Available online at www.sciencedirect.com
SCIENCE DIRECT®
automatica
Automatica 41 (2005) 1157–1171
www.elsevier.com/locate/automatica

Optimal statistical fault detection with nuisance parameters☆

Mitra Fouladirad, Igor Nikiforov*

ISTIT, FRE CNRS 2732, Université de Technologie de Troyes, BP 2060, 10010 Troyes Cedex, France

Received 19 January 2004; received in revised form 9 August 2004; accepted 10 February 2005

Abstract

Fault detection is addressed within a statistical framework. The goal of this paper is to propose an optimal statistical tool to detect a fault in a linear stochastic (dynamical) system with uncertainties (nuisance parameters or nuisance faults). 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). Examples of ground station based and receiver autonomous Global Positioning System (GPS) integrity monitoring illustrate the proposed method.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Fault detection; Linear systems; Stochastic systems; Decision theory; Tests; Nuisance parameter; Parity space

1. Introduction

A key issue in fault detection/isolation is to state the significance of the observed deviation (fault) with respect to random noises, deterministic uncertainties (also called nuisance parameters) or nuisance faults (in case of fault isolation). The purpose of this paper is to propose an optimal statistical tool to detect a fault in a linear stochastic (dynamical) system with nuisance parameters or nuisance faults. It is supposed that the nuisance parameters (faults) are unknown but non-random. In automatic control community, the problem of nuisance parameters (or nuisance faults) 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 interested readers can find the details from the survey papers (Mironovski, 1980; Patton & Chen, 1991; Frank, 1990; Staroswiecki, 2001) and the books (Patton, Frank, & Clark, 1989; Gertler, 1998; Chen & Patton, 1999). The theory of analytical redundancy is mainly deterministic; the random noises are only heuristically treated. Recently, the deterministic analytical redundancy approach, namely the fault detection problem with nuisance parameters or nuisance faults, has been extended to the case where the process and sensor noises are also considered (Chen, Mingori, & Speyer, 2003). The stochastic fault detection filter is derived by maximizing the transmission from the target fault to the projected output while minimizing the transmission from the nuisance parameters (nuisance faults) and from the process and sensor noises. Let us recall that the performance of any statistical fault detector is defined with the probabilities of false decisions (false alarm and non-detection). However, the cost criterion proposed in Chen et al. (2003) does not include the probabilities of false alarm and non-detection. Instead of this, the cost criterion includes the covariance matrices representing the transmissions from the target fault, nuisance faults and noises, as in traditional analytical redundancy papers.

When the impact of random noise is non-negligible the problem of statistical fault detection should be addressed (Basseville, 1997; Basseville & Nikiforov, 2002). This problem is especially important in safety-critical applications where the probabilities of false decisions are widely used to describe the minimum operational performance

☆ Part of this work has been presented at Qualita 2003, Nancy, France, March 18–20, 2003, at the American Control Conference, Denver, Colorado, USA, June 4–6, 2003 and at the SAFEPROCESS 2003, Washington, D.C., USA, June 9–11, 2003.
* Corresponding author. Tel.: +33 3 25 71 56 78; fax: +33 3 25 71 56 99.
E-mail address: nikiforov@utt.fr (I. Nikiforov).

0005-1098/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2005.02.004

requirements. Traditionally, the fault detection problem is split into two steps (Basseville, 1997): residual generation and residual evaluation (design of decision rules based on these residuals). Even if each step is designed by using an optimal approach (with a particular optimality criterion), the optimality of the total procedure with respect to the probabilities of false alarm and non-detection is not warranted.

On the other hand, there exist an invariant hypotheses 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 $\mathscr{H}_0$ (no faults) and $\mathscr{H}_1$ (there exists a fault). But the traditional hypotheses testing approaches use relatively simple (with respect to automatic control community demands) statistical models (i.e. parameterized distributions). Hence, these methods should be adapted to more complex models used in fault detection. Another difficulty of hypotheses testing approach is the optimality criterion definition in the case of vector parameter and/or composite hypotheses, which represents a considerable practical interest in fault detection.

Therefore, the following paradoxical situation takes place in the fault detection. The ideas of analytical redundancy are used more then 30 years and almost considered as standard. Recently, the statistical fault detection methods are integrated into two steps procedure: residual generation by using analytical redundancy and residual evaluation by using statistical tests, but the following problems remain unsolved:

(1) What is a lower bound for the probability of non-detection over the class of detectors with a given probability of false alarm and what is the fault detection method that attains this lower bound when dealing with an unknown vector of fault and nuisance parameters?
(2) It is well-known that the choice of the parity space can be multiple. Hence, what is the optimal with respect to the false alarm—non-detection criterion nuisance parameter rejector (or residual generator)?
(3) Sometimes it is necessary to deal with a parity subspace. How to compare the performance index (false alarm against non-detection) of the full-set parity vector against a subset one?
(4) Sometimes the dynamical model of the nuisance parameters is available. How to estimate the impact of this model on the performance (false alarm against non-detection) of the optimal fault detector?

The contribution of this paper consists in the development of an optimal statistical approach of fault detection 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 (for example, by a hostile environment) to maximize its negative impact on the monitored system (for instance, to mask a fault). It is worth noting that this is a so-called "worst-case" approach (with respect to the nuisance parameter) that is especially interesting for monitoring the systems operating in a hostile environment. We propose a lower bound for the probability of non-detection (optimality theorem). To our knowledge no proof of optimality in a mathematically precise sense exists. It is shown that the optimal solution coincides with the generalized likelihood ratio test applied to the parity vector computed by using traditional analytical redundancy approach. By using the proposed optimality theorem, the problem of optimal residual generation is solved, the full-set parity vector is compared against a subset one and the impact of the dynamic nuisance parameters is evaluated.

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. Examples of the GPS integrity monitoring illustrating the relevance of the proposed tools are described in Section 6.

2. Problem statement

2.1. Statistical hypotheses testing problem

There are two problem statements in the theory of statistical fault detection: the hypotheses testing problem and the change detection problem (Basseville & Nikiforov, 2002). In the present paper, we address the binary hypotheses testing problem. The quality of a test δ is defined with the probability of false alarm: $\alpha = \Pr_0(\delta \neq \mathscr{H}_0)$, where $\Pr_i$ stands for observations $Y_1, \ldots, Y_k$ being generated by distribution P_i, and the *power function*: $\beta_\delta(\theta) = \Pr_\theta(\delta = \mathscr{H}_1)$. In case of a vector parameter θ, 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 the parameter θ is scalar, the family of distributions has a monotone likelihood ratio, and the test is one-sided (Borovkov, 1998; Lehmann, 1986). 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 (Wald, 1943) and the theory of invariant tests. We also adapt the Wald's theory to the case where the parameter of interest (fault) θ 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 θ detection in the following linear gaussian model (some motivation for this model can be found in Basseville (1997)):

$$Y = HX + M\theta + \xi, \tag{1}$$

where $Y \in \mathbb{R}^n$ is the measured output, $\theta \in \mathbb{R}^r$ is the parameter (fault) of model (1), $X \in \mathbb{R}^m$ is an unknown and non-random state vector (nuisance parameter), M is a full column rank matrix of size $(n \times r)$ with $r < n$ and H is a matrix of size $(n \times m)$ with rank $H = q$. It is assumed that $n \geqslant q + r$ and that the noise ξ follows a zero-mean gaussian distribution, i.e. $\xi \sim \mathcal{N}(0, \sigma^2 I_n)$, with the known variance $\sigma^2 > 0$ and the $(n \times n)$ identity matrix I_n. The assumption that the noise covariance matrix is scalar is not restrictive. It is used to obtain a simple and lucid presentation. As it will be shown in Section 3, the general covariance matrix case immediately follows from the scalar matrix one. Second, the developed tool is used to compare the full-set parity vector against a subset one, to solve the problem of "optimal residual generation" and to examine the impact of nuisance parameter modeled by a state-space model. To conclude this section, let us discuss the practical relevance of the "false alarm-non-detection" criterion and the invariant hypotheses testing approach and also let us provide the reader with some ideas how to use in practice a lower bound for the probability of non-detection.

2.2. Practical motivation of the criterion of optimality

For many safety-critical applications (such as aircraft navigation systems), a major problem of the existing systems consists in its lack of integrity. In the case of navigation system, the integrity monitoring concept defined by the International Civil Aviation Organization (ICAO) requires that a navigation system detects faults and removes them from the navigation solution before they sufficiently contaminate the output. The recent researches show that the detection/exclusion of the navigation message contamination is crucially important for the radio-navigation. It is proposed "to encourage all the transportation modes to give attention to autonomous integrity monitoring of GPS signals" (John Volpe National Transportation System Center, 2001). It is intuitively obvious that the criterion which must be used should favor reliable detection with few false alarms. A small probability of non-detection is necessary because the abnormal measurements are taken in the navigation systems, which leads to the unacceptably large position error and, hence, is clearly very undesirable. On the other hand, false alarms result in lower accuracy of the position estimate because some correct information is not used. The optimal solution involves a tradeoff between these two contradictory requirements. It is worth to note that the probabilistic criterion "false alarm-non-detection" is traditionally used by ICAO as an engineering language to specify and analyse the minimum operational performance requirements (see John Volpe National Transportation System Center (2001), RTCA/DO-229A (1998)).

2.3. Practical motivation of the invariant hypotheses testing approach

The idea of the invariant hypotheses testing approach is based on the existence of the natural invariance of the detection problem with respect to a certain group of transformation. The magnitude of the nuisance parameter (fault) X is unlimited. Its impact, modeled by the term HX in Eq. (1) or by a state-space model in Section 5, defines a subspace in the observation space $\mathscr{Y} = \mathbb{R}^n$. The ability of the invariant hypotheses testing approach to provide us with algorithms which detect the target fault θ while being insensitive to the nuisance parameters (or faults) X makes this theory a serious candidate for the design of the monitoring systems operating in hostile environment. In such a context the deterministic nuisance X with unlimited magnitude can be used by a hostile environment to mask a target fault impact modeled by $M\theta$. Let us also stress the drawbacks of the Bayesian approach in such a situation: this approach exploits some a priori information on the distribution of X but this information may not be reliable (hostile environment!). Hence, the Bayesian approach is irrelevant to the case of nuisance parameters governed by a hostile environment. In contrast with the Bayesian approach, the invariant hypotheses testing theory is based on the nuisance X rejection and, therefore, does not use any a priori information on the distribution of X.

2.4. Discussion: how to use in practice a lower bound for the probability of non-detection

There exists an opinion among some specialists that the problem of lower bound for the probability of non-detection in a certain class of tests has a purely theoretical character. But this problem is also of practical interest. Let us discuss now how to use in practice a lower bound for the probability of non-detection.

First, if the integrity requirements are defined in terms of "false alarm-non-detection" criterion, then it is important to know the best theoretically achievable level of integrity. It can happen that the integrity level required by minimum performance requirements cannot be achieved even theoretically. Such a conclusion is very important for the system designers: either they reduce the required integrity level or they improve the sensor performances (augment, for instant, the signal-to-noise ratio).

Second, a lower bound for the probability of non-detection is useful as a *benchmark* for suboptimal algorithms with a reduced dimension of parity space or with a nonlinear model.

Third, often, the problem of integrity monitoring is twofold. Namely, the fault tolerance is achieved by implementing an efficient fault detection algorithm and optimal sensor configurations. Hence, to solve the problem of optimal sensor configurations we should use a relation between the statistical performances of the fault detection algorithm and some indexes of sensor configurations. For example, we would like to verify that the requirements of integrity monitoring can be achieved (at least theoretically) by using a given number and orientation of sensors. In such a situation, it is not reasonable to use the performance index of a particular algorithm to prove this fact, on the contrary, we have to use a lower bound that characterizes a certain class of detection algorithms. The number of sensors and their orientations (together with the model of faults) define the matrices H and M and the lower bounds provide us with the relation between the sensor configuration and the achievable integrity level in a certain class of fault detection algorithms. Another task of sensor configurations arise when we have to estimate the potential capability of fault detection scheme for a given sensor configuration, for example, for the "worst case" GPS navigation satellite geometries. And again a lower bound for the probability of non-detection plays a key role in this investigation.

