(Y) = \begin{cases} H_0 & \text{if } \Lambda(Y) < h(\alpha) \\ H_1 & \text{if } \Lambda(Y) \geq h(\alpha) \end{cases} \quad (12)$$

with  $\Lambda(Y) = \frac{1}{\sigma^2} Y^T P_H M (M^T P_H M)^{-1} M^T P_H Y$  which is UBCP over the following family of surfaces

$$S = \left\{ S_c : \frac{1}{\sigma^2} \theta^T M^T P_H M \theta = c^2 \right\}. \quad (13)$$

To compute the power function and the probability of false alarm, equations (8) - (9) can be used with the following elementary modification : the noncentrality parameter  $\lambda$  is equal now to  $c^2 = \frac{1}{\sigma^2} \theta^T M^T P_H M \theta$ .

### 4.3 Discussion of the UBCP invariant test

Let us discuss some issues in deciding between  $H_0 : \{\theta = 0\}$  and  $H_1 : \{\theta \neq 0\}$  (10) while considering  $X$  as an unknown nuisance parameter. As in the previous case, the invariant test  $\delta^*(Y)$  coincides with the GLR test. It can be shown that the following equality is satisfied

$$\widehat{\Lambda}(Y) = 2 \log \frac{\sup_{\theta, X} f_{\theta, X}(Y)}{f_X(Y)} = \frac{1}{\sigma^2} Y^T (P_H - P_{M, H}) Y = \Lambda(Y),$$

where  $P_{M, H} = I_n - \widetilde{H}(\widetilde{H}^T \widetilde{H})^{-1} \widetilde{H}^T$ ,  $\widetilde{H} = (H M)$  and  $\Lambda(Y)$  is given by equation (12). Hence, in the case of the regression model (1), the GLR is an optimal invariant test.

The statistical properties of the GLR test have been examined in [6]. It has been mentioned in [6] that the GLR test is (UMP) optimal because the  $\chi^2$  distribution of  $\widehat{\Lambda}(Y) = \Lambda(Y)$  is monotone in the noncentrality parameter  $\lambda$ . Therefore, the authors of [6] have reduced a vector parameter detection problem to a scalar one (without any guarantee that this reduction does not change the sense of optimality). We would like to stress that in contrast with [6] we prove the optimality of the test *directly* in the parameter space.

Another interesting question arises in connection with the above mentioned problem of "optimal residual generation" (see, for instance, the survey [7]). The recommendation of the theory of invariant tests is to use the maximal invariant statistics  $Z = WY$  to design an optimal test. Let us consider now a subspace of the parity space (defined by  $W$ ) and estimate the power function of such a test based on such a "subset" parity vector. To obtain this subspace, let us define another nuisance rejection matrix  $W_1$  of size  $(k \times n)$ , where  $k$  is so chosen that  $r < k < n - m$ , from the matrix  $W$  by deleting  $n - m - k$  rows of  $W$ . It is easy to see that  $W_1 H = 0$  and  $W_1 W_1^T = I_k$ . Finally, we obtain

$$Z_1 = W_1 Y = W_1 M \theta + W_1 \xi = W_1 M \theta + \zeta_1, \quad (14)$$

where  $\zeta_1 \sim N(0, \sigma^2 I_k)$ ,  $\sigma^2 > 0$ . Putting together equations (6) and (11) we get the new invariant test :

$$\delta_1^*(Z_1) = \begin{cases} H_0 & \text{if } \Lambda(Z_1) < h(\alpha) \\ H_1 & \text{if } \Lambda(Z_1) \geq h(\alpha) \end{cases} \quad (15)$$

with  $\Lambda(Z_1) = \frac{1}{\sigma^2} Z_1^T W_1 M ((W_1 M)^T W_1 M)^{-1} M^T W_1^T Z_1$  which is UBCP over the family of ellipsoids

$$S = \left\{ S_c : \frac{1}{\sigma^2} \| W_1 M \theta \|_2^2 = c^2 \right\}. \quad (16)$$

**Lemma 1** *Let us assume that  $\Pr_0(\delta_1^* \neq H_0) = \Pr_0(\delta^* \neq H_0) = \alpha$  ( $0 < \alpha < 1$ ). Then the following inequality is satisfied for any  $\theta \neq 0$  :  $\beta_{\delta_1^*}(\theta) \leq \beta_{\delta^*}(\theta)$ .*

