Testing the Mahalanobis distance between a random signal with unknown distribution and a known deterministic model in additive and independent standard Gaussian noise: the random distortion testing problem.

Dominique Pastor (Télécom Bretagne, Labsticc)
Testing the Mahalanobis distance between a random signal with unknown distribution and a known deterministic model in additive and independent standard Gaussian noise: the random distortion testing problem.

Rapport de Recherche
Institut Télécom / Télécom Bretagne

Dominique Pastor
Institut Télécom; Télécom Bretagne
UMR CNRS 6285 Lab-STICC,
Université européenne de Bretagne
Technopôle Brest Iroise-CS 83818,
29238 Brest Cedex, France

Email: dominique.pastor@telecom-bretagne.eu

April 3, 2012
Abstract

This paper addresses the problem of deciding whether the Mahalanobis distance between a random signal $\Theta$ and a known deterministic model $\theta_0$ exceeds some given non-negative real number or not, when the $\Theta$ has unknown probability distribution and is observed in additive independent Gaussian noise with positive definite covariance matrix. A new optimality criterion based on the invariance of both the problem and the noise distribution is introduced, via conditional notions of power and size. The tests optimal with respect to this criterion are given. The results established in this paper extend those of Wald's for testing the mean of a Gaussian distribution, which is the particular case where $\tau = 0$ and the signal is deterministic unknown. Application to the detection of random signals in additive independent Gaussian noise is also addressed.

Keywords

Event testing, hypothesis testing, invariance, Mahalanobis norm, random distortion, test with uniformly best constant power, test with uniformly best constant conditional power.
# Contents

1 Introduction 4

2 Outline of main results 5  
   2.1 Random distortion testing (RDT) in Gaussian noise . . . . . . . . . . . . 5  
   2.2 Application to deterministic distortion testing of a Gaussian distribution 5  
   2.3 Application to signal detection . . . . . . . . . . . . . . . . . . . . . . 6

3 Preliminary material 6  
   3.1 Tests, power function and invariance . . . . . . . . . . . . . . . . . . . . 6  
   3.2 Groups $\mathcal{G}$ and $\overline{\mathcal{G}}$ . . . . . . . . . . . . . . . . . . . 7  
   3.3 Thresholding tests on the Mahalanobis distance . . . . . . . . . . . . . . 7  
   3.4 Map $\mathcal{R}$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

4 Random distortion testing in additive independent Gaussian noise 8  
   4.1 Mathematical statement . . . . . . . . . . . . . . . . . . . . . . . . . . . 8  
   4.2 Invariance of the RDT problem . . . . . . . . . . . . . . . . . . . . . . . 9  
   4.3 Conditional power with respect to ellipsoids . . . . . . . . . . . . . . . 10  
   4.4 Tests with uniformly best constant conditional power on family $\mathcal{F}$ . . 10

5 Application to deterministic distortion testing (DDT) of a Gaussian distribution 12  
   5.1 problem statement . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12

6 Application to signal detection 14  
   6.1 Detection in noise with known covariance matrix . . . . . . . . . . . . . 15  
   6.2 Detection in noise with unknown standard deviation . . . . . . . . . . . 17

7 Conclusion and perspectives 20

Appendix I: Proof of Lemma 1 21
Appendix II: Proof of Lemma 2 22
Appendix III: No UMP test for RDT in the random case 22
Appendix IV: Proof of Proposition 1 23
Appendix V: Proof of Theorem 1 24
1 Introduction

In many signal processing applications (radar, sonar, remote sensing, telecommunication, communication electronic support, among others), the observation captured by a sensor is assumed to be a real random vector $Y$, whose probability distribution involves some $d$-dimensional parameter of interest or signal $\theta$ and possibly other parameters, which are the so-called nuisance parameters when unknown. A fundamental problem is to decide on whether $Y$ is a corrupted version of some specified deterministic model $\theta_0$ for $\theta$ or not. This problem is usually posed as the testing of the simple hypothesis $\theta = \theta_0$ against the composite alternative $\theta \neq \theta_0$. No Uniformly Most Powerful (UMP) test exists for this two-sided vector parameter testing problem and sub-optimal tests have then been proposed. For instance, if $Y$ is assumed to be Gaussian distributed with mean $\theta$ and known definitive positive covariance matrix $C$ — we hereafter write $Y \sim \mathcal{N}(\theta, C)$ with $C > 0$ —, sub-optimal tests are given by Wald in [1, Section 6, Proposition III, p. 450] to optimize the alternative criterion [1, Definition III, Section 6, p. 450]. These tests directly relate to the invariance of the problem because they have constant power on the orbits of the group that leaves the problem invariant. In the general case where the distribution of $Y$ involves possibly nuisance parameters, the so-called holy trinity — that is, the generalized likelihood ratio test (GLRT) [2], the Rao score test [3, Sec. 3, p. 53] and the Wald test [4, Sec. 13, p. 478, Theorem VIII] — are of great interest when a sufficiently large number of independent observations is available to benefit from the asymptotic properties of maximum likelihood estimates and Fisher’s information matrix. In particular, these tests satisfy invariance properties and exhibit similar performance [5–7], asymptotically. In absence of nuisance parameters, they coincide when the distribution of $Y$ is exponential [8].

When the signal is a distorted version of $\theta_0$ because of unknown environmental fluctuations regardless of noise, small deviations from the model can be of poor interest for the user, who may want to detect big enough ones only. Therefore, testing that the signal is exactly $\theta_0$ can be too severe, and even paradoxical because of physics. It should therefore be more reasonable to allow for possible distortions around $\theta_0$, by testing the composite hypothesis that the signal lies in a neighborhood of $\theta_0$. For an unknown deterministic signal, the Wald and Rao tests can cope with distortions by choosing a neighborhood of $\theta_0$ satisfying the conditions of [4, Sec. 13, p. 478, Theorem VIII] or [3, Sec. 3, p. 53]. For small deviations, a robust test [9, III.E.2] can also be derived from an uncertainty model. As in [10], the present work is motivated by situations where the degree of uncertainty or the number of unknown parameters is so great that standard likelihood theory, including the holy trinity, may not apply to allow for possible distortions of the model. In this respect, we suppose that the signal is a $d$-dimensional real random vector $\Theta$ with unknown distribution and that the distribution of $Y$ conditioned to $\Theta = \theta$ belongs to some family of distributions parameterized by $\theta$. Postponing to further work the general case with possibly nuisance parameters, we hereafter assume that the distribution of $Y$ conditioned to $\Theta = \theta$ is Gaussian, centered with known positive definite covariance matrix. This additive and signal-independent Gaussian noise model is reasonable in many applications and amounts to assuming that $Y = \Theta + X$, where signal $\Theta$ is random with unknown distribution and independent of $X \sim \mathcal{N}(0, C)$ with $C > 0$. By compensating the variation induced by the noise covariance matrix, the Mahalanobis norm [11] of $\Theta - \theta_0$ is relevant to evaluate how far $\Theta$ deviates from its model $\theta_0$. The sequel then addresses the problem of testing whether the Mahalanobis norm of $\Theta - \theta_0$ ex-
ceeds some tolerance or not. For simplicity sake, we henceforth call this problem the random distortion testing (RDT) problem. The next section outlines the main contributions of the paper. These contributions are then detailed in sections 4, 5 and 6, respectively, after introducing some material in Section 3. Perspectives are discussed in Section 7.