3. Simple linear model

The goal of this section is to apply the Wald's (1943) 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.

3.1. Design of the test and its optimality

Originally, Wald has solved the following problem: the observation $Y \in \mathbb{R}^n$ is generated by one of two gaussian distributions: $\mathcal{N}(0, \Sigma)$ and $\mathcal{N}(\theta \neq 0, \Sigma)$, where θ is the mean vector and Σ is a positive definite covariance matrix. The hypotheses testing problem consists in deciding between $\mathscr{H}_0$: $\{\theta = 0\}$ and $\mathscr{H}_1$: $\{\theta \neq 0\}$. The peculiarity of the above problem is that a uniformly most powerful (UMP) test does not exist in case of a vector parameter θ. To overcome this difficulty, Wald proposes to impose an additional constraint on the class of considered tests, namely, a *constant power function* over a family of surfaces $\mathscr{S}$ defined on the parameter space Ω, in order to avoid the existence of UMP tests over a subspace $\overline{\Omega}$ of Ω which are very inefficient over $\Omega \backslash \overline{\Omega}$. Let us assume now the following gaussian linear model:

$$Y = M\theta + \xi, \tag{2}$$

where $Y \in \mathbb{R}^n$ is the observation vector, $\theta \in \mathbb{R}^r$ is the parameter of model (2), i.e. the vector of faults, M is a full column rank matrix of size $(n \times r)$ with $r < n$ and ξ is a zero-mean gaussian noise $\xi \sim \mathcal{N}(0, \sigma^2 I_r)$, $\sigma^2 > 0$. As in the previous case, the hypotheses testing problem consists in deciding between $\mathscr{H}_0 : \{\theta = 0\}$ and $\mathscr{H}_1 : \{\theta \neq 0\}$. Let us apply the general Wald's idea to design a test that can potentially be UBCP:

$$\delta^*(Y) = \begin{cases} \mathscr{H}_0 & \text{if } \Lambda(Y) = \widehat{\theta}^{\mathrm{T}} \mathscr{F}_{\widehat{\theta}} \widehat{\theta} < h(\alpha), \\ \mathscr{H}_1 & \text{if } \Lambda(Y) = \dfrac{1}{\sigma^2} Y^{\mathrm{T}} M (M^{\mathrm{T}} M)^{-1} M^{\mathrm{T}} Y \\ & \quad \geqslant h(\alpha), \end{cases} \tag{3}$$

where $\widehat{\theta} = (M^{\mathrm{T}} M)^{-1} M^{\mathrm{T}} Y$ is the least-square (LS) estimator of θ and $\mathscr{F}_{\widehat{\theta}} = 1/\sigma^2 M^{\mathrm{T}} M$, is the Fisher matrix. Now, it is necessary to prove that the test δ^* given by Eq. (3) is UBCP over the following family of ellipsoids:

$$\mathscr{S}_M = \left\{ S_c : \theta^{\mathrm{T}} \mathscr{F}_{\widehat{\theta}} \theta = \frac{1}{\sigma^2} \| M\theta \|_2^2 = c^2, c > 0 \right\}, \tag{4}$$

because the results of Wald (1943) does not cover the regression model given (2). The main result of this section is established in the following.[1]

Theorem 1. *Let us consider the regression model* (2). *The test $\delta^*(Y) \in \mathscr{K}_\alpha$, given by Eq.* (3), *is UBCP for deciding between the hypotheses $\mathscr{H}_0 : \{\theta = 0\}$ and $\mathscr{H}_1 : \{\theta \neq 0\}$ over the family of ellipsoids* (4).

Proof. See Appendix A. □

3.2. Discussion of the UBCP test

Let us discuss now some issues in detecting the fault vector θ in the model $Y = M\theta + \xi$ (or statistically speaking, deciding between $\mathscr{H}_0 : \{\theta = 0\}$ and $\mathscr{H}_1 : \{\theta \neq 0\}$). First of all, the test $\delta^*(Y)$ coincides with the generalized likelihood ratio (GLR) test. The GLR test is widely used in practice, but the problem of its non-asymptotic optimality remains unsolved. The interested reader can find the details in Nikiforov (2002, Chapter 2). As it follows from Koch (1999, Chapter 2.7), the statistics $\Lambda(Y)$ is distributed according to the χ^2 law with r degrees of freedom. This law χ^2 is central under $\mathscr{H}_0$ and non-central under $\mathscr{H}_1$ with the non-centrality parameter $\lambda = 1/\sigma^2 \| M\theta \|_2^2$. Hence, the power function $\beta_{\delta^*}(\theta : (1/\sigma^2)\theta^{\mathrm{T}} M^{\mathrm{T}} M \theta = c^2) = \mathrm{Pr}_{c^2}(\Lambda(Y) \geqslant h(\alpha))$ is constant on the same surface $S_c : 1/\sigma^2 \| M\theta \|_2^2 = c^2$ (see (4)). Because the power is constant on the same surface S_c, it is reasonable to present the power as a function of c^2: $c^2 \mapsto \beta_{\delta^*}(c^2)$. A typical family of concentric ellipses and the power function β_{δ^*} are depicted in Fig. 1.

Finally, let us consider that the noise ξ of model (2) follows the gaussian distribution $\mathcal{N}(0, \Sigma)$, where Σ is a known (positive definite) covariance matrix. By using the change of variables $g(X) = R^{-1} X$, where the symmetric matrix R is define by $\Sigma = RR$, and the invariance properties of the gaussian family $\mathcal{N}(\theta, \Sigma)$ (see Borovkov (1998)), the new hypotheses testing problem can be reduced to the problem presented in Theorem 1.

4. Linear model with nuisance parameters

The goal of this section is to apply the results obtained in Section 3 to a linear gaussian model with a nuisance parameter.

[1] It is worth emphasizing that in contrast with the asymptotic hypotheses testing theory of Wald (when $N \to \infty$) (Wald, 1943), the non-asymptotic theory is discussed in the paper. Hence, the known results on asymptotic optimality is not applicable here and the extension of Wald's results is given by Theorem 1.

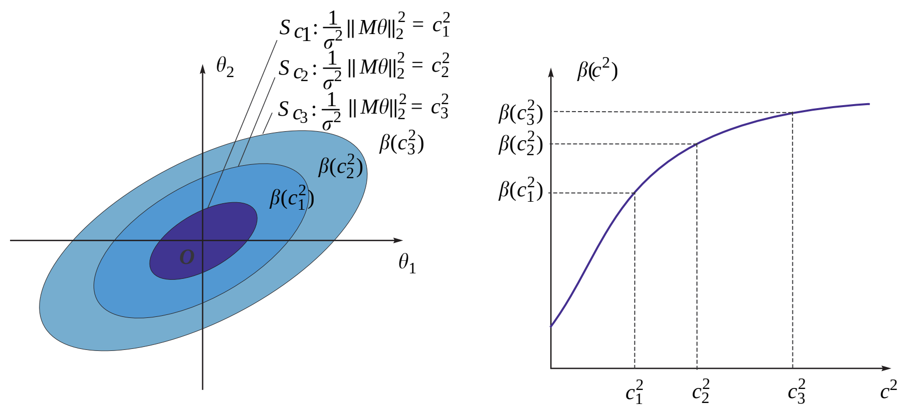

Fig. 1. Family of concentric ellipses and the power function $\beta(c^2)$.

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 $\mathscr{H}_0 : \theta = 0$ and $\mathscr{H}_1 : \theta \neq 0$, 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 Λ. First of all, let us note that the above mentioned hypotheses testing problem remains invariant under the group of translations $G = \{g : g(Y) = Y + HC\}$, $C \in \mathbb{R}^m$. 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 (Frank, 1990). The parity vector $Z = WY$ is the transformation of the measured output Y into a set of $n - q$ linearly independent variables by projection onto the left null space of the matrix H. The matrix $W^{\mathrm{T}} = (w_1, \ldots, w_{n-q})$ of size $n \times (n - q)$ is composed of the eigenvectors $w_1, \ldots, w_{n-q}$ of the projection matrix $P_H = I_n - H(H^{\mathrm{T}}H)^{-}H^{\mathrm{T}}$, where A^{-} is a generalized inverse of A (Koch, 1999, Chapter 1.5.3), corresponding to eigenvalue 1. If the matrix H is full column rank, i.e. $q = m$, then $P_H = I_n - H(H^{\mathrm{T}}H)^{-1}H^{\mathrm{T}}$. The matrix W satisfies the following conditions:

$$WH = 0, \quad W^{\mathrm{T}}W = P_H, \quad WW^{\mathrm{T}} = I_{n-q}. \tag{5}$$

Example 1 (*A simple case*). The rejection of the nuisance parameters in a simple case where $Y \in \mathbb{R}^3$ and $x \in \mathbb{R}$ is depicted in Fig. 2. Let us assume that $H = (1, 1, 1)^{\mathrm{T}}$. Therefore, the column space of H is $R(H) = \{Y | Y = Hx, x \in \mathbb{R}\}$. Its orthogonal complement (parity space) $R(H)^{\perp} = \{Y | Y = w_1a_1 + w_2a_2, \ (a_1, a_2) \in \mathbb{R}^2\}$ is spanned by the vectors $w_1 = (\sqrt{6}/3, -\sqrt{6}/6, -\sqrt{6}/6)^{\mathrm{T}}$ and $w_2 = (0, \sqrt{2}/2, -\sqrt{2}/2)^{\mathrm{T}}$ (see Fig. 2). The "cloud" of random observations Y_i around the mean value $\mathbb{E}(Y) = Hx + M\theta$ is shown in Fig. 2. The projection $P_HM\theta$ of the deterministic part $\mathbb{E}(Y)$ onto the parity space (plane) spanned by the vectors w_1 and w_2 and the "cloud" of random residuals $e_i = P_HY_i$ are also shown in Fig. 2. As it follows from the first of the above conditions (5), the transformation by W completely removes the interference of the nuisance parameter x. It is illustrated by Fig. 2, where the residual's $e_i = P_HY_i$ position is independent of the nuisance x, as desired.

Moreover, as it can be shown, the statistics $Z = WY$ is maximal invariant to the group G. For this reason all invariant tests should depend on Y only via the vector $Z = WY$. Therefore, the measurement model (1) can be rewritten by the following manner:

$$Z = WY = WM\theta + W\xi = WM\theta + \zeta, \tag{6}$$

where $\zeta \sim \mathcal{N}(0, \sigma^2 I_{n-q})$, $\sigma^2 > 0$. It is assumed that $r \leqslant n - q$ (see below Eq. (1)), let us additionally assume that the matrix WM is full column rank of size $((n-q) \times r)$. Hence, the results of Section 3 can be directly applied to the model given by Eq. (6) for deciding between $\mathscr{H}_0 : \{\theta = 0\}$ and $\mathscr{H}_1 : \{\theta \neq 0\}$.