The above equation shows that a "subset" parity vector cannot improve the power of the test and the full-set parity space is better (at least not worse) than a subset one. But the choice of the full-set rejection matrix  $W$  of size  $(n - m \times n)$  is not unique. As it is mentioned in the survey [7], the product  $AW$ , where  $A$  is a matrix of size  $(n - m \times n - m)$  such that  $\det A \neq 0$ , leads to another rejection matrix  $\widetilde{W} = AW$ . It is asked in [7] whether such an operation can improve the residual generation step or not. It is clear that  $\widetilde{W}H = 0$ . The following lemma shows that using the rejection matrix  $\widetilde{W} = AW$  does not change the power function of the test.

**Lemma 2** *Let us assume that the UBCP invariant test (12) is designed by using a rejection matrix  $\widetilde{W} = AW$ , where  $\det A \neq 0$ . Then the power function of the test (12) remains independent of  $A$ .*

## 5 Dynamical model with unknown inputs

The goal of this section is to apply the statistical tools developed in sections 3 and 4 to a dynamical model where the nuisance is modeled by a deterministic state equation :

$$X_k = F X_{k-1} + B U_k, \quad (17)$$

$$Y_k = H X_k + M \theta_k + \xi_k, \quad (18)$$

where  $Y_k \in \mathbb{R}^n$  is the measured output,  $X_k \in \mathbb{R}^m$  is the state vector,  $U_k \in \mathbb{R}^p$  is the unknown input vector (nuisance parameter),  $\theta_k \in \mathbb{R}^r$  is the vector of faults at time  $k$ , and  $\xi_k$  is a zero mean gaussian white noise,  $\xi_k \sim N(0, \sigma^2 I_n)$ ,  $\sigma^2 > 0$ . It is assumed that all the matrices of (17) - (18) and the variance  $\sigma^2$  are known. The characteristic feature of the dynamical model (17) - (18) is that the nuisance  $X$  is modeled by a deterministic state equation (17) which is governed by the input vector  $U$ . For instance, the nuisance  $X_k$ , say, the altitude of aircraft, certainly depends on its previous value  $X_{k-1}$ , via the aircraft dynamics of modeled by the matrix  $F$ , and some unknown factors like positions of control surfaces.

A fixed  $N$ -size sample of measured outputs  $Y_1, \dots, Y_N$  is available and supposed to be generated by one of two alternative hypotheses. There are two methods to deal with the nuisance parameters  $X_1, \dots, X_N$ .

*Method 1 : Ignoring the state equation.* Here, the UBCP invariant test is designed by using uniquely the measurement equation (18). Equation (17) is ignored. By putting together  $N$  measurement equations we get :

$$Y_{1,N} = H X_{1,N} + M \theta_{1,N} + \xi_{1,N}. \quad (19)$$

where  $X_{1,N} = (X_1^T, \dots, X_N^T)^T$ ,  $\theta_{1,N} = (\theta_1^T, \dots, \theta_N^T)^T$ ,  $\xi_{1,N} = (\xi_1^T, \dots, \xi_N^T)^T$  and the block diagonal matrices  $\widetilde{M}$  et  $\widetilde{H}$  are composed of  $M$  and  $H$  respectively. The UBCP invariant test (12) is applicable now by replacing the vector  $Y$ ,

matrices  $H$ ,  $M$  and  $P_H$  by  $Y_{1,N}$ ,  $H$ ,  $M$  and  $P_H = I_{kn} - H(H^T H)^{-1} H^T$  respectively. This test  $\delta_H^*$  is UBCP over the following family of ellipsoids

$$S_H = \left\{ S_c : \frac{1}{\sigma^2} \| WM\theta_{1,k} \|_2^2 = c^2 \right\},$$

where the matrix  $W$  satisfies the following conditions  $WH = 0$ ,  $W^T W = P_H$  and  $WW^T = I_{N(n-q)}$ .