2 Outline of main results

2.1 Random distortion testing (RDT) in Gaussian noise

In this problem, the observation is assumed to be $Y = \Theta + X$, where $\Theta$ and $X$ are independent $d$-dimensional real random vectors and $X \sim N(0, C)$ with $C > 0$. No assumption is made on the probability distribution of $\Theta$. RDT in Gaussian noise $N(0, C)$ is thus the problem of testing whether the Mahalanobis of $\Theta - \theta_0$ exceeds some specified non-negative real number $\tau$ or not, on the basis of observation $Y$. The value $\tau$ is called the tolerance of the RDT problem. It can be specified by the user himself on the basis of his experience and know-how with respect to a given environment or context. To the best of our knowledge, RDT is addressed in this form for the first time. In contrast to standard approaches in statistical inference, RDT is not a binary hypothesis testing problem in the usual sense, since the decision concerns random events involving a random vector with unknown distribution. Therefore, RDT cannot be tackled via usual hypothesis testing and likelihood theory. The crux of the approach followed below is then the invariance of the RDT problem under the action of a specific group $G$. The orbits of $G$ are ellipsoids $\Upsilon_\rho = \{y \in \mathbb{R}^d : \|y - \theta_0\| = \rho\}$ where $\rho$ is any non-negative real number and $\|\cdot\|$ denotes the Mahalanobis norm in $\mathbb{R}^d$. Throughout, this family of orbits is denoted by $F$. Since $\Theta$ is random with unknown distribution, the RDT invariance cannot be exploited through direct application of the standard invariance principle [12–15]. A suitable theoretical framework is therefore proposed to cope with this invariance. The new criterion deriving from this framework involves optimizing a conditional notion of power, defined with respect to the elements of $F$, with a constraint on a notion of size tailored to RDT. Therefore, the theoretical framework applies to deterministic distortion testing (DDT) of a Gaussian distribution, that is, the problem of testing whether the Mahalanobis distance between the mean $\theta$ of $Y \sim N(\theta, C)$ with $C > 0$ and $\theta_0$ remains within a given tolerance $\tau$ or not. It then turns out that, for DDT of a Gaussian distribution, the UBCCP tests of Theorem 1 have uniformly best constant power (UBCP), in a somewhat extended sense with respect to [1, Definition III, Section 6, p. 450]. These tests are thus UMP among the tests with $F$-invariant power function and thus, uniformly most powerful invariant with respect to group $G$ ($G$-UMPI). Testing the mean of a Gaussian distribution being the DDT problem with null tolerance, Wald’s result [1, Section 6, Proposition III, p. 450] derives from ours.
2.3 Application to signal detection

The detection of random signals with unknown distributions in independent and additive white Gaussian noise (AWGN) is cast in the RDT framework. Theorem 1 then applies and the detection performance of UBCCP tests is discussed. The detection problem in case of an estimated noise standard distortion is considered as well because it can also be regarded as an RDT problem. When the noise standard deviation is unknown, the results in RDT are adapted via an estimate-and-plug-in detector [7] based on auxiliary data of noise alone. The use of a positive tolerance partly compensates the performance loss induced by the noise standard deviation estimation.

3 Preliminary material

All the random vectors and variables encountered henceforth are assumed to be defined on the same probability space \((\Omega, \mathcal{B}, P)\). As usual, we write (a-s) for almost surely. The set of all \(d\)-dimensional real random vectors defined on \((\Omega, \mathcal{B})\) is hereafter denoted by \(\mathcal{M}(\Omega, \mathbb{R}^d)\). Given any \(Z \in \mathcal{M}(\Omega, \mathbb{R}^d)\) (resp. any random variable), \(PZ^{-1}\) stands for the probability distribution of \(Z\), that is the probability measure defined for any Borel subset \(B\) of \(\mathbb{R}^d\) (resp. \(\mathbb{R}\)) by \(PZ^{-1}(B) = P[Z \in B]\).

Throughout the paper, we are given the noise positive definite \(d \times d\) covariance matrix \(C\). The Mahalanobis norm in \(\mathbb{R}^d\) defined with respect to \(C\) is then denoted by \(\| \cdot \|\). We thus have \(\| y \| = \sqrt{y^T C^{-1} y}\) for any \(y \in \mathbb{R}^d\), where \(A^T\) henceforth stands for the transpose of any matrix or vector \(A\).

3.1 Tests, power function and invariance

In the sequel, a test is any measurable map of \(\mathbb{R}^d\) into \([0, 1]\). Randomized tests [12, Section 3.1, p. 58] are hereafter useless because of the absolute continuity of the observation distribution with respect to Lebesgue’s measure in \(\mathbb{R}^d\). Given a family \(\mathcal{T} = \{P_\theta : \theta \in \mathbb{R}^d\}\) of distributions parameterized by \(\theta \in \mathbb{R}^d\), we recall that the power function of any test \(T\) is the map that assigns to every given \(\theta \in \mathbb{R}^d\) the value

\[
\beta_\theta(T) = P[T(Y) = 1],
\]

when \(Y\) is any random vector or variable whose probability distribution is \(P_\theta\). With the notation introduced above, we now recall some basics on invariance.