Example 2. To explain the above-mentioned assumptions on the matrix WM, let us briefly discuss here two practical examples when this is not so. First, sometimes the observation model allows the existence of fault vector θ of size n, for instance, within the framework of GPS integrity monitoring (the detailed discussion can be found in Section 6), all satellite channels potentially can be simultaneously contaminated by additional pseudorange biases, hence the linearized GPS observation model is given by $Y = HX + \theta + \xi$. Such a situation ($M = I_n$ and the inequality $r \leqslant n - q$ is not satisfied) is especially important in the case of so-called "intentional" contamination (John Volpe National Transportation System Center, 2001). This leads to the problem of fault detectability (see Section 6.1.2): i.e. some combinations of pseudorange biases are undetectable. In respect to the problem of GPS integrity monitoring, it is worth to note that traditionally only one individual satellite channel fault is assumed

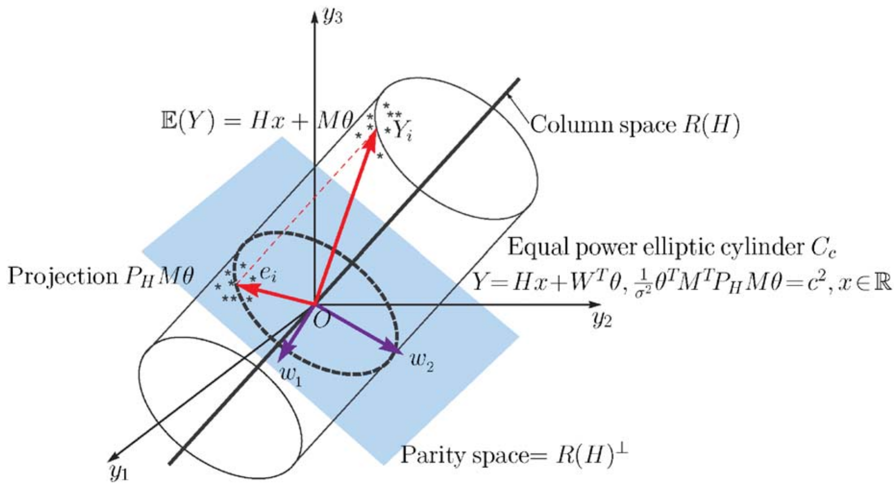

Fig. 2. Column space and its orthogonal complement (parity space).

at a time (RTCA/DO-229A, 1998) but now recently arising integrity monitoring scenarios ("intentional" contamination) stress the importance of fault detectability. To solve (at least partially) this problem, it is proposed in Section 6.1.2 to split the total fault space into two subspaces: a subspace of "detectable" faults and a subspace of "undetectable" ones. More generally, let us assume that the matrix H is full column rank (i.e. $q = m$) and the number of nuisance (state) parameters m plus the number of possible faults r is greater than the number of observations n (again the inequality $r \leqslant n - q$ is not satisfied). In this case (exactly as in the previous one) some fault combinations will be undetectable. For example, in the case of autonomous GPS integrity monitoring the number of nuisance parameters m is equal to 4 (3 space coordinates and the user's clock error). Let us assume that $n = 7$ satellites are visible. If four GPS channels ($r = 4$) can be contaminated simultaneously, then some individual fault combinations will be undetectable because the dimension of parity space $n - m = 3$ is insufficient.

Second, let us assume that the inequality $r \leqslant n - q$ is satisfied but some columns of M belong to the column space $R(H)$ of the matrix H. The following numerical example illustrates this situation:

$$Y = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} x + \begin{pmatrix} 4 & 2 \\ 2 & 1 \\ 10 & 4 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \xi,$$

where $\xi \sim \mathcal{N}(0, I_3)$.

It is easy to see that the first column of M belongs to $R(H)$. Now, $n = 3$, $m = 1$, $r = 2$, $q = \operatorname{rank} H = 1$, $\operatorname{rank} M = 2$ and $r = 2 = n - q$ but the matrix in question

$$WM = \begin{pmatrix} 0 & 0.1221 \\ 0 & -0.3896 \end{pmatrix}$$

is not full column rank, $\operatorname{rank}(WM) = 1$. The consequence of this fact is that the first fault component θ_1 is undetectable. In summary, it can be concluded that the practical goal of the above-mentioned assumptions on the matrix WM is to warrant the detectability of any fault vectors $\theta \neq 0$.

4.2. Design of the UBCP invariant test

Let us apply Theorem 1 to design the UBCP invariant test. In the case of model (6), the LS estimator of θ is given by: $\widehat{\theta} = ((WM)^{\mathrm{T}}(WM))^{-1}(WM)^{\mathrm{T}}Z$ and the Fisher matrix is $\mathscr{F}_{\widehat{\theta}} = 1/\sigma^2(WM)^{\mathrm{T}}WM$. It directly follows from Eq. (3) that the UBCP test is given by

$$\delta^*(Y) = \begin{cases} \mathscr{H}_0 & \text{if } \Lambda(Y) < h(\alpha), \\ \mathscr{H}_1 & \text{if } \Lambda(Y) \geqslant h(\alpha), \end{cases} \tag{7}$$

where $\Lambda(Y) = 1/\sigma^2 Y^{\mathrm{T}} P_H M (M^{\mathrm{T}} P_H M)^{-1} M^{\mathrm{T}} P_H Y$. The test given by Eq. (7) is UBCP over the following family of surfaces (see Fig. 2)

$$\mathscr{S}_{WM} = \left\{ S_c : \frac{1}{\sigma^2} \|WM\theta\|_2^2 = c^2, \; c > 0 \right\}. \tag{8}$$

4.3. Discussion of the UBCP invariant test

As in the previous case, the invariant test $\delta^*(Y)$ coincides with the GLR test. The statistical properties of the GLR test have been examined in Scharf and Friedlander (1994) in the context of invariance. It has been mentioned in Scharf and Friedlander (1994) that the GLR test is (UMP) optimal because the χ^2 distribution is monotone in the non-centrality parameter λ (in other words the family of distributions possesses a monotone likelihood ratio (Borovkov, 1998; Lehmann, 1986)). A min-max approach with respect to the unknown input has been developed in Rougée, Basseville, Benveniste, and Moustakides (1987) where the hypotheses have also been formulated in terms of non-centrality parameter. Therefore, Scharf and Friedlander (1994) and Rougée et al. (1987) 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 Scharf and Friedlander (1994), Rougée et al. (1987) we prove the optimality of the test *directly* in the parameter space.

The statistics $\Lambda(Y)$ of the UBCP invariant test given by Eq. (7) is distributed according to the χ^2 law with r degrees of freedom (see Section 3.2).

Example 3 (*A simple case—continued*). Let us recall a simple model discussed in Example 1. It is assumed here that $\theta \in \mathbb{R}^2$ (see Fig. 2). Sometimes it is interesting to draw an equal power surface of the UBCP invariant in the observation space ($Y \in \mathbb{R}^3$). This surface is an elliptic cylinder C_c given by the following parametric representation: $Y = Y(x, \theta)$: $C_c = \{Y = Hx + W^{\mathrm{T}}\theta\}$, where $\theta : (1/\sigma^2)\theta^{\mathrm{T}}M^{\mathrm{T}}P_H M\theta = c^2$ and $x \in \mathbb{R}$. The intersection of the plane spanned by the vectors w_1 and w_2 (parity space) with the elliptic cylinder C_c is an ellipse given by equation $c^2 = (1/\sigma^2)\theta^{\mathrm{T}}M^{\mathrm{T}}P_H M\theta$ in the plane (parity space). This situation is depicted in Fig. 2.

4.3.1. *Full-set parity vector against a subset one*

This question arises in connection with the above mentioned problem of "optimal residual generation" (see, for instance, the survey (Staroswiecki, 2001)). 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 chosen so that $r < k < n - q$, from the matrix W by deleting $n - q - k$ rows of W. It is easy to see that $W_1 H = 0$ and $W_1 W_1^{\mathrm{T}} = I_k$. Finally, we obtain $Z_1 = W_1 Y = W_1 M\theta + W_1\xi = W_1 M\theta + \zeta_1$, where $\zeta_1 \sim \mathcal{N}(0, \sigma^2 I_k)$, $\sigma^2 > 0$. For this model the LS estimator of θ is given by: $\widehat{\theta}_1 = ((W_1M)^{\mathrm{T}}(W_1M))^{-1}(W_1M)^{\mathrm{T}}W_1Y$ and the Fisher matrix is $\mathcal{F}_{\widehat{\theta}_1} = 1/\sigma^2 M^{\mathrm{T}}W_1^{\mathrm{T}}W_1M$. It follows from (3) that the new test is given by

$$\delta_1^*(Z_1) = \begin{cases} \mathcal{H}_0 & \text{if } \Lambda(Z_1) < h_1(\alpha), \\ \mathcal{H}_1 & \text{if } \Lambda(Z_1) \geqslant h_1(\alpha), \end{cases} \tag{9}$$

where $\Lambda(Z_1) = 1/\sigma^2 Z_1^{\mathrm{T}}W_1M((W_1M)^{\mathrm{T}}W_1M)^{-1}M^{\mathrm{T}}W_1^{\mathrm{T}}Z_1$ has a constant power over the family of ellipsoids

$$\mathcal{S}_{W_1M} = \left\{S_c : \frac{1}{\sigma^2}\|W_1M\theta\|_2^2 = c_1^2,\ c_1 > 0\right\}. \tag{10}$$

Lemma 1. *Let us assume that* $\Pr_0(\delta_1^* \neq \mathcal{H}_0) = \Pr_0(\delta^* \neq \mathcal{H}_0) = \alpha$, $(0 < \alpha < 1)$. *Then the following inequality is satisfied for any* $\theta \neq 0$: $\beta_{\delta_1^*}(\theta) \leqslant \beta_{\delta^*}(\theta)$.

Proof. See Appendix B. □

4.3.2. *Discussion*

The above-mentioned inequality becomes strict under some additional conditions. Let us assume, without loss of generality, that the matrix W of size $((n-q) \times n)$ is composed of two blocks W_1 and D of size $(k \times n)$ and $((n-q-k) \times n)$, respectively: $W = (W_1^{\mathrm{T}}D^{\mathrm{T}})^{\mathrm{T}}$. As it is shown in Appendix B, to compare the power functions of the tests δ^* and δ_1^* it is sufficient to compare the noncentrality parameters $c^2(\theta)$ and $c_1^2(\theta)$ for the same fault vector $\theta \neq 0$. Let us consider two particular cases:

(1) It is assumed that $k = n - q - 1$ and $r \geqslant 2$. Therefore, the matrix D is of size $(1 \times n)$. It is possible to find a direction in the parameter space defined by the vector $\theta^* \neq 0$ for which both tests $\delta^*(Z)$ and $\delta_1^*(Z)$ will have the same power, $c^2(\theta^*) - c_1^2(\theta^*) = 0 \iff \|DM\theta^*\|_2^2 = \|w_{n-q}^{\mathrm{T}}M\theta^*\|_2^2 = 0 \iff M^{\mathrm{T}}w_{n-q} \perp \theta^*$, where w_{n-q} is the $n-q$th eigenvector of the projection matrix P_H corresponding to the eigenvalue 1 (see Section 4.1).
(2) Let us assume now that $\mathrm{rank}(DM) = \dim(\theta)$. In this case $c_1^2(\theta) - c^2(\theta) = -1/\sigma^2\|DM\theta\|_2^2 < 0$ and, hence, $\beta_{\delta_1^*}(c_1^2(\theta)) < \beta_{\delta^*}(c^2(\theta))$ for any $\theta \neq 0$.