*Method 2 : Using both (state-space and measurement) equations.* In contrast with the previous case, the UBCP invariant test is designed now by using both equations (17) and (18). The nuisance  $X_i$  is replaced by  $F^k X_0 + \sum_{j=0}^{i-1} F^j B U_{i-j}$  for  $i \geq 1$ . This implies that  $Y_i = H F^i X_0 + H F^{i-1} B U_1 + \dots + H B U_i$  for  $1 \leq i \leq N$ . By putting together  $N$  measurement equations, we get :

$$Y_{1,N} = U X_0 + V U_{1,N} + M \theta_{1,N} + \xi_{1,N}, \quad (20)$$

where the nuisance is defined by  $X_0$  and  $U_{1,N} = (U_1^T, \dots, U_N^T)^T$  and the matrices  $U$  and  $V$  are functions of the matrices  $H$ ,  $F$  and  $B$ . Let us apply the UBCP invariant test (12) to the model given by equation (20), by replacing the vector  $Y$ , matrices  $H$ ,  $M$  and  $P_H$  by  $Y_{1,N}$ ,  $(U \ V)$ ,  $M$  and  $P_{(U \ V)} = I_{Nn} - (U \ V) ((U \ V)^T (U \ V))^{-1} (U \ V)^T$  respectively. This test  $\delta_{(U \ V)}^*$  is UBCP over the following family of ellipsoids

$$S_{(U \ V)} = \left\{ S_c : \frac{1}{\sigma^2} \| W_1 M \theta_{1,k} \|_2^2 = c^2 \right\},$$

where matrix  $W_1$  satisfies the following conditions  $W_1 (U \ V) = 0$ ,  $W_1 W_1^T = I_{Nn - \text{rank}(U \ V)}$ , and  $W_1^T W_1 = P_{(U \ V)}$ . If the matrices  $H$  and/or  $(U \ V)$  are not full column rank then  $A^{-1}$  should be replaced by  $A^-$  in the above equations.

*Comparison of the above mentioned methods.* Let us consider the above mentioned tests  $\delta_H^*$  and  $\delta_{(U \ V)}^*$ . The result of their comparison is established by the following Lemma.

**Lemma 3** *Let us assume that  $\text{Pr}_0(\delta_H^* \neq H_0) = \text{Pr}_0(\delta_{(U \ V)}^* \neq H_0) = \alpha$  ( $0 < \alpha < 1$ ). Then the following inequality is satisfied for any  $\theta_{1,k} \neq 0$  :  $\beta_{\delta_H^*}(\theta_{1,N}) \leq \beta_{\delta_{(U \ V)}^*}(\theta_{1,N})$ .*

The proof of Lemma 3 is based on the fact that the column space of  $(U \ V)$  (see equation(20)) is contained in the column space of  $H$  (see equation(19)). This leads to the orthogonal complement which can be richer in the case of using both (state and measurement) equations. The importance of state equation to model the nuisance has been heuristically discussed in the analytical redundancy literature. Lemma 3 formally proves this result. The UBCP invariant test based on the state-space model (17) - (18) performs better (at least not worse) than the test based uniquely on the measurement equation (18).

## 6 Application : ground station based GPS integrity monitoring

Integrity monitoring requires that a navigation system detects faulty measurement sources, and if possible isolates and removes them from the navigation solution before they sufficiently contaminate the output. The recent researches show that the detection of the GPS navigation message degradation is crucially important for many transportation systems [9].

The ground station based GPS integrity is considered (see figure 3). For some safety critical navigation modes, land-

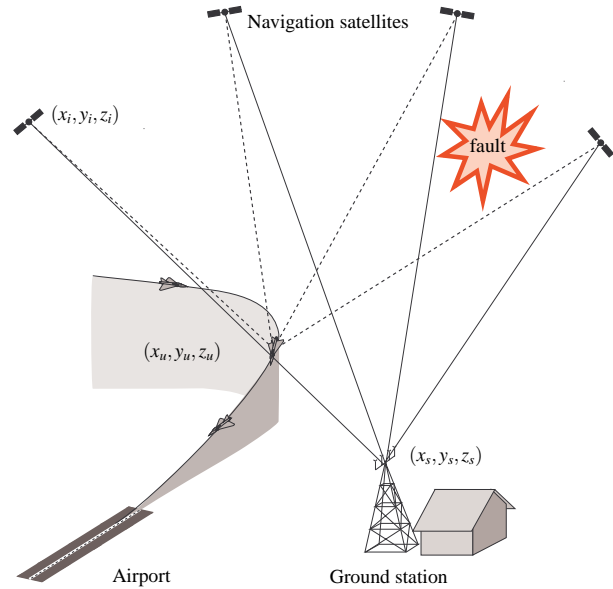


Figure 3: Ground station integrity monitoring.