Let \(G\) be some group of measurable transforms of \(\mathbb{R}^d\) into itself. Test \(T\) is then said to be \(G\)-invariant if \(T \circ g = T\) for any \(g \in G\) and to have \(G\)-invariant power function if \(\beta_\theta(T \circ g) = \beta_\theta(T)\) for any \(g \in G\) and any \(\theta \in \mathbb{R}^d\). Therefore, a \(G\)-invariant test has necessarily \(G\)-invariant power function. However, the converse is not true [12, Chapter 6, pp. 227 – 228]. In the sequel, the following remark will prove useful as a particular case of [13, Sec. 47, Chapter III, p. 281]. Suppose that \(\theta \in \mathbb{R}^d \rightarrow P_\theta \in \mathcal{T}\) is a one-to-one mapping and that, given any \(\theta \in \mathbb{R}^d\), \(g(Y)\) has probability distribution \(P_{g(\theta)}\) for any \(Y\) with distribution \(P_\theta\). It then turns out that \(T\) has \(G\)-invariant power function if and only if \(T\) has constant power on every orbit of \(G\), that is, for any orbit of \(G\) and any elements \(\theta\) and \(\theta'\) of this orbit, \(\beta_\theta(T) = \beta_{\theta'}(T)\). This is the situation encountered in the standard Gaussian model considered below. Indeed, we henceforth cope with the family of Gaussian distributions \(\mathcal{N}(\theta, C)\) parameterized by \(\theta \in \mathbb{R}^d\) with known \(C > 0\). The power function thus handled in the sequel is given by (1)
with $Y \sim N(\theta, C)$. In Section 3.2 and the rest of the paper, we then exhibit a group $\mathcal{G}$ such that the tests with $\mathcal{G}$-invariant power function are exactly those with constant power on every orbit of this group.

### 3.2 Groups $\mathcal{G}$ and $\mathcal{G}$

Consider any eigenvector decomposition $C = U\Delta U^T$ of the noise covariance matrix $C > 0$, where $\Delta = \text{diag}(\xi_1, \xi_2, \ldots, \xi_d)$ is a diagonal matrix whose diagonal elements $\xi_1, \xi_2, \ldots, \xi_d$ are the eigenvalues of $C$ and $U$ is an element of the orthogonal group $O_d$ of $\mathbb{R}^d$. Set $\Delta^{-1/2} = \text{diag}(\xi_1^{-1/2}, \xi_2^{-1/2}, \ldots, \xi_d^{-1/2})$ and $\Phi = \Delta^{-1/2}U^T$. For any $y \in \mathbb{R}^d$, we then have

$$\|y\| = \|\Phi y\|_2,$$

where $\|y\|_2 = \sqrt{y^Ty}$ henceforth stands the standard Euclidean norm of $y$. Group $\mathcal{G} = \{\Phi^{-1}R\Phi : R \in O_d\}$ is then the group of isometries of the normed space $(\mathbb{R}^d, \|\cdot\|)$. This group leaves invariant the distribution of any $X \sim N(0, C)$ and does not depend on the orthogonal and diagonal matrices $U$ and $\Delta$ involved in the eigenvector decomposition of $C$. To each $g \in \mathcal{G}$, we assign the affine isometry $\mathcal{G}$ defined for every $y \in \mathbb{R}^d$ by $\mathcal{G}(y) = g(y - \theta_0) + \theta_0$. We then define group $\mathcal{S} = \{\mathcal{G} : g \in \mathcal{G}\}$, which is induced by $\mathcal{G}$ in that $\mathcal{G} \circ \mathcal{H} = \mathcal{G} \circ \mathcal{H}$ for any $g, h \in \mathcal{S}$. The orbits of $\mathcal{S}$ are exactly the elements of $\mathcal{S} = \{Y_{\rho} : \rho \geq 0\}$ of ellipsoids $Y_{\rho} = \{y \in \mathbb{R}^d : \|y - \theta_0\| = \rho\}$. The reader will then verify that $\mathcal{G}(Y) \sim N(\mathcal{G}(\theta), C)$ for any $Y \sim N(\theta, C)$ so that the tests with constant power on every $Y_{\rho} \in \mathcal{S}$ are those with $\mathcal{S}$-invariant power function. The invariance of the RDT problem will be characterized thanks to $\mathcal{G}$ and $\mathcal{S}$.

### 3.3 Thresholding tests on the Mahalanobis distance

Given any $\eta \geq 0$, a thresholding test with threshold height $\eta \geq 0$ on the Mahalanobis distance to $\theta_0 \in \mathbb{R}^d$ is any test $\mathcal{T}_\eta$ such that, for any $y \in \mathbb{R}^d$,

$$\mathcal{T}_\eta(y) = \begin{cases} 1 & \text{if } \|y - \theta_0\| > \eta \\ 0 & \text{if } \|y - \theta_0\| < \eta. \end{cases}$$

(3)

The handling of equality in (3) plays no role in the sequel because of the absolute continuity of the probability distribution of the observation with respect to Lebesgue’s measure. Any of these tests has constant power on every $Y_{\rho} \in \mathcal{S}$ or, equivalently, $\mathcal{S}$-invariant power function. In fact, according to [1, Section 6, Proposition III, p. 450], any thresholding test with threshold height $\eta \geq 0$ on the Mahalanobis distance to $\theta_0$ has uniformly best constant power [1, Definition III, Section 6, p. 450] — and we say that this test is UBCP — on $\mathcal{S}$ for testing the mean of a Gaussian distribution, that is, testing $\theta = \theta_0$ against $\theta \neq \theta_0$, when the observation has distribution $N(\theta, C)$.