Example 4. Let us consider the following model:

$$Y = \begin{pmatrix} 10 \\ 1 \\ 10 \\ 1 \\ 1 \end{pmatrix} x + \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 3 & -4 \\ 4 & -1 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} + \xi,$$

where $\xi \sim \mathcal{N}(0, I_5)$. The nuisance rejection matrix is given by

$$W = \begin{pmatrix} -0.0702 & 0.9971 & -0.0289 & -0.0029 & -0.0029 \\ -0.7019 & -0.0289 & 0.7105 & -0.0289 & -0.0289 \\ -0.0702 & -0.0029 & -0.0289 & 0.9971 & -0.0029 \\ -0.0702 & -0.0029 & -0.0289 & -0.0029 & 0.9971 \end{pmatrix}.$$

Let us define two other rejection matrices W_1 and W_2. The matrix W_1 is obtained by deleting the last row of the matrix W and the matrix W_2 by deleting two last rows of W. We consider the full-set parity test $\delta^*(Z)$ and two subset parity tests $\delta_1^*(Z_1)$ and $\delta_2^*(Z_2)$ designed by using the statistics $Z_1 = W_1Y$ and $Z_2 = W_2Y$, respectively. All these statistics follow χ^2 central distribution with 2 degrees of freedom under $\mathcal{H}_0$. The comparison of these three tests are shown in Figs. 3 and 4. First of all, to show the advantage of the full-set parity vector, let us depict the equal power $\beta(c^2)$ ellipses of the tests δ^* (solid line), δ_1^* (dash-dotted line) and δ_2^* (dashed line) for the same constant, say $c = 1$ (see Fig. 3). Fig. 3 shows that the full-set parity test $\delta^*(Z)$ performs better than other tests. It is easy to see from Fig. 3 that the relative efficiency of the tests depends on the orientation (angle γ) of the vector $\theta = (\theta_1\theta_2)^{\mathrm{T}}$: the advantage of the full-set parity test over a subset one can be somewhat limited for some directions in the parameter space (see the discussion given in Section 4.3.2 for some additional comments). For this reason it is reasonable to compare the probability of non-detection $\Pr_\theta(\delta^*) = 1 - \beta(\theta)$ as a function of two variables: the signal-to-noise ratio $\|\theta\|_2/\sigma$ and the angle γ.

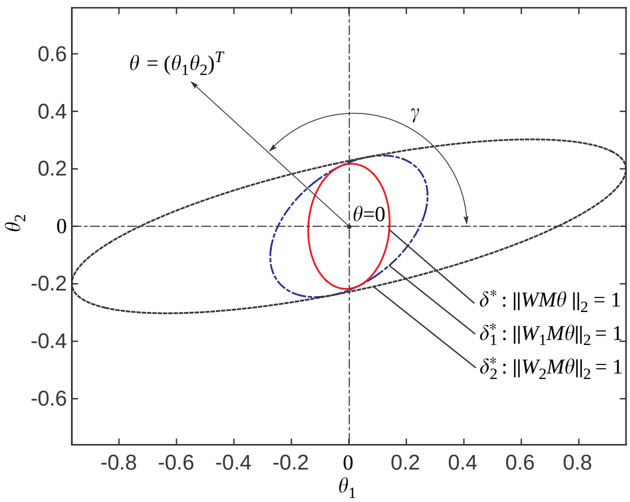

Fig. 3. The equal power ellipses of the tests δ^* (solid line), δ_1^* (dash-dotted line) and δ_2^* (dashed line).

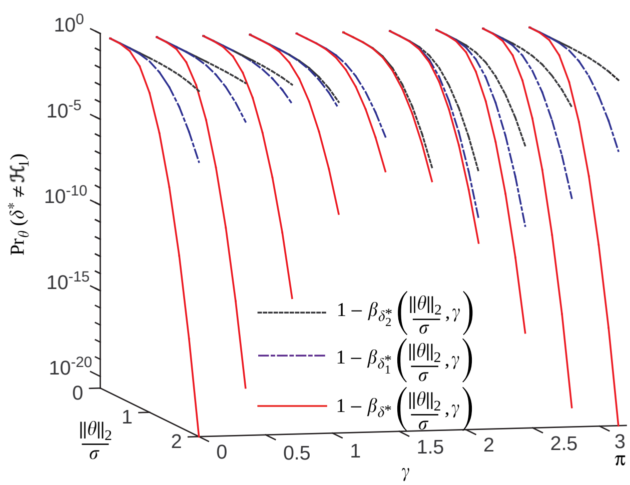

Fig. 4. Probability of non-detection for the full-set parity test $\delta^*(Z)$ (solid lines) and the subset parity tests $\delta_1^*(Z_1)$ (dash-dotted lines) and $\delta_2^*(Z_2)$ (dashed lines).

The functions

$$\left(\frac{\|\theta\|_2}{\sigma},\gamma\right) \mapsto 1-\beta\left(\frac{\|\theta\|_2}{\sigma},\gamma\right)$$

are shown in Fig. 4. Here, the probability of false alarm is chosen to be $\alpha = 10^{-5}$, the signal-to-noise ratio $\|\theta\|_2/\sigma$ varies between 0 and 2 and the angle γ varies between 0 and π. It is easy to see from Fig. 4 that the probability of non-detection $1-\beta_{\delta^*}(\|\theta\|_2/\sigma,\gamma)$ (solid lines) of the full-set test δ^* is typically much smaller than the probabilities of non-detection $1-\beta_{\delta_1^*}(\|\theta\|_2/\sigma,\gamma)$ (dash-dotted line) and $1-\beta_{\delta_2^*}(\|\theta\|_2/\sigma,\gamma)$ (dashed lines) of the subset tests δ_1^* and δ_2^*. However, if the angle γ is close to $\pi/2$, all the probabilities are comparable.

4.3.3. "Optimal" residual generation

As it is mentioned in the survey (Staroswiecki, 2001), the choice of the full-set rejection matrix W of size $((n-q)\times n)$ is not unique: the product AW, where A is a matrix of size $((n-q)\times(n-q))$ such that $\det A \neq 0$, leads to another rejection matrix $\widetilde{W} = AW$. It is asked in Staroswiecki (2001) 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 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* (7) *is designed by using a rejection matrix* $\widetilde{W} = AW$, *where* $\det A \neq 0$. *Then the power function of the test* (7) *remains independent of* A.

Proof. See Appendix C. □

5. Dynamical model with unknown inputs

The goal of this section is to apply the statistical tools developed in Sections 3 and 4 to the following linear state-space model where the nuisance is modeled by a deterministic state equation:

$$X_k = FX_{k-1} + BU_k, \tag{11}$$

$$Y_k = HX_k + M\theta_k + \xi_k, \tag{12}$$

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 ξ_k is a zero-mean gaussian white noise, $\xi_k \sim \mathcal{N}(0, \sigma^2 I_n)$, $\sigma^2 > 0$. It is assumed that all the matrices of (11)–(12) are known. The characteristic feature of the state-space model (11)–(12) is that the nuisance X_k is modeled by a deterministic state equation (11) which is governed by the input vector U_k. Let us assume that a fixed N-size sample of measured outputs $Y_1, \ldots, Y_N$ is available and supposed to be generated by one of the two alternative hypotheses. There are two methods to deal with the nuisance parameters $X_1, \ldots, X_N$: (i) to ignore the presence of the state equation (11) and to uniquely use the measurement equation (12); (ii) to use both Eqs. (11) and (12). It can happen that the parity space obtained by using both state and measurement equations is reachable than the parity space obtained by ignoring the state equation. It will be formally proven now that this fact can improve the performance of the UBCP invariant test based uniquely on the measurement equation (12).

5.1. Ignoring the state equation

Here, the UBCP invariant test is designed uniquely by putting together N measurement Eqs. (12). Eq. (11) is ignored.

$$\mathcal{Y} = \mathcal{H}\mathcal{X} + \mathcal{M}\Theta + \Xi, \tag{13}$$

where the definitions of the vectors $\mathscr{Y}=(Y_1^{\mathrm{T}},\ldots,Y_N^{\mathrm{T}})^{\mathrm{T}}$, $\mathscr{X}=(X_1^{\mathrm{T}},\ldots,X_N^{\mathrm{T}})^{\mathrm{T}}$, $\Theta=(\theta_1^{\mathrm{T}},\ldots,\theta_N^{\mathrm{T}})^{\mathrm{T}}$, $\Xi=(\xi_1,\ldots,\xi_N)^{\mathrm{T}}$, $\mathscr{H}$ and $\mathscr{M}$ are $(Nn\times Nm)$ and $(Nn\times Nr)$ block matrices, respectively, in which the blocks off the "diagonal" are the zero matrices and the "diagonal matrices" are H and M, respectively. The UBCP invariant test (7) is applicable now by replacing the vector Y, matrices H, M and P_H by $\mathscr{Y}$, $\mathscr{H}$, $\mathscr{M}$ and $\mathscr{P}_{\mathscr{H}}=I_{Nn}-\mathscr{H}(\mathscr{H}^{\mathrm{T}}\mathscr{H})^{-}\mathscr{H}^{\mathrm{T}}$, respectively. This test $\delta^*_{\mathscr{H}}$ is UBCP over the following family of ellipsoids $\mathscr{S}_{\mathscr{W}_{\mathscr{H}}\mathscr{M}}=\{S_c: 1/\sigma^2\|\mathscr{W}_{\mathscr{H}}\mathscr{M}\Theta\|_2^2=c^2_{\mathscr{H}}, c_{\mathscr{H}}>0\}$, where the matrix $\mathscr{W}_{\mathscr{H}}$ satisfies the following conditions $\mathscr{W}_{\mathscr{H}}\mathscr{H}=0$, $\mathscr{W}^{\mathrm{T}}_{\mathscr{H}}\mathscr{W}_{\mathscr{H}}=\mathscr{P}_{\mathscr{H}}$ and $\mathscr{W}_{\mathscr{H}}\mathscr{W}^{\mathrm{T}}_{\mathscr{H}}=I_{Nn-Nq}$.

5.2. Using both state-space and measurement equations

The nuisance X_k is replaced by $X_k=F^{k-1}X_1+\sum_{j=0}^{k-2}F^jBU_{k-j}$ for $k\geqslant 2$. This implies that $Y_1=HX_1+M\theta_1+\xi_1$ and $Y_k=HF^{k-1}X_1+H\sum_{j=0}^{k-2}F^jBU_{k-j}+M\theta_k+\xi_k$ for $k\geqslant 2$. By putting together N measurement equations, we get:

$$\begin{aligned}\mathscr{Y}&=\mathscr{H}_F X_1+\mathscr{H}_{F,B}\mathscr{U}+\mathscr{M}\Theta+\Xi\\&=\widetilde{\mathscr{H}}\widetilde{\mathscr{X}}+\mathscr{M}\Theta+\Xi,\end{aligned}\tag{14}$$

where the definitions of the vector $\mathscr{U}$ and the matrices $\mathscr{H}_F$, $\mathscr{H}_{F,B}$, $\mathscr{M}$ are trivial, $\widetilde{\mathscr{H}}=(\mathscr{H}_F\mathscr{H}_{F,B})$ and the nuisance is defined by the following vector $\widetilde{\mathscr{X}}=(X_1^{\mathrm{T}},\mathscr{U}^{\mathrm{T}})^{\mathrm{T}}$. Let us apply the UBCP invariant test given by Eq. (7) to the state-space model (14). To do this, the vector Y, matrices H, M and P_H are replaced by $\mathscr{Y}$, $\widetilde{\mathscr{H}}$, $\mathscr{M}$ and $\mathscr{P}_{\widetilde{\mathscr{H}}}=I_{Nn}-\widetilde{\mathscr{H}}(\widetilde{\mathscr{H}}^{\mathrm{T}}\widetilde{\mathscr{H}})^{-}\widetilde{\mathscr{H}}^{\mathrm{T}}$, respectively. This test $\delta^*_{\widetilde{\mathscr{H}}}$ is UBCP over the following family of ellipsoids $\mathscr{S}_{\mathscr{W}_{\widetilde{\mathscr{H}}}\mathscr{M}}=\{S_c: 1/\sigma^2\|\mathscr{W}_{\widetilde{\mathscr{H}}}\mathscr{M}\Theta\|_2^2=c^2_{\widetilde{\mathscr{H}}}, c_{\widetilde{\mathscr{H}}}>0\}$, where matrix $\mathscr{W}_{\widetilde{\mathscr{H}}}$ satisfies the following conditions $\mathscr{W}_{\widetilde{\mathscr{H}}}\widetilde{\mathscr{H}}=0$, $\mathscr{W}_{\widetilde{\mathscr{H}}}\mathscr{W}^{\mathrm{T}}_{\widetilde{\mathscr{H}}}=I_{Nn-\mathrm{rank}\,\widetilde{\mathscr{H}}}$, and $\mathscr{W}^{\mathrm{T}}_{\widetilde{\mathscr{H}}}\mathscr{W}_{\widetilde{\mathscr{H}}}=\mathscr{P}_{\widetilde{\mathscr{H}}}$.