ing, for instance, the GPS channels integrity monitoring is realised by using the measurements of monitoring station at a known position  $X_s = (x_s, y_s, z_s)^T$  close to the airport (see figure 3). When a fault is detected, the corresponding information is transmitted via the integrity channel. The GPS solution is based upon accurate measuring the distance (*range*) from several satellites with known locations to a user (station or vehicle). Let us assume that there are  $n$  satellites located in three-space at the known positions  $X_i = (x_i, y_i, z_i)^T, i = 1, \dots, n$ . The *pseudorange*  $r_i$  from the  $i$ -th satellite to the user can be written as

$$y_i = r_i - d_i = c \rho + \xi_i \quad (21)$$

where :  $d_i = \|X_i - X_s\|_2 = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2}$  is the true distance from the  $i$ -th satellite to the station,  $\rho$  is an unknown (non random) station clock bias  $\rho \in \mathbb{R}^1$ ,

<sup>1</sup>It is assumed that the ground station is equipped by a lower accuracy clock. To explain why we assume this model of nuisance ( $\rho \in \mathbb{R}$ ) let us recall that a clock error of  $10^{-3}$  s leads to an additional pseudorange bias of 300km ! In fact, nominal pseudorange standard deviation is approximately equal to 12.5 m now and will be limited by 6.6 m or even by 3.8 m in the near future. Roughly speaking, for  $\rho$  the space  $\mathbb{R}$  is "reduced" to the interval  $[-10^{-3}; +10^{-3}]$  s.

$c \simeq 2.9979 \cdot 10^8$  m/s is the speed of light and  $\xi_i$  is an additive pseudorange error,  $\xi = (\xi_1, \dots, \xi_n)^T$  is the vector of additive pseudorange errors at the ground station position. A fault is modeled by the vector  $\theta$  of additional pseudorange biases. Hence, we get the following measurement model with a fault :

$$Y = C\rho + \xi (+\theta), \quad \xi \sim N(0, \sigma^2 I_n), \quad Y = (y_1, \dots, y_n)^T, \quad (22)$$

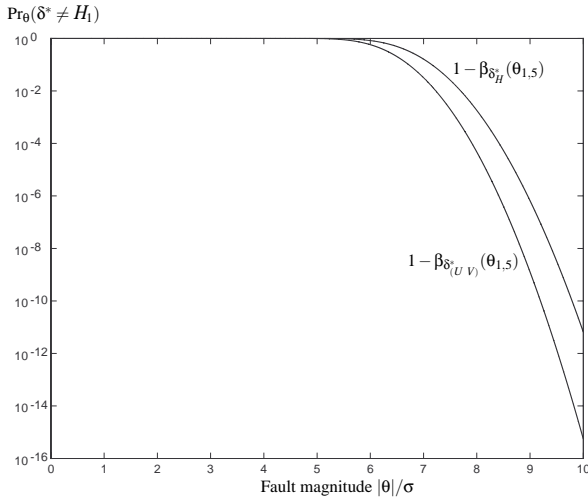
where  $C = (c, \dots, c)^T$ . The problem consists in deciding between the null hypothesis  $H_0$  (no contaminated pseudoranges) and the alternative hypothesis  $H_1$  (there are contaminated pseudoranges)

$$H_0 : \theta = 0, \rho \in \mathbb{R} \text{ against } H_1 : \theta \neq 0, \rho \in \mathbb{R}$$

As it follows from section 4, the UBCP invariant test is based on the following statistics (see details in [5])

$$\chi^2 = \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \bar{y})^2 \quad (23)$$

where  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ . To illustrate Lemma 3 and to show



**Figure 4:** Probability of non detection by ignoring the state equation and by using both (state and measurement) equations.