### 3.4 Map $\mathcal{R}$

For any $\rho \in [0, \infty)$, $\mathcal{R}(\rho, \cdot)$ hereafter stands for the cumulative distribution function of the square root of any random variable that follows the non-central $\chi^2$ distribution with $d$ degrees of freedom and non-central parameter $\rho^2$. Given any $Y \sim N(\theta, C)$, $\Phi Y \sim N(\Phi \theta, I_d)$ where $\Phi$ is defined as in Section 3.2 and $I_d$ henceforth stands for the $d \times d$ identity matrix. It then follows from (2) that

$$P[\|Y\| \leq \eta] = \mathcal{R}(\|\theta\|, \eta).$$

(4)
Given any \( \rho \in [0, \infty) \), \( \mathcal{R}(\rho, \cdot) \) is strictly increasing and continuous and, thus, a one-to-one mapping of \([0, \infty)\) into \([0, 1)\). No analytical expression of \( \mathcal{R} \) is needed. On the other hand, the following properties of \( \mathcal{R} \) lead to some specified non-negative real number. Given \( \omega \in \Omega \), we want to know whether \( \|\Theta - \theta_0(\omega)\| \leq \tau \) or not, when we are given \( Y(\omega) \) and the probability distribution of \( \Theta \) is unknown. By analogy with standard terminology in statistical inference, we say that the testing of the event \( \|\Theta - \theta_0\| \leq \tau \) is called the probability of those \( \rho \)-dimensional real random vectors \( \Theta \) such that \( P\left[\|\Theta - \theta_0\| \leq \tau \right] \in (0, 1) \). We hereafter focus on RDT from above, which is summarized as follows for further reference in the sequel:

\[
\begin{align*}
\text{Observation:} & \quad Y = \Theta + X, \\
\text{Tested or null event:} & \quad \Omega_0 = \left\{ \|\Theta - \theta_0\| \leq \tau \right\}, P(\Omega_0) \in (0, 1) \\
\text{Alternative event:} & \quad \Omega_1 = \left\{ \|\Theta - \theta_0\| > \tau \right\}.
\end{align*}
\]

At the end of the section, the results in RDT from above will then be transposed to RDT from below.

The testing of \( \Omega_0 \) amounts to choosing some map of \( \Omega \) into \([0, 1]\) such that, for every \( \omega \in \Omega \), the value returned by this map is the decision on whether \( \|\Theta - \theta_0(\omega)\| \leq \tau \) or not. Of course, there are infinitely many possible choices for maps of \( \Omega \) into \([0, 1]\). In the sequel, we restrict our attention to the rather natural class of those composite maps \( \mathcal{T} \) or \( \mathcal{T}(Y) \) where \( \mathcal{T} \) is any test. The performance of a given test \( \mathcal{T} \) is then assessed for the RDT problem (5) via the following two quantities, whose

**Lemma 1** Given any \( \eta \in (0, \infty) \), the map \( \mathcal{R}(\cdot, \eta) \) is strictly decreasing.

**Proof:** See appendix I.

**Lemma 2**

(i) given \( \gamma \in (0, 1) \) and \( \rho \in [0, \infty) \), there exists a unique solution \( \lambda_{\gamma}(\rho) \in [0, \infty) \) in \( \eta \) to \( 1 - \mathcal{R}(\rho, \eta) = \gamma \).

(ii) given \( \gamma \in (0, 1) \), \( \lambda_{\gamma} \) is a strictly increasing and everywhere continuous map of \([0, \infty)\) into \([0, \infty)\).

(iii) given \( \rho \in [0, \infty) \), the map \( \gamma \in (0, 1) \mapsto \lambda_{\gamma}(\rho) \in [0, \infty) \) is strictly decreasing and continuous everywhere.

**Proof:** See appendix II.
The size of $\mathcal{I}$ for the RDT problem (5) is defined by

$$a^{\Theta}(\mathcal{I}) = \sup_{\Theta \in \vartheta_r} \mathbb{P}[\mathcal{I}(Y) = 1 \mid \Omega_0].$$

(6)

Test $\mathcal{I}$ is then said to have level (resp. size) $\gamma \in [0, 1]$ for testing $\Omega_0$ if $a^{\Theta}(\mathcal{I}) \leq \gamma$ (resp. $a^{\Theta}(\mathcal{I}) = \gamma$). Given $\gamma \in [0, 1]$, $\mathcal{X}^{\Theta \beta}_{\mathcal{I}}$ will henceforth denote the class of those tests $\mathcal{I}$ such that $a^{\Theta}(\mathcal{I}) \leq \gamma$. Second, the power of $\mathcal{I}$ for the RDT problem (5) is defined as the map that assigns to each $\Theta \in \vartheta_r$ the conditional probability

$$\beta^{\Theta}(\mathcal{I}) = \mathbb{P}[\mathcal{I}(Y) = 1 \mid \Omega_1].$$

(7)

For the RDT problem (5), a test $\mathcal{I}$ with level $\gamma$ is said to be unbiased if, for any $\Theta \in \vartheta_r$, $\beta^{\Theta}(\mathcal{I}) \geq \gamma$ and to be more powerful than another test $\mathcal{I}'$ with same level if $\beta^{\Theta}(\mathcal{I}) \geq \beta^{\Theta}(\mathcal{I}')$ for every $\Theta \in \vartheta_r$. The notions of size and power defined for the RDT problem (5) relate to the power function defined by (1) since we have the following lemma, whose proof is left to the reader.

**Lemma 3** Let $\Theta$ and $\Theta'$ be any elements of $\mathcal{X}^d$ such that $\|\Theta - \Theta_0\| \leq \tau$ and $\|\Theta' - \Theta_0\| > \tau$.

For any Bernoulli random variable $\epsilon$ independent of $X$ and valued in $[0, 1]$ with $P[\epsilon = 1] \in (0, 1)$, the random vector $\Theta = (1 - \epsilon)\Theta + \epsilon\Theta'$ is an element of $\vartheta_r$, $\beta^{\Theta}(\mathcal{I}) = \beta^{\Theta'}(\mathcal{I})$ and $\beta(\mathcal{I}) \leq a^{\Theta}(\mathcal{I})$.

Our purpose is to pinpoint tests in $\mathcal{X}^{\Theta \beta}_{\mathcal{I}}$ whose power is optimal, with respect to a certain criterion, for testing $\Omega_0$ in (5). This criterion must necessarily concern a restricted family of tests. Indeed, as shown in Appendix III, there is no UMP test with level $\gamma \in (0, 1)$ for the RDT problem (5). By UMP test with level $\gamma$, we mean some $\mathcal{I} \in \mathcal{X}^{\Theta \beta}_{\mathcal{I}}$ such that $\beta^{\Theta}(\mathcal{I}) \geq \beta^{\Theta}(\mathcal{I}')$ for any $\Theta \in \vartheta_r$ and any $\mathcal{I}' \in \mathcal{X}^{\Theta \beta}_{\mathcal{I}}$. In addition to the non-existence of UMP tests for the RDT problem (5), neither standard results — such as Karlin-Rubin’s Theorem [12, Theorem 3.4.1, p. 65, corollary 3.4.1, p. 67] or [12, Theorem 3.7.1, p. 81] — nor results based on the invariance principle [12–14] directly apply to RDT, where the signal is assumed to be random with unknown probability distribution. Nevertheless, the RDT problem (5) is invariant in the following sense.