5.3. Comparison of the tests $\delta^*_{\mathscr{H}}$ and $\delta^*_{\widetilde{\mathscr{H}}}$

Let us consider the above mentioned tests $\delta^*_{\mathscr{H}}$ and $\delta^*_{\widetilde{\mathscr{H}}}$. The result of their comparison is summarized in the following lemma.

Lemma 3. *Let us assume that* $\Pr_0(\delta^*_{\mathscr{H}}\neq\mathscr{H}_0)=\Pr_0(\delta^*_{\widetilde{\mathscr{H}}}\neq\mathscr{H}_0)=\alpha$ $(0<\alpha<1)$. *Then the following inequality is satisfied for any* $\Theta\neq 0$: $\beta_{\delta^*_{\mathscr{H}}}(\Theta)\leqslant\beta_{\delta^*_{\widetilde{\mathscr{H}}}}(\Theta)$.

Proof. See Appendix D. As it is shown in Appendix D, the above-mentioned inequality becomes strict under some additional conditions. □

The proof of Lemma 3 is based on the fact that the column space of $\widetilde{\mathscr{H}}$ (see Eq. (14)) is contained in the column space of $\mathscr{H}$ (see Eq. (13)). 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 (11)–(12) performs at least not worse than the test based uniquely on the measurement Eq. (12). Nevertheless, as it follows from Appendix D, the column spaces of $\mathscr{H}$ and $\widetilde{\mathscr{H}}$ may coincide. It happens if rank $\widetilde{\mathscr{H}}=$ rank $\mathscr{H}$. In such a case, the state equation does not improve the quality of the UBCP invariant test: $\beta_{\delta^*_{\mathscr{H}}}(\Theta)=\beta_{\delta^*_{\widetilde{\mathscr{H}}}}(\Theta)$. A practical example of the above situation will be discussed in Section 6.3 to show the actual relevance of the approach which ignores the state equation.

6. Application: ground station based and receiver autonomous GPS integrity monitoring

Integrity monitoring, a major issue for the GPS in many safety-critical applications, requires that a navigation system *detects*, *isolates* faulty measurement sources (channels), and *removes* them from the navigation solution before they significantly contaminate the output. For some safety-critical navigation modes, landing, for instance, the GPS channels integrity monitoring is realized by using the measurements of ground based monitoring station at a known position $X_s=(x_s,y_s,z_s)^{\mathrm{T}}$ close to the airport. When a fault is detected, the corresponding information is transmitted via the integrity channel. The contribution of this example is twofold: first, it will be mathematically rigorously shown that the widely used LS residual statistics is in fact an optimal (UBCP invariant) solution without a receiver clock aiding; second, it will also be shown that the quality of fault detection can be improved by using clock aiding. The contribution of the second example (receiver autonomous integrity monitoring (RAIM) algorithm) is to show the actual relevance of Section 5. Here, the UBCP test is designed with and without taking into account the model of vehicle dynamics.

6.1. Ground station-based GPS integrity monitoring: algorithm obtained by ignoring the clock model

Let us assume that an inexpensive crystal oscillator is used as a frequency source in the station receiver. This means that the station receiver clock bias ϱ relative to the GPS time is an unknown (non-random) value, $\varrho\in\mathbb{R}$. Some additional motivation can be found in Basseville and Nikiforov (2002).

6.1.1. Measurement model

The scheme of ground station-based GPS integrity monitoring is depicted in Fig. 5. The GPS solution is based upon accurate measuring the distance (*range*) from n satellites with known locations $X_i=(x_i,y_i,z_i)^{\mathrm{T}}$, $i=1,\ldots,n$ to a user. The ground station model is: $y_i=r_i-d_i=c\varrho+\xi_i$, $i=1,2,\ldots,n$, where r_i is the *pseudorange* from the ith satellite to the user and $d_i=\|X_i-X_s\|_2$ is the known distance from the ith satellite to the user, $c\simeq 2.9979\times$

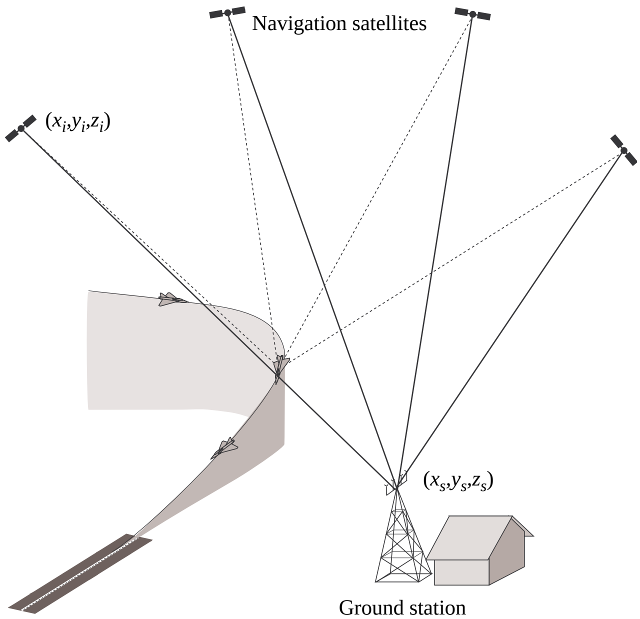

Fig. 5. Ground station-based GPS integrity monitoring.

10^8 m/s is the speed of light and ξ_i is an additive pseudorange error, $\xi = (\xi_1, \ldots, \xi_n)^{\mathrm{T}}$ is the vector of additive pseudorange errors at the ground station position. A fault is modeled by the vector θ of additional pseudorange biases: $Y_s = \mathbf{1}_n\mu + \xi\ (+\theta),\ \ \xi = (\xi_1, \ldots, \xi_n)^{\mathrm{T}} \sim \mathcal{N}(0, \sigma^2 I_n)$, where $Y_s = (y_1, y_2, \ldots, y_n)^{\mathrm{T}}$ is the vector of station measurements, $\mathbf{1}_n = (1, \ldots, 1)^{\mathrm{T}}$ and $\mu = c\varrho,\ \mu \in \mathbb{R}$, is the impact of the station clock bias measured in meter and considered as a nuisance parameter.

6.1.2. Fault detectability

Several GPS channels can be contaminated simultaneously. This leads to the situation when some combinations of individual channel faults are undetectable. To better understand the situation, let us represent the additional biases vector θ in the following manner[2] $\theta = \mathbf{1}_n\theta_\mu + W^{\mathrm{T}}\theta_W,\ \theta_\mu \in \mathbb{R},\ \theta_W \in \mathbb{R}^{n-1}$, where the nuisance rejection $((n-1)\times n)$ matrix $W : W\mathbf{1}_n = 0$ is composed of $n-1$ basis vectors which span the orthogonal complement of $R(\mathbf{1}_n)$ and it satisfies the conditions defined by Eq. (5): $WW^{\mathrm{T}} = I_{n-1}$ and $W^{\mathrm{T}}W = P_{\mathbf{1}_n} = (I_n - (1/n)\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})$. It follows from the above definition of θ that the sub-vector θ_μ is undetectable (the subspace of θ_μ coincides with the subspace of the nuisance parameter μ and, hence, θ_μ is masked by μ). Therefore, the only "detectable part" of $\theta$is represented by $W^{\mathrm{T}}\theta_W$ and we get $Y_s = \mathbf{1}_n\mu + \xi\ (+W^{\mathrm{T}}\theta_W),\ \ \xi = (\xi_1, \ldots, \xi_n)^{\mathrm{T}} \sim \mathcal{N}(0, \sigma^2 I_n)$. Let us analyze the impact of this undetectable part $\mathbf{1}_n\theta_\mu$ of the vector fault θ on the user's positioning. We consider a user (aircraft) at the positions $X_u = (x_u, y_u, z_u)^{\mathrm{T}}$. By linearizing the pseudorange equation with respect to the state vector $X = (x_u, y_u, z_u, \mu_u)^{\mathrm{T}} = (X_u^{\mathrm{T}}, \mu_u)^{\mathrm{T}}$ around the working point $X_0 = (X_{u_0}^{\mathrm{T}}, 0)^{\mathrm{T}}$, we get the linearized measurement equation of user with a fault

$$Y_u = R - D_0 \simeq H_0(X - X_0) + \xi\ (+\theta), \tag{15}$$

where $R = (r_1, \ldots, r_n)^{\mathrm{T}}$, $D_0 = (d_{1_0}, \ldots, d_{n_0})^{\mathrm{T}}$, $d_{i_0} = \|X_i - X_{u_0}\|_2$, $\xi = (\xi_1, \ldots, \xi_n)^{\mathrm{T}}$, $H_0 = \partial R/\partial X|_{X=X_0}$ is the Jacobian matrix of size $n \times 4$. As it follows from Eq. (15), a fault θ affecting the GPS channels implies an additional error $b = \mathbb{E}(\widehat{X} - X) = (H_0^{\mathrm{T}}H_0)^{-1}H_0^{\mathrm{T}}\theta$ in the vector $\widehat{X}$. Fortunately, the impact $b = \theta_\mu(H^{\mathrm{T}}H)^{-1}H^{\mathrm{T}}\mathbf{1}_n$ of such an undetectable bias $\mathbf{1}_n\theta_\mu$ on the first three components $\widehat{x}_u, \widehat{y}_u, \widehat{z}_u$ 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.

6.1.3. Fault detection algorithm

The problem consists in deciding between the null hypothesis $\mathscr{H}_0 : \{Y_s \sim \mathcal{N}(\mathbf{1}_n\mu, \sigma^2 I_n),\ \ \mu \in \mathbb{R}\}$ (no contaminated pseudoranges) and the alternative hypothesis $\mathscr{H}_1 : \{Y_s \sim \mathcal{N}(\mathbf{1}_n\mu + W^{\mathrm{T}}\theta_W, \sigma^2 I_n), \theta_W \neq 0, \mu \in \mathbb{R}\}$. The family $\mathscr{P} = \{\mathcal{N}(\mathbf{1}_n\mu + W^{\mathrm{T}}\theta_W, \sigma^2 I_n), \theta \in \mathbb{R}^r\}$ is invariant to the group $G = \{Y_s \mapsto g(Y_s) = Y_s + \mathbf{1}_n\mu\}$ and the induced group $\overline{G}$ is given by $\overline{g}(\theta) = \theta + \mathbf{1}_n x\ (x \in \mathbb{R})$. Let us assume that only one measurement vector Y_s is available to decide between two hypotheses. As it follows from Section 4, the test $\delta^*(Y_s)$ based on the following well-known LS residual statistics, widely used in fault detection: $\Lambda(Y_s) = 1/\sigma^2\sum_{i=1}^n (y_i - \overline{y})^2$,where $\overline{y} = 1/n\sum_{i=1}^n y_i$, is UBCP invariant over the family of surfaces[3]

$$S_c : \frac{1}{\sigma^2}\theta_W^{\mathrm{T}}WW^{\mathrm{T}}WW^{\mathrm{T}}\theta_W = \frac{1}{\sigma^2}\|\theta_W\|_2^2 = c^2.$$

6.2. Clock-aided ground station-based GPS integrity monitoring

The idea to use the clock model in order to improve the fault detection algorithm in the GPS integrity monitoring has been originally proposed and motivated in Misra, Muchnik, and Manganis (1995). It has been shown that by using a receiver clock short-term stability the performance of GPS integrity monitoring (a vital safety concern in civil aviation) can be seriously improved. Let us consider a clock-aided ground station based GPS monitoring to illustrate Lemma 3 and to show how the state equation improves the performance of the test. Here, the following very simplified clock model is used (Brown & Hwang, 1992): the long-term biases are modeled by an unknown low-frequency

[2] Instead of the total vector θ used in Basseville and Nikiforov (2002) and Nikiforov (2002, Chapter 2), only the "detectable" sub-vector θ_W is considered here. It is necessary because: (1) the matrix M of size $(n \times r)$ from Eq. (1) should be full column rank $(r < n)$; (2) the interpretation of the fault detection test performance is simpler.