how the use of state equation improves the performance of the test, let us consider that the station clock bias is modeled by the following state equation :  $\rho_k = \rho_{k-1} + 0.5\rho_{k-2} + u_k$ . Let us additionally assume that :  $n = 4$  and  $N = 5$ . The comparison between two methods presented in section 5 (by ignoring and by using the state equation) is shown in Figure 4. The probability of false alarm is chosen to be  $\alpha = 10^{-3}$ . It is easy to see that the probability of non detection  $1 - \beta_{\delta_{(U V)}}^*(\theta_{1,5})$  by using the state equation is much smaller than the probability of non detection  $1 - \beta_{\delta_H}^*(\theta_{1,5})$  obtained by ignoring the state equation. Let us finish this example with a comment on the fault detectability problem. We consider a user (aircraft) at the positions  $X_u =$

$(x_u, y_u, z_u)^T$ . By linearizing the pseudorange equation with respect to the state vector  $X = (x_u, y_u, z_u, \rho_u)^T = (X_u^T, \rho_u)^T$  around the working point  $X_0$ , we get the linearized measurement equation with a fault

$$Y = R - R_0 \simeq H_0(X - X_0) + \xi (+\theta), \quad (24)$$

where  $R_0 = (r_{10}, \dots, r_{n0})^T$ ,  $r_{i0} = \|X_i - X_{u0}\|_2 + c\rho_{u0}$ ,  $\xi = (\xi_1, \dots, \xi_n)^T$ ,  $H_0 = \frac{\partial R}{\partial X} \Big|_{X=X_0}$  is the Jacobian matrix of size  $n \times 4$ . As it follows from equations (24), a fault  $\theta$  affecting the GPS channels implies an additional error  $b = \mathbb{E}(\hat{X} - X) = (H_0^T H_0)^{-1} H_0^T \theta$  in the vector  $\hat{X}$ . It is obvious that the set of undetectable biases  $\theta$  is expressed by the equation  $W\theta = 0$ . All these vectors  $\theta$  can be expressed as :  $\theta = x\mathbf{1}_n$  with  $x \in \mathbb{R}$  and  $\mathbf{1}_n = (1, \dots, 1)^T$ . Fortunately, the impact  $b = x(H^T H)^{-1} H^T \mathbf{1}_n$  of such an undetectable bias  $x\mathbf{1}_n$  on the first three components  $\hat{x}_u, \hat{y}_u, \hat{z}_u$  is equal to zero, i.e.  $b_x = b_y = b_z = 0$ . Therefore, undetectable (by a ground monitoring station) pseudorange biases are not dangerous for the navigation.

## References

- [1] Basseville M., and Nikiforov I. Fault isolation for diagnosis: nuisance rejection and multiple hypotheses testing. Survey Paper, 15 th IFAC World Congress, Barcelona, Spain 21 th - 26 th July, 2002, pp. 179-190
- [2] Borovkov A. *Statistique mathématique*. Edition Mir, Moscou, 1987.
- [3] Frank P. M. Fault diagnosis in dynamic systems using analytical and knowledge based redundancy - A survey and new results, *Automatica*, vol.26, no.3, pp. 459-474, 1990.
- [4] Lehman E. L. *Testing Statistical Hypotheses*, Chapman and Hall, 1986.
- [5] Nikiforov I. Eléments de théorie de la décision statistique II : compléments, Chap. 2 in book *Décision et reconnaissance des formes en signal*, pp. 43-84, Hermes Science Publications (Traité IC2), Edited by R. Lengellé, 2002
- [6] L.L. Scharf and B. Friedlander Matched subspace detectors, *IEEE Trans. Sign. Proc.*, vol. 42, no. 8, August, 2146-2157, 1994
- [7] Staroswiecki M. *Redondance analytique in Automatique et statistiques pour le diagnostic*, Hermas Science Europe, 43-68, 2001.
- [8] Strang G. *Linear Algebra And Its Applications*, Harcourt Brace and Company International Edition, 1986.
- [9] Vulnerability Assessment Of the Transportation Infrastructure Relying on the GPS, John A. Volpe National Transportation System Center, Final Report, U.S. Department of Transportation, August 29, 2001.
- [10] Wald A. Tests of Statistical Hypotheses Concerning Several Parameters when the number of Observations is Large. *Trans. Amer. Math. Soc.*, 54, 426-482, 1943.