### 4.2 Invariance of the RDT problem

With the notation of Section 3.2, given any $\mathcal{G} \in \mathcal{S}$, $\|\mathcal{G}(y) - \theta_0\| = \|y - \theta_0\|$ for all $y \in \mathcal{X}^d$. Therefore, for the RDT problem (5), $\mathcal{G}(\Theta_t) = \Theta_t$, $\mathcal{G}(\Omega_0) = \Omega_0$ and $\mathcal{G}(\Omega_1) = \Omega_1$, so that $\Theta_t$, the null and alternative events are invariant under the action of $\mathcal{G}$. Moreover, there exists $g \in \mathcal{G}$ such that $\mathcal{G}(y) = g(y - \theta_0) + \theta_0$ for any $y \in \mathcal{X}^d$. The observation $Y$ in (5) is thus $\mathcal{G}$-invariant in that, for any $\Theta \in \vartheta_r$ and any $\mathcal{G} \in \mathcal{S}$, $\mathcal{G}(Y) = \mathcal{G}(\Theta) + g(X)$ where $\mathcal{G}(\Theta) \in \Theta_t$ and $g(X) \sim N(0, C)$. These facts characterize the invariance of the RDT problem (5) under the action of groups $\mathcal{G}$ and $\tilde{\mathcal{G}}$, which thus play very similar roles to those of groups intervening in invariant hypothesis testing problems [12, Chapter 6, Section 6.1]. In particular, since the orbits of $\tilde{\mathcal{G}}$ are the ellipsoids $\mathcal{Y}_\rho$ of $\mathcal{G}$, $\mathcal{G}(\Theta(\omega))$ belongs to the same ellipsoid as $\Theta(\omega)$ for any $\omega \in \Omega$ and any $\mathcal{G} \in \mathcal{S}$. We thus look for tests that optimize some suitable notion of constant power on every $\mathcal{Y}_\rho$, as Wald's tests in [1, Section 6, Proposition III, p. 450] have constant power on each
ellipsoid \( Y_\rho \) because the problem of testing the mean of a Gaussian distribution is also invariant under the action of \( \Xi \). Since \( \Theta \) has unknown distribution, the probability that \( \Theta \) belongs to a given \( Y_\rho \) is unknown as well and the suitable notion of power for taking the invariance of the RDT problem into account will then be conditional. The next section specifies this conditional notion of power on which our invariance-based criterion for optimality relies.

### 4.3 Conditional power with respect to ellipsoids

**Definition 1** Let \( \mathcal{T} \) be some test. Given any \( \Theta \in \varnothing_\tau \) and any ellipsoid \( Y_\rho \), the power of \( \mathcal{T} \) conditioned to \( \Theta \in Y_\rho \) is defined as the conditional probability

\[
\beta_\Theta(\mathcal{T} | \Theta \in Y_\rho) = P[\mathcal{T}(Y) = 1 | \Theta - \theta_0 = \rho, \Theta \in Y_\rho].
\]

The reader will verify the next lemma, which is very similar to Lemma 3 and relates the power of \( \mathcal{T} \) conditioned to \( \Theta \in Y_\rho \) to the power function of \( \mathcal{T} \).

**Lemma 4** Let \( \rho \) and \( \rho' \) be any elements of \( \mathbb{R}^d \) such that \( \| \Theta - \theta_0 \| \leq \tau \) and \( \| \Theta' - \theta_0 \| > \tau \). For any Bernoulli random variable \( \epsilon \) independent of \( X \) and valued in \( \{0, 1\} \) with \( P[\epsilon = 1] = \rho, \Theta = (1 - \epsilon)\theta + \epsilon \theta' \in \varnothing_\tau, \beta_\Theta(\mathcal{T} | \Theta \in Y_{\theta - \theta_0}) = \beta_\Theta(\mathcal{T}) \) and \( \beta_\Theta(\mathcal{T} | \Theta \in Y_{\theta' - \theta_0}) = \beta_{\rho'}(\mathcal{T}) \).

The size and power of \( \mathcal{T} \) for the RDT problem (5) also relate to \( \beta_\Theta(\mathcal{T} | \Theta \in Y_\rho) \) when \( \rho \) varies since Bayes’ rule and definition 1 imply that

\[
P[\mathcal{T}(Y) = 1 | \Omega_0] = \frac{1}{P(\Omega_0)} \int_{[0, 1]} \beta_\Theta(\mathcal{T} | \Theta \in Y_\rho) P[\| \Theta - \theta_0 \|^{-1}](d\rho)\tag{8}
\]

and

\[
\beta_{\rho'}(\mathcal{T}) = \frac{1}{P(\Omega_1)} \int_{(\tau, \infty)} \beta_\Theta(\mathcal{T} | \Theta \in Y_\rho) P[\| \Theta - \theta_0 \|^{-1}](d\rho).\tag{9}
\]

There is no test \( \mathcal{T} \in \mathcal{X}_\Theta \) such that, for all \( \mathcal{T}' \in \mathcal{X}_\Theta \) and all \( \Theta \in \varnothing_\tau, \beta_\Theta(\mathcal{T} | \Theta \in Y_\rho) > \beta_{\rho'}(\mathcal{T} | \Theta \in Y_\rho) \) for \( P[\| \Theta - \theta_0 \|^{-1} = 0 \) almost every \( \rho > \tau \). Indeed, if such a test \( \mathcal{T} \) existed, it follows from (9) that this test would be UMP with level \( \gamma \) for testing \( \Omega_0 \), a contradiction.

### 4.4 Tests with uniformly best constant conditional power on family \( \mathcal{X} \)

To deal with the invariance of the problem, we introduce a conditional notion of constant power based on definition 1. As such, the following definition characterizes the invariance of a test \( \mathcal{T} \), given observation \( Y \) and regardless of the testing problems for which \( \mathcal{T} \) could be used. This definition leads to the criterion optimized to test \( \Omega_0 \).