[3] Unlike Basseville and Nikiforov (2002) and Nikiforov (2002, Chapter 2), where the total fault vector θ is used, the new family of surfaces $S_c : 1/\sigma^2\|\theta_W\|_2^2 = c^2$ is defined on the "detectable" subspace.

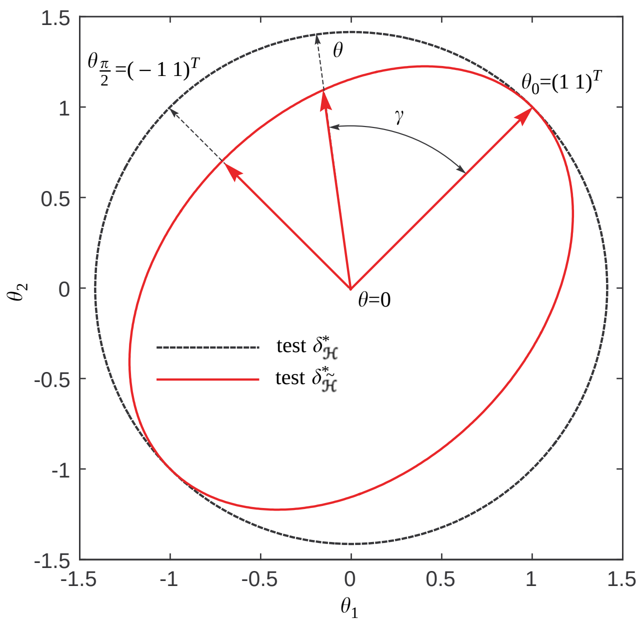

Fig. 6. Two equal power ellipses obtained by ignoring (dashed line) and by using the state equation (solid line).

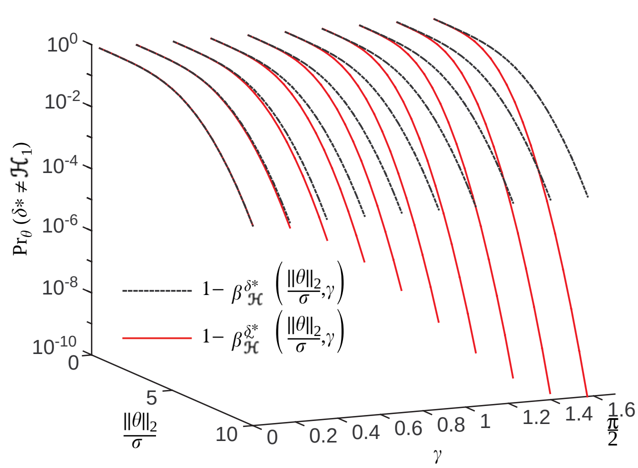

Fig. 7. Probability of non-detection by ignoring the state equation (dashed lines) and by using both state and measurement equations (solid lines).

function (due to the temperature changes, for example) but the short-term biases (1–10 s) are assumed to be practically constant. The choice of this simple model is especially motivated by the fact that the fault detection delay (time-to-alert) is bounded by 2–5 s for safety critical applications. Hence, the state-space model is given by: $\mu_k = \mu_{k-1}$, $Y_{s,k} = \mathbf{1}_n \mu_k + M\theta_k + \xi_k$, $\xi_k = (\xi_{k,1}, \ldots, \xi_{k,n})^{\mathrm{T}} \sim \mathcal{N}(0, \sigma^2 I_n)$, where $k = 1, \ldots, N$. Let us consider the following particular case: $n=2$, $N=2$, $M=(10)^{\mathrm{T}}$, to illustrate Lemma 3. The comparison between two fault detection tests (obtained by ignoring and by using the state equation) is presented in Figs. 6 and 7. Fig. 6 shows two equal power ellipses obtained by ignoring and by using the state equation: the ellipse (circle) $\|\mathscr{W}_{\mathscr{H}} \mathscr{M} \theta\|_2^2 = \sigma^2 c^2$, where $\sigma c = 1$, corresponding to the UBCP invariant test $\delta^*_{\mathscr{H}}$ obtained by ignoring the clock model, is given by $\theta_1^2 + \theta_2^2 = 2$. It is shown by a dashed line in Fig. 6. The ellipse $\|\mathscr{W}_{\widetilde{\mathscr{H}}} \mathscr{M} \theta\|_2^2 = \sigma^2 c^2$, where $\sigma c = 1$, corresponding to the UBCP invariant test $\delta^*_{\widetilde{\mathscr{H}}}$ obtained by using the clock model, is given by $\frac{3}{2}\,\theta_1^2 - \theta_1\theta_2 + \frac{3}{2}\,\theta_2^2 = 2$. It is shown by a solid line in Fig. 6. It is easy to see from Fig. 6 that the efficiency of clock-aided monitoring depends on the orientation (angle γ) of the vector $\theta = (\theta_1\ \theta_2)^{\mathrm{T}}$. If $\theta_1 = \theta_2$ ($\gamma = 0$ or π) then both tests $\delta^*_{\mathscr{H}}$ and $\delta^*_{\widetilde{\mathscr{H}}}$ are equivalent. If $\theta_1 = -\theta_2$ ($\gamma = \pi/2$ or $3\pi/2$) then the clock-aided test $\delta^*_{\widetilde{\mathscr{H}}}$ is more efficient than the test $\delta^*_{\mathscr{H}}$ obtained by ignoring the clock model. For this reason it is reasonable to compare the probability of non-detection $\Pr_\theta(\delta^* \neq \mathscr{H}_1) = 1 - \beta(\theta)$ as a function of two variables: the signal-to-noise ratio $\|\theta\|_2/\sigma$ and the angle γ. The functions $(\|\theta\|_2/\sigma, \gamma) \mapsto 1 - \beta(\|\theta\|_2/\sigma, \gamma)$ are shown in Fig. 7. Here, the probability of false alarm is chosen to be $\alpha = 10^{-3}$, the signal-to-noise ratio $\|\theta\|_2/\sigma$ varies between 0 and 10 and the angle γ varies between 0 and $\pi/2$. It is easy to see from Fig. 7 that the probability of non-detection $1 - \beta_{\delta^*_{\widetilde{\mathscr{H}}}}(\|\theta\|_2/\sigma, \gamma)$ (solid lines) of the clock-aided test $\delta^*_{\widetilde{\mathscr{H}}}$ is much smaller than the probability of non-detection $1 - \beta_{\delta^*_{\mathscr{H}}}(\|\theta\|_2/\sigma, \gamma)$ (dashed lines) of the test $\delta^*_{\mathscr{H}}$ obtained by ignoring the clock model when the angle γ is close to $\pi/2$. If the angle γ is close to 0, the both probabilities are comparable.

6.3. RAIM: the UBCP test with and without taking into account the model of vehicle dynamics

The goal of this subsection is to continue the illustration of the theoretical results of Section 5, especially to discuss again Lemma 3, and also to provide the readers with some additional explanation on the relevance of the UBCP invariant test given in Section 5.1 which is obtained by ignoring the state equation. To explain the subtlety of the usage of the dynamical models with unknown inputs, two different models of the vehicle dynamics will be used to design the GPS RAIM algorithms. Particularly, it will be shown that the state equation does not improve the quality of fault detector for one of these models.

Let us put together the linearized GPS measurement equation of the user (15) and the deterministic state equation that models the vehicle dynamics (the known term $H_0 X_0$ in (15) is omitted to simplify the notations):

$$X_k = \begin{pmatrix} aI_3 & 0 \\ 0 & 0 \end{pmatrix} X_{k-1} + BU_k,$$
$$X_k = (x_k, y_k, z_k, \mu_k)^{\mathrm{T}}, \tag{16}$$

$$Y_k = H_0 X_k + M\theta_k + \xi_k,$$
$$M = (1\ \ 0\ \ \cdots\ \ 0)^{\mathrm{T}}, \quad \theta_k \in \mathbb{R}, \tag{17}$$

where $a = 0.9$ is a coefficient and the vector U_k models the unknown input (nuisance parameter). Let us consider two

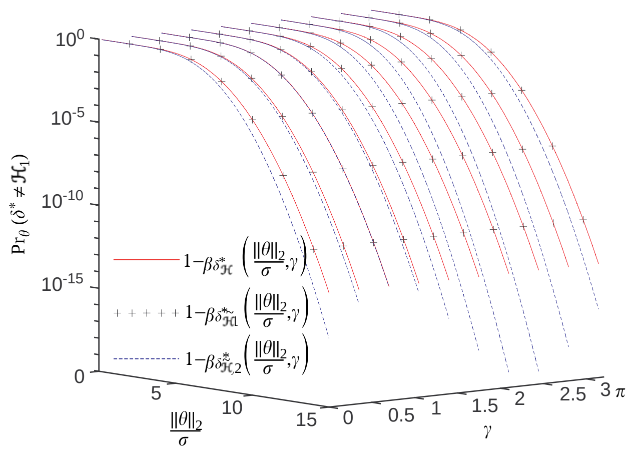

Fig. 8. Probability of non-detection by ignoring the vehicle dynamics (test $\delta^*_{\mathcal{H}}$, solid line), by using the model of dynamics with an unknown "motor thrust force" (test $\delta^*_{\widetilde{\mathcal{H}},1}$, "+" signs) and by using the model of dynamics with an unknown magnitude of the "motor thrust force" but known direction (test $\delta^*_{\widetilde{\mathcal{H}},2}$, dashed line).

particular case of the matrix B:

$$B_1 = I_4 \quad \text{with } U_k \in \mathbb{R}^4 \quad \text{and} \quad B_2 = \begin{pmatrix} 10 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{\mathrm{T}},$$

with $U_k \in \mathbb{R}^2$. In the case of $B = B_1$, the first three components of the vector U_k represent the "motor thrust force" that is responsible for the vehicle propulsion and the forth component "drives" the receiver clock error. In the case of $B = B_2$, the propulsion of the vehicle is realized by the first component of U_k (the first column of the matrix B_2 defines the direction of the "thrust vector") and the second component "drives" the receiver clock error. The above models of the vehicle dynamics differ by the level of a priori information: in contrast with the first model, the second model (with B_2) assumes that the direction of the "thrust vector" is known and only the magnitude of the thrust is unknown. Let us compare three following fault detectors:

- $\delta^*_{\mathcal{H}}$, obtained by ignoring the vehicle dynamics (state equation (16) is ignored);
- $\delta^*_{\widetilde{\mathcal{H}},1}$, obtained by using the vehicle dynamics with a minimal level of a priori information: the "motor thrust force" is unknown (state equation (16) is used with the matrix B_1);
- $\delta^*_{\widetilde{\mathcal{H}},2}$, obtained by using the vehicle dynamics with a known direction of the "thrust vector" and unknown thrust magnitude (state equation (16) is used with the matrix B_2).