**Definition 2** Given any \( \Theta \in \varnothing_\tau \) independent of \( X \), a test \( \mathcal{T} \) is said to have constant power conditioned to \( \Theta \in Y_\rho \) if \( \beta_\Theta(\mathcal{T} | \Theta \in Y_\rho) = \beta_0(\mathcal{T}) \) for any \( \Theta \in Y_\rho \).

Let \( \mathcal{T} \) be some test. For any \( \Xi \in M(\Omega, \mathbb{R}^d) \), independent of \( X \) and such that \( \Xi \in Y_\rho \) (a.s.), we have

\[
P[\mathcal{T}(\Xi + X) = 1] = f_{\mathcal{T}}(\mathcal{T}) (\Xi^{-1}) \] (d\Xi).\]

Hence, the following lemma.

**Lemma 5** For any \( \Theta \in \varnothing_\tau \), test \( \mathcal{T} \) has constant power conditioned to \( \Theta \in Y_\rho \) if and only if \( \beta_\Theta(\mathcal{T} | \Theta \in Y_\rho) = P[\mathcal{T}(\Xi + X) = 1] \) for any \( \Xi \in M(\Omega, \mathbb{R}^d) \), independent of \( X \) and such that \( \Xi \in Y_\rho \) (a.s.).
4.4 Tests with uniformly best constant conditional power on family $\tilde{\mathbb{F}}$

If test $\mathcal{T}$ has constant power conditioned to $\Theta \in \mathbb{Y}_\rho$ for some $\Theta \in \vartheta_r$, it has constant power on $\mathbb{Y}_\rho$. The next result extends this easy remark.

**Proposition 1** A given test $\mathcal{T}$ has constant power on every $\mathbb{Y}_\rho \in \tilde{\mathbb{F}}$ if and only if, for any $\Theta \in \vartheta_r$, $\mathcal{T}$ has constant power conditioned to $\Theta \in \mathbb{Y}_\rho$ for $P[\Theta - \Theta_0]^{\rho} - \text{almost every } \rho \geq 0$.

**Proof:** See Appendix IV. □

With respect to the RDT problem (5), we can expect, for a given $\Theta \in \vartheta_r$ and $P[\Theta - \Theta_0]^{\rho} - \text{almost every } \rho \geq 0$, the existence of a test with specified level $\gamma \in (0, 1)$ and constant power conditioned to $\Theta \in \mathbb{Y}_\rho$, that is, a test $\mathcal{T} \in \mathcal{K}_\vartheta^\Theta$ with constant power conditioned to $\Theta \in \mathbb{Y}_\rho$ and such that $\beta(\mathcal{T} \mid \Theta \in \mathbb{Y}_\rho) \geq \beta_0(\mathcal{T} \mid \Theta \in \mathbb{Y}_\rho)$ for any other $\mathcal{T}' \in \mathcal{K}_\vartheta^\Theta$ with constant power conditioned to $\Theta \in \mathbb{Y}_\rho$. The optimality of such a test would be limited to $\Theta$, whereas our goal is to point out tests that are optimal, in a certain sense related to spherical invariance, to test $\Omega_0$ for all $\Theta \in \vartheta_r$. Thence, the following criterion.

**Definition 3** A given test $\mathcal{T}^*$ is said to have level (resp. size) $\gamma \in (0, 1)$ and uniformly best constant power conditioned on $\mathbb{Y} = \{\mathbb{Y}_\rho : \rho \geq 0\}$ — and we say that $\mathcal{T}^*$ is UBCCP — for the RDT problem (5) if:

(i) for $P[\Theta - \Theta_0]^{\rho} - \text{almost every } \rho \geq 0, \mathcal{T}^*$ has constant power conditioned to $\Theta \in \mathbb{Y}_\rho$,

(ii) for $P[\Theta - \Theta_0]^{1} - \text{almost every } \rho \in (\tau, \infty), \beta\left(\mathcal{T}^* \mid \Theta \in \mathbb{Y}_\rho\right) \geq \beta_p(\mathcal{T} \mid \Theta \in \mathbb{Y}_\rho)$ when $\mathcal{T}$ is any test in $\mathcal{K}_\vartheta^\Theta$ with constant power conditioned to $\Theta \in \mathbb{Y}_\rho$.

With the notation of this definition, property (ii) is rather strong since it means that, for $P[\Theta - \Theta_0]^{\rho} - \text{almost every } \rho \in (\tau, \infty)$, the power of $\mathcal{T}^*$ conditioned to $\Theta \in \mathbb{Y}_\rho$ must not be less than that of $\mathcal{T}$, whatever the behaviour of $\mathcal{T}$ on any ellipsoid other than $\mathbb{Y}_\rho$. As a straightforward application of (9) and definition 3, we have:

**Proposition 2** Let $\mathcal{T}^*$ be some UBCCP test with level $\gamma \in (0, 1)$ for the RDT problem (5). Given $\Theta \in \vartheta_r, \mathcal{T}^*$ is UMP within the class of those elements of $\mathcal{K}_\vartheta^\Theta$ that have constant power conditioned to $\Theta \in \mathbb{Y}_\rho$ for $P[\Theta - \Theta_0]^{\rho} - \text{almost every } \rho \in (\tau, \infty)$.

The following theorem establishes the existence of UBCCP tests.

**Theorem 1** The thresholding tests with threshold height $\lambda_\tau(\tau)$ on the Mahalanobis distance to $\Theta_0$ have size $\gamma$, are unbiased and UBCCP for the RDT problem (5). For any $\Theta \in \vartheta_r$ and $P[\Theta - \Theta_0]^{\rho} - \text{almost every } \rho \in (\tau, \infty)$, the constant power conditioned to $\Theta \in \mathbb{Y}_\rho$ of any of these tests is $1 - R(\rho, \lambda_\tau(\tau))$.

**Proof:** See Appendix V. □

**Remark 1** If $\tau > 0$, Theorem 1 still holds true if $\Omega_0 = [\|\Theta - \Theta_0\| < \tau]$ and $\Omega_1 = [\|\Theta - \Theta_0\| \geq \tau]$ in the RDT problem (5), by opening $[0, \tau]$ to the right and closing $(\tau, \infty)$ to the left in definition 3.
Remark 2  In RDT from below, UBCCP tests are defined by replacing $(\tau, \infty)$ with $[0, \tau)$ in definition 3. Now, say that a thresholding test from below $\eta \geq 0$ on the Mahalanobis distance to model $\theta_0$ is any test $T$ defined for any $y \in \mathbb{R}^d$ by $T(y) = 1$ (resp. $T(y) = 0$) if $\|y - \theta_0\| < \eta$ (resp. $\|y - \theta_0\| > \eta$). By comparison, UBCCP tests of Theorem 1 are thresholding tests from above $\lambda_\tau(\tau)$, as the RDT problem (5) for which they are designed is itself from above $\tau$. Theorem 1 then holds true in RDT from below by using thresholding tests from below $\lambda_1 - \gamma(\tau)$ instead of thresholding tests (from above) with threshold height $\lambda_\tau(\tau)$ and replacing $1 - \mathcal{R}(\rho, \lambda_\tau(\tau))$ by $\mathcal{R}(\rho, \lambda_1 - \gamma(\tau))$. Thresholding tests from below $\lambda_1 - \gamma(\tau)$ are also UBCCP for testing $\|\theta - \theta_0\| \geq \tau$ against $\|\theta - \theta_0\| < \tau$ with $\tau > 0$, for the same reasons as those of Remark 1.

Remark 3  Theorem 1 implies [1, Section 6, Proposition III, p. 450] for $\tau = 0$. This will be made clearer in the next section.

5 Application to deterministic distortion testing (DDT) of a Gaussian distribution

5.1 problem statement

The probability distribution of the received signal is now assumed to be a Dirac mass centred at some unknown $\theta \in \mathbb{R}^d$, so that the observation is $Y \sim N(\theta, C)$. Considering that $\theta$ is a distorted version of $\theta_0$, we hereafter focus on DDT of $N(\theta, C)$ from above $\tau \geq 0$, that is, the composite hypothesis testing problem summarized by:

\[
\begin{align*}
\text{Observation:} && Y & \sim N(\theta, C), \\
\text{Tested or null hypothesis:} && \|\theta - \theta_0\| & \leq \tau, \\
\text{Alternative hypothesis:} && \|\theta - \theta_0\| & > \tau.
\end{align*}
\]

As in Section 4, the results in DDT from above will be transposed to DDT from below, that is, the hypothesis testing problem where the null and alternative hypotheses in (10) are exchanged.

Following standard terminology, the size of a given test $T$ for testing $\|\theta - \theta_0\| \leq \tau$ is

\[
\alpha(T) = \sup_{\theta \in \mathbb{R}^d} \beta_0(T) \quad \text{for} \theta \in \mathbb{R}^d \text{ such that } \|\theta - \theta_0\| \leq \tau.
\]

Given $\gamma \in [0, 1]$, $T$ is said to have level (resp. size) $\gamma$ for testing $\|\theta - \theta_0\| \leq \tau$ if $\alpha(T) \leq \gamma$ (resp. $\alpha(T) = \gamma$). Hereafter, $\mathcal{K}_\gamma$ denotes the class of those tests $T$ such that $\alpha(T) \leq \gamma$. Second, the power of $T$ for testing $\|\theta - \theta_0\| \leq \tau$ is defined as the restriction of the power function of $T$ to vectors $\theta$ such that $\|\theta - \theta_0\| > \tau$. According to the standard definition of unbiased tests, an unbiased test for testing $\|\theta - \theta_0\| \leq \tau$ is any test $T$ such that $\beta_0(T) \geq \alpha(T)$ for any $\theta \in \mathbb{R}^d$ such that $\|\theta - \theta_0\| > \tau$. As usual, we say that test $T$ is $\textit{UMP within some class } \mathcal{K}$ of tests that have all same level $\gamma$ for the DDT problem (10) if (i) $T$ belongs to $\mathcal{K}$, (ii) $\alpha(T) \leq \gamma$ and (iii) $\beta_0(T) \geq \beta_0(T')$ for any $T' \in \mathcal{K}$ and any $\theta$ such that $\|\theta - \theta_0\| > \tau$.

The DDT problem (10) is invariant, in the usual sense [12, Chapter 6, Section 6.1], under the action of $\mathcal{G}$ introduced in Section 3.2. Therefore, according to [12, Theorem 6.2.1] — or [13, Theorem 1, Sec. 47, chapter III, p. 281] — and the Karlin-Rubin Theorem [12, Theorem 3.4.1], there exists a $\mathcal{G}$-UMPI test with specified level $\gamma \in (0, 1)$, that is, a UMP test in the class of those tests that are $\mathcal{G}$-invariant with level
\(\gamma\), for testing \(\|\theta - \theta_0\| \leq \tau\). Although DDT basically differs from RDT because no deterministic vector belongs to \(\theta_0\), the DDT problem (10) can be cast in the RDT framework so as to state a more general result. The following lemma begins by establishing links between the notions of size and power in DDT and RDT.

**Lemma 6** Given any test \(\mathcal{T}\), \(\beta_0(\mathcal{T}) \leq \alpha(\mathcal{T}) \leq \alpha_0(\mathcal{T})\) for any \(\theta \in \mathbb{R}^d\) such that \(\|\theta - \theta_0\| \leq \tau\). Moreover, if test \(\mathcal{T}\) has constant power on every ellipsoid \(Y_\rho \in \mathfrak{F}\) and has level (resp. size) \(\gamma \in (0, 1)\) for the DDT problem (10), then \(\mathcal{T}\) has level (resp. size) \(\gamma\) for the RDT problem (5).

**Proof:** The first statement of this lemma derives directly from Lemma 3 and (11). If \(\mathcal{T}\) has level (resp. size) \(\gamma \in [0, 1]\) for the DDT problem (10) and constant power on every ellipsoid \(Y_\rho \in \mathfrak{F}\), the first statement of the present lemma and Proposition 1 imply that, for every \(\Theta \in \Theta\), \(\beta_0(\mathcal{T} \mid \Theta \in Y_\rho) \leq \gamma\) (resp. \(\beta_0(\mathcal{T} \mid \Theta \in Y_\rho) = \gamma\)) for \(P(\Theta - \Theta_0)\) almost every \(\rho \leq \tau\). The second statement of this lemma then follows from (6) and (8).

As an easy extension of [1, Definition III, Section 6, p. 450] in the Gaussian case, we put the following definition.