The comparison between the above three fault detection algorithms is presented in Fig. 8. Eight satellites are visible at this moment ($n = 8$). Two consequent GPS observations (epochs) Y_1 and Y_2 ($N = 2$) are used to detect a fault $\theta = (\theta_1, \theta_2)^{\mathrm{T}} \in \mathbb{R}^2$ affected the first satellite channel (see the matrix M in (17)). The probability of false alarm is chosen to be $\alpha = 10^{-5}$. As in the previous case (see Section 6.2) it is reasonable to compare the probability of non-detection $\Pr_\theta(\delta^* \neq \mathcal{H}_1) = 1 - \beta(\theta)$ as a function of two variables: the signal-to-noise ratio $\|\theta\|_2/\sigma$ and the angle $\gamma = \arctan \theta_2/\theta_1$. The functions

$$\left(\frac{\|\theta\|_2}{\sigma}, \gamma\right) \mapsto 1 - \beta\left(\frac{\|\theta\|_2}{\sigma}, \gamma\right)$$

are shown in Fig. 8. The probabilities of non-detection for the tests $\delta^*_{\mathcal{H}}$ (ignoring the vehicle dynamics) and $\delta^*_{\widetilde{\mathcal{H}},1}$ (unknown "motor thrust force") are shown by a solid line and by "+" signs, respectively. It can be concluded from Fig. 8 that the probabilities of non-detection for both tests $\delta^*_{\mathcal{H}}$ and $\delta^*_{\widetilde{\mathcal{H}},1}$ are the same, in other words, the model of vehicle dynamics with the matrix B_1 does not improve the quality of the UBCP invariant test. Why it happens? As it follows from the proof of Lemma 3 (see Appendix D), the both column subspaces of $\mathcal{H}$ (Eq. (13)) and $\widetilde{\mathcal{H}}$ (Eq. (14)) may coincide: $R(\widetilde{\mathcal{H}}) = R(\mathcal{H})$. This is what happens in the case of vehicle dynamics with the matrix B_1. In fact, rank $\widetilde{\mathcal{H}}(B_1) = \text{rank } \mathcal{H} = 8$. In such a case, the power functions of both tests (with and without state equation) are equal. Intuitively, this result is easily understandable in the light of the vehicle dynamics model (see Eq. (16)). In fact, the matrix $B_1 = I_4$ does not reduce the dimension of the unknown input subspace and for this reason it is difficult to expect an improvement in the quality of test. On the contrary, the matrix B_2 reduces the dimension of the unknown input subspace and this leads immediately to a serious improvement in the quality of the test $\delta^*_{\widetilde{\mathcal{H}},2}$ obtained by using the vehicle dynamics with a known direction of the "thrust vector" and unknown thrust magnitude. It is confirmed by the probabilities of non-detection for the tests $\delta^*_{\widetilde{\mathcal{H}},2}$, which is shown in Fig. 8 (see a dashed line). To conclude this example let us stress the relevance of the approach obtained by ignoring the state equation in the case where the column spaces of $\mathcal{H}$ and $\widetilde{\mathcal{H}}$ coincide. The UBCP test, which uniquely uses the measurement equation is simpler and potentially more robust (because does not use the state equation coefficients) then the test based on the full-size state-space model.

7. Conclusion

The problem of fault detection in a linear gaussian model with nuisance parameters (or nuisance faults) has been addressed from the statistical point of view. First, the uniformly best constant power (UBCP) test is derived from solving an optimal hypotheses testing problem for the gaussian linear model by using the Wald's theory. The optimality of the proposed test is established in Theorem 1. Second, the invariant UBCP test is obtained for the linear gaussian model with nuisance parameters (or nuisance faults) by using Theorem 1 and the theory of invariant tests. Several critical issues concerning the design and the properties of fault detection algorithms (full-set parity vector against a subset one, "optimal" residual generation, etc.) have been addressed. The main results are given by Lemmas 1 and 2. Third, by using the proposed statistical tools, the linear state-space model with nuisance parameters defined by a deterministic state equation has been considered. It has been shown in Lemma 3 that the UBCP test designed by using both state and measurement equations performs at least not worse than the test based uniquely on the measurement equation.

Appendix A. Proof of Theorem 1

The idea of this proof is inspired by Wald (1943). It is necessary to show that the proposed test coincides with a specially designed Neyman–Person (N–P) test which realizes a constant power function over the family of surfaces $\mathscr{S}$. To apply this idea to model (2), let us assume the following hypotheses testing problem $\mathscr{H}_0 : \{\theta = 0\}$ and $\mathscr{H}_{1,c} : \{\theta : 1/\sigma^2\|M\theta\|_2^2 = c^2,\ c > 0\}$. The alternative hypothesis $\mathscr{H}_{1,c}$ corresponds to the ellipsoid S_c given by equation $\theta^{\mathrm{T}}M^{\mathrm{T}}M\theta = c^2\sigma^2$ from the family $\mathscr{S}_M$ (4). Let us define the following a priori density $p(\theta) = b = 1/A(S_c)$ of the parameter θ over the ellipsoid S_c, where $A(S_c)$ is the air of surface S_c. Hence, the density of the observation vector Y is given by the following surface integral $\int \cdots \int_{S_c} f_\theta(Y)p(\theta)\,\mathrm{d}S(\theta)$ when the hypothesis $\mathscr{H}_{1,c}$ is true. The N–P decision rule for testing the hypothesis $\mathscr{H}_0 = \{f_0(Y)\}$ against $\mathscr{H}_{1,c} = \{\int \cdots \int_{S_c} f_\theta(Y)p(\theta)\,\mathrm{d}S(\theta)\}$ is given by

$$\delta_{\mathrm{N-P}}(Y) = \begin{cases} \mathscr{H}_0 & \text{if } \Lambda_{\mathrm{N-P}}(Y) \leqslant d(c), \\ \mathscr{H}_{1,\ c} & \text{if } \Lambda_{\mathrm{N-P}}(Y) \geqslant d(c), \end{cases} \tag{A.1}$$

where

$$\Lambda_{\mathrm{N-P}}(Y) = \int \cdots \int_{S_c} \frac{f_\theta(Y)}{f_0(Y)}\, p(\theta)\,\mathrm{d}S(\theta)$$

is the likelihood ratio (LR) between $\mathscr{H}_0$ and $\mathscr{H}_{1,c}$,

$$f_\theta(Y) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left\{-\frac{1}{2\sigma^2}\|Y - M\theta\|_2^2\right\}$$

is the density of observations and $d(c)$ is the threshold so chosen that the probability of false alarm is equal to a prescribed level: $\Pr_0(\delta_{\mathrm{N-P}}(Y) \neq \mathscr{H}_0) = \alpha$. By developing f_θ, the integral $\Lambda_{\mathrm{N-P}}(Y)$ can be written as follows: $\Lambda_{\mathrm{N-P}}(Y) = b\int \cdots \int_{S_c} \exp\{1/2\sigma^2[2\theta^{\mathrm{T}}M^{\mathrm{T}}Y - \theta^{\mathrm{T}}(M^{\mathrm{T}}M)\theta]\}\,\mathrm{d}S(\theta)$. To prove the Theorem, it is necessary to show that the LR $\Lambda_{\mathrm{N-P}}(Y)$ of the N–P test is a non-decreasing function of the decision function $\Lambda(Y) = (1/\sigma^2)Y^{\mathrm{T}}M(M^{\mathrm{T}}M)^{-1}M^{\mathrm{T}}Y$ of the test $\delta^*(Y)$ (3) for any value of $c > 0$. The matrix $\Sigma = M^{\mathrm{T}}M$ is positive definite, hence it can be rewritten as follow: $\Sigma = RR^{\mathrm{T}}$ and $\Sigma^{-1} = R^{-\mathrm{T}}R^{-1}$. Let us define the vector $Z = R^{-1}M^{\mathrm{T}}Y$. The decision function $\Lambda(Y)$ of the test (3) can be represented as a function of Z, i.e. $\Lambda(Z) = 1/\sigma^2\|Z\|_2^2$ and the family of surfaces $\mathscr{S}_H$ can be also rewritten as $\mathscr{S}' = \{S'_c : \|\theta'\|_2^2 = c^2\sigma^2, c > 0\}$, where $\theta' = R^{\mathrm{T}}\theta$ and $\theta = R^{-\mathrm{T}}\theta'$ since the matrix R is non-singular. By using the change of variables $\theta' = R\theta$ ($\det R \neq 0$) and putting $M^{\mathrm{T}}Y = RZ$, we get:

$$\Lambda_{\mathrm{N-P}}(Z) = b' \int \cdots \int_{S'_c} \exp\left\{\frac{1}{\sigma^2}\theta'^{\mathrm{T}}Z\right\} \times \exp\left\{-\frac{1}{2\sigma^2}\|\theta'\|_2^2\right\} \mathrm{d}S'(\theta'),$$

where b' is a constant and the second term under the sign of integral $\exp\{-1/2\sigma^2\|\theta'\|_2^2\}$ is constant over the surface S'_c. As it follows from Wald (1943), the function $\|Z\|_2 \mapsto \Lambda_{\mathrm{N-P}}(\|Z\|_2)$ is non-decreasing for any $c > 0$. Therefore, it has been shown that the LR of the N–P test $\Lambda_{\mathrm{N-P}}(Y)$ is a non-decreasing function of $\Lambda(Y) = (1/\sigma^2)Y^{\mathrm{T}}M(M^{\mathrm{T}}M)^{-1}M^{\mathrm{T}}Y$. It proves that the test $\delta^*(Y)$ given by Eq. (3) has uniformly best constant power with respect to S_c and, hence, the test $\delta^*(Y)$ is UBCP over the family of ellipsoids $\mathscr{S}_M$ given by Eq. (4). □

Appendix B. Proof of Lemma 1

First of all, let us note that the statistics $\Lambda(Z)$ and $\Lambda(Z_1)$ given by Eqs. (7) and (9), respectively, follow χ^2 distributions with r degrees of freedom. These χ^2 distributions are central under $\mathscr{H}_0$. Therefore $h(\alpha) = h_1(\alpha)$. The tests $\delta^*(Z)$ and $\delta_1^*(Z)$ given by Eqs. (7) and (9) have constant power functions over the family of surfaces $\mathscr{S}_{WM}$ and $\mathscr{S}_{W_1M}$ given by Eqs. (8) and (10), respectively. The power functions $\beta_{\delta^*(Z)}(c^2)$ and $\beta_{\delta_1^*(Z)}(c_1^2)$ of the tests $\delta^*(Z)$ and $\delta_1^*(Z)$ are non-decreasing functions of their non-centrality parameters c^2 and c_1^2, respectively, given the equal thresholds $h(\alpha) = h_1(\alpha)$. Hence, to compare the power functions of the above tests it is sufficient to compare the non-centrality parameters $c^2(\theta)$ and $c_1^2(\theta)$ for the same fault vector $\theta \neq 0$. It follows immediately from paragraph 4.3.2 and an elementary matrix algebra that $c_1^2(\theta) - c^2(\theta) = -1/\sigma^2\|DM\theta\|_2^2 \leqslant 0$. Finally, we obtain $\beta_{\delta_1^*}(c_1^2(\theta)) \leqslant \beta_{\delta^*}(c^2(\theta))$ for any $\theta \neq 0$ and thus Lemma 1 is proved. □

Appendix C. Proof of Lemma 2

The goal of Lemma 2 is to prove that the utilization of the matrix $\widetilde{W} = AW : AWH = 0$, where A is a matrix of size $(n-q)\times(n-q)$ such that $\det A \neq 0$, instead of matrix W, does not change the power function of the test. Let $\delta^*(Z)$ be the test based on the statistics $Z = WY$ and given by Eq. (7) and $\delta^*(\widetilde{Z})$ be the same test based on the statistics $\widetilde{Z} = \widetilde{W}Y$. The proof is elementary: it is worth to note that the matrix A establishes a one-to-one linear transformation $\widetilde{Z} = AZ$ in $\mathscr{L}(\mathscr{Z}, \widetilde{\mathscr{Z}})$, where $Z \in \mathscr{Z}$ and $\widetilde{Z} \in \widetilde{\mathscr{Z}}$. Hence, the statistical properties of the tests based on the statistics Z and $\widetilde{Z}$ are exactly the same: $\beta_{\delta^*(Z)}(c^2) = \beta_{\delta^*(\widetilde{Z})}(c^2)$ for any $c^2 \geqslant 0$. Moreover, by applying $\widetilde{W}$ to $Y = HX + M\theta + \xi$, we get the new model $\widetilde{Z} = \widetilde{W}Y = \widetilde{W}M\theta + \widetilde{W}\xi$, where $\widetilde{W}\xi \sim \mathscr{N}(0, \sigma^2 V)$, with $V = AA^{\mathrm{T}}$. The direct calculation shows that the linear transformation $\widetilde{Z} = AZ$ does not change the

Fisher matrix $\mathscr{F}_{\widehat{\theta}}$, the statistics of the test: $\Lambda(Z) = \Lambda(\widetilde{Z})$ and the family of ellipsoids (8).

Appendix D. Proof of Lemma 3

The goal of Lemma 3 is to prove that the power function $\beta_{\delta^*_{\widetilde{\mathscr{H}}}}$ of the test $\delta^*_{\widetilde{\mathscr{H}}}$ obtained by using both (state (11) and measurement (12)) equations, is uniformly greater than or equal to the power function $\beta_{\delta^*_{\mathscr{H}}}$ of the test $\delta^*_{\mathscr{H}}$ obtained by ignoring the state equation. The proof is organized as follows: first, a relation between the matrices $\mathscr{H}$ (see Eq. (13)) and $\widetilde{\mathscr{H}} = (\mathscr{H}_F \mathscr{H}_{F,B})$ (14) will be established. Next, a relation between the parity spaces of $\mathscr{H}$ and $\widetilde{\mathscr{H}}$ will be shown. And finally, by using the same method as in Lemma 1, the power functions $\beta_{\delta^*_{\widetilde{\mathscr{H}}}}$ and $\beta_{\delta^*_{\mathscr{H}}}$ will be compared. Let us recall the regression model (13) obtained by ignoring the state equation, i.e. $\mathscr{Y} = \mathscr{H}\mathscr{X} + \mathscr{M}\Theta + \Xi$, where $\mathscr{Y} \in \mathbb{R}^{Nn}$, $\mathscr{X} \in \mathbb{R}^{Nm}$, $\Theta \in \mathbb{R}^{Nr}$ and the model (14) obtained by using both equations: $\mathscr{Y} = \widetilde{\mathscr{H}}\widetilde{\mathscr{X}} + \mathscr{M}\Theta + \Xi$, where $\widetilde{\mathscr{X}} \in \mathbb{R}^{m+(N-1)p}$. It follows from Eqs. (13) and (14) and an elementary matrix algebra (Strang, 1986) that $\widetilde{\mathscr{H}} = \mathscr{H}\mathscr{H}'$, where the $(Nm \times (m + (N-1)p))$ matrix $\mathscr{H}'$ is a function of blocks F and B. It also follows from Strang (1986) that $\operatorname{rank}\widetilde{\mathscr{H}} = \operatorname{rank}(\mathscr{H}\mathscr{H}') \leqslant \operatorname{rank}\mathscr{H}$ because the column space (range) $R(\widetilde{\mathscr{H}})$ of $\widetilde{\mathscr{H}}$ is contained in the range $R(\mathscr{H})$ of $\mathscr{H}$ (each column of $\widetilde{\mathscr{H}}$ is a combination of the columns of $\mathscr{H}$). There are two possible cases: (i) if $\operatorname{rank}\widetilde{\mathscr{H}} = \operatorname{rank}\mathscr{H}$ then both ranges coincide: $R(\widetilde{\mathscr{H}}) = R(\mathscr{H})$ and, hence, $R(\widetilde{\mathscr{H}})^{\perp} = R(\mathscr{H})^{\perp}$; (ii) if $\operatorname{rank}\widetilde{\mathscr{H}} < \operatorname{rank}\mathscr{H}$ then the orthogonal complement $R(\mathscr{H})^{\perp}$ of the range $R(\mathscr{H})$ is a subspace of the orthogonal complement $R(\widetilde{\mathscr{H}})^{\perp}$ of the range $R(\widetilde{\mathscr{H}}) : R(\mathscr{H})^{\perp} \subset R(\widetilde{\mathscr{H}})^{\perp}$. The first case is trivial: the power functions of both tests $\delta^*_{\mathscr{H}}$ and $\delta^*_{\widetilde{\mathscr{H}}}$ are equal: $\beta_{\delta^*_{\mathscr{H}}} = \beta_{\delta^*_{\widetilde{\mathscr{H}}}}$. To discuss the second case, let us assume that $s = \operatorname{rank}\mathscr{H} - \operatorname{rank}\widetilde{\mathscr{H}} > 0$. Hence, $\dim R(\mathscr{H})^{\perp} = Nn - \operatorname{rank}\mathscr{H} = Nn - Nq$ and $\dim R(\widetilde{\mathscr{H}})^{\perp} = Nn - Nq + s$. It follows from Strang (1986) that, since $R(\mathscr{H})^{\perp}$ is a subspace of $R(\mathscr{H})^{\perp}_1$, the vectors which span $R(\mathscr{H})^{\perp}$ can be extended to a basis of $R(\widetilde{\mathscr{H}})^{\perp}$ by adding s linearly independent vectors. Therefore, for any matrix $\mathscr{W}_{\mathscr{H}}$ ($\mathscr{W}_{\mathscr{H}}\mathscr{H} = 0$) there exists a matrix $\mathscr{D}$ composed of the lines which span the orthogonal complement of $R(\mathscr{H})^{\perp}$ in $R(\widetilde{\mathscr{H}})^{\perp}$: $\mathscr{W}_{\widetilde{\mathscr{H}}} = (\mathscr{W}^{\mathrm{T}}_{\mathscr{H}}\mathscr{D}^{\mathrm{T}})^{\mathrm{T}}$. It follows immediately from paragraph 4.3.2 and Lemma 1 applied with $\mathscr{W}_{\widetilde{\mathscr{H}}}$ and $\mathscr{W}_{\mathscr{H}}$ instead of W and W_1, respectively, that: $c^2_{\mathscr{H}}(\Theta) - c^2_{\widetilde{\mathscr{H}}}(\Theta) = -1/\sigma^2\|\mathscr{D}\mathscr{M}\Theta\|_2^2 \leqslant 0$ and, hence, $\beta_{\delta^*_{\mathscr{H}}}(\Theta) \leqslant \beta_{\delta^*_{\widetilde{\mathscr{H}}}}(\Theta)$ for any $\Theta \neq 0$. Let us additionally assume that $\operatorname{rank}(\mathscr{D}\mathscr{M}) = \dim(\Theta)$. In this case $-1/\sigma^2\|\mathscr{D}\mathscr{M}\Theta\|_2^2 < 0$ and, hence, $\beta_{\delta^*_{\mathscr{H}}}(\Theta) < \beta_{\delta^*_{\widetilde{\mathscr{H}}}}(\Theta)$ for any $\Theta \neq 0$.

References

Basseville, M. (1997). Information criteria for residual generation and fault detection and isolation. *Automatica*, *33*(5), 783–803.

Basseville, M., & Nikiforov, I. V. (2002). Fault isolation for diagnosis: Nuisance rejection and multiple hypothesis testing. *Annual Reviews in Control*, *26*(2), 189–202.

Borovkov, A. A. (1998). *Mathematical statistics*. Amsterdam: Gordon and Breach Sciences Publishers.

Brown, R. G., & Hwang, P. Y. C. (1992). *Introduction to random signals and applied Kalman filtering*. (2nd ed), New York: Wiley.

Chen, J., & Patton, R. J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Boston, MA: Kluwer Academic Publishers.

Chen, R. H., Mingori, D. L., & Speyer, J. L. (2003). Optimal stochastic fault detection filter. *Automatica*, *39*, 377–390.

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

Gertler, J. J. (1998). *Fault detection and diagnosis in engineering systems*. New York: Marcel Dekker.

A. John Volpe National Transportation System Center (2001). US Department of Transportation. *Vulnerability assessment of the transportation infrastructure relying on the GPS*. Technical report, August 2001. [Online]. Available: http://www.navcen.uscg.gov/pubs/default.htm.

Koch, K.-R. (1999). *Parameter estimation and hypothesis testing in linear models*. Berlin, Heidelberg, New York: Springer.

Lehmann, E. L. (1986). *Testing statistical hypotheses*. London: Chapman & Hall.

Mironovski, L. A. (1980). Functional diagnosis of dynamic systems—a survey. *Automation Remote Control*, *41*, 1122–1143.

Misra, P., Muchnik, R., & Manganis, B. (1995). A general RAIM algorithm based on receiver clock. In *Proceedings of the eighth international meeting of the satellite division of the institute of navigation* (pp. 1941–1949), Palm Springs, CA.

Nikiforov, I. (2002). Eléments de théorie de la décision statistique II: compléments. in: R. Lengellé (Ed.), *Décision et reconnaissance des formes en signal* (pp. 43–84). Hermes Science Publications, (Traité IC2).

Patton, R. J., & Chen, J. (1991). A review of parity space approaches to fault diagnosis. In *Preprints of IFAC/IMACS symposium SAFEPROCESS'91*, vol. 1 (pp. 239–255), Baden-Baden, September 10–13.

Patton, R. J., Frank, P. M., & Clark, R. N. (Eds), (1989). *Fault diagnosis in dynamic systems*. Englewood Cliffs, NJ: Prentice-Hall.

Rougée, A., Basseville, M., Benveniste, A., & Moustakides, G. (1987). Optimum robust detection of changes in the ar part of a multivariable arma process. *IEEE Transactions on Automatic Control*, *32*(12), 1116–1120.

RTCA/DO-229A (1998). *Minimum operational performance standards for global positioning systems/wide area augmentation system airborne equipment*. Washington, DC, USA: RTCA, Inc.

Scharf, L. L., & Friedlander, B. (1994). Matched subspace detectors. *IEEE Transactions on Signal Processing*, *42*(8), 2146–2157.

Staroswiecki, M, (2001). Redondance analytique. In *Automatique et statistiques pour le diagnostic* (pp. 43–68). Hermes Science Europe.

Strang, G. (1986). *Linear algebra and its applications*. Harcourt Brace and Company International Edition.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of American Mathematical Society*, *54*, 426–482.

Mitra Fouladirad received the M.S. degree in mathematical statistics from the University of Paris VI in 2001. Actually she is a Ph.D. student in the University of Technology of Troyes. Her scientific interest is statistical decision theory.

Igor V. Nikiforov received the M.S. degree in automatic control from the Moscow Physical—Technical Institute in 1974, and Ph.D. in automatic control from the Institute of Control Sciences, Moscow, in 1981. He joined the University of Technology of Troyes in 1995, France, where he is the professor and head of the Institute of Computer Sciences and Engineering of Troyes, FRE CNRS 2732. His scientific interests include statistical decision theory, fault detection, signal processing, navigation.