**Definition 4** A test \(\mathcal{T}^*\) with level \(\gamma\) is hereafter said to be UBCP on \(\mathfrak{F} = \{Y_\rho; \rho \geq 0\}\) — and we simply say that \(\mathcal{T}^*\) is UBCP — for the DDT problem (10) if \(\mathcal{T}^*\) is UMP within the class of tests with level \(\gamma\) and constant power on every \(Y_\rho\) of \(\mathfrak{F}\).

A UBCP test on \(\mathfrak{F}\) is also \(\mathfrak{F}\)-UMPI, since a test with constant power on every \(Y_\rho \in \mathfrak{F}\) has \(\mathfrak{F}\)-invariant power function. Note also that, in the definition above, the optimality is defined in a class of tests larger than that involved in [1, Section 6, Proposition III, p. 450] and based on [1, Definition III, Section 6, p. 450]. We then have the following result.

**Theorem 2** If \(\mathcal{T}^*\) is UBCP with level (resp. size) \(\gamma\) for the RDT problem (5), then \(\mathcal{T}^*\) is UBCP with level (resp. size) \(\gamma\) for the DDT problem (10).

**Proof:** Let \(\mathcal{T}\) be any test with constant power on every \(Y_\rho \in \mathfrak{F}\) and level \(\gamma\) for the DDT problem (10). According to Lemma 6, \(\mathcal{T} \in \mathfrak{X}_\gamma^{\mathfrak{F}}\). Given any two elements \(\theta\) and \(\theta'\) of \(\mathbb{R}^d\) such that \(\|\theta - \theta_0\| \leq \tau\) and \(\|\theta' - \theta_0\| > \tau\), let us consider \(\Theta = (1 - \epsilon)\theta + \epsilon\theta'\) \(\in \Theta\), where \(\epsilon\) is independent of \(X\) and valued in \([0, 1]\) with \(P[\epsilon = 1] \in (0, 1).\) Proposition 1 then implies that \(\mathcal{T}\) has constant power conditioned to \(\Theta \in Y_{\theta_0-\theta_0}\). In addition, we obtain \(\beta_\rho(\mathcal{T}) = \beta_0(\mathcal{T} \mid \Theta \in Y_{\theta_0-\theta_0})\) from Lemma 4.

If \(\mathcal{T}^*\) is UBCP with level (resp. size) \(\gamma\) for the RDT problem (5), it follows from Proposition 1 and Lemma 6 that this test has constant power on every \(Y_\rho \in \mathfrak{F}\) and level (resp. size) \(\gamma\) for testing \(\|\theta - \theta_0\| \leq \tau\). It thus satisfies the same properties as \(\mathcal{T}\) and, in particular, \(\beta_\rho(\mathcal{T}^*) = \beta_0(\mathcal{T}^* \mid \Theta \in Y_{\theta_0-\theta_0})\). Since \(\beta_{\Theta}(\mathcal{T}^* \mid \Theta \in Y_{\theta_0-\theta_0}) \geq \beta_0(\mathcal{T}^* \mid \Theta \in Y_{\theta_0-\theta_0})\), it follows that \(\beta_\rho(\mathcal{T}^*) \geq \beta_\rho(\mathcal{T})\), which concludes the proof, since \(\rho'\) is arbitrarily chosen.

As a direct consequence of Lemma 3, an unbiased test for the RDT problem (5) is unbiased as well for the DDT problem (10). The next result then derives straightforwardly from Theorems 1, 2 and this remark.

**Theorem 3** Given \(\gamma \in (0, 1)\), any thresholding test with threshold height \(\lambda_\gamma(\tau)\) on the Mahalanobis distance to \(\theta_0\) is unbiased, has size \(\gamma\), is UBCP and, thus, \(\mathfrak{F}\)-UMPI for the DDT problem (10).
Remark 4 The standard two-sided testing problem of accepting or rejecting $\theta = \theta_0$ against $\theta \neq \theta_0$ on the basis of observation $Y \sim N(\theta, C)$ with $C > 0$ is the DDT problem (10) with null tolerance. Therefore, [1, Section 6, Proposition III, p. 450] is a direct corollary of Theorem 3.

Remark 5 When $\tau > 0$, DDT from below is the testing problem where the null and alternative hypotheses in (10) are exchanged. As in the random case, the results of this section remain valid in DDT from below by swapping inequalities $\|\theta - \theta_0\| \leq \tau$ and $\|\theta - \theta_0\| > \tau$ in (10) and using thresholding tests from below (see Remark 2) instead of thresholding tests from above, in the statement of Theorem 3. The results of this section hold also true in DDT from above and from below when inequalities in the wide-sense are replaced by strict inequalities and vice-versa in the null and alternatives hypotheses.

Remark 6 In the one-dimensional case, two additional results derive from properties of $R$ and standard results. On the one hand, from [12, Theorem 3.7.1] and the properties of $R$, any test such that $T(y) = 1$ (resp. $T(y) = 0$) for any $y \in \mathbb{R}$ such that $|y - \theta_0| < \sigma \lambda_1(\tau/\sigma)$ (resp. $|y - \theta_0| > \sigma \lambda_1(\tau/\sigma)$) is UMP with size $\gamma$ for testing any of the hypotheses $|\theta - \theta_0| > \tau$ and $|\theta - \theta_0| \geq \tau$, given observation $Y \sim N(\theta, \sigma^2)$. On the other hand, there is no UMP test for testing any of the hypotheses $|\theta - \theta_0| \leq \tau$ and $|\theta - \theta_0| < \tau$. However, it follows from [12, Eqs. (4.2) & (4.3), Section 4.2] and the properties of $R$ that any test $T$ such that $T(y) = 1$ (resp. $T(y) = 0$) for any $y \in \mathbb{R}$ such that $|y - \theta_0| > \sigma \lambda_1(\tau/\sigma)$ (resp. $|y - \theta_0| < \sigma \lambda_1(\tau/\sigma)$) is UMPU with size $\gamma \in (0, 1)$ for testing any of these hypotheses, given observation $Y \sim N(\theta, \sigma^2)$.

6 Application to signal detection

The detection of an unknown and non-null $d$-dimensional signal in additive independent Gaussian noise is a problem of most interest in practice. In many papers and textbooks, the unknown signal is considered to be deterministic. Depending on the geometrical structure that this deterministic unknown signal may satisfy — for instance, if this signal obeys a linear subspace model —, the spherical- and scale-invariance of the problem makes it possible to reduce it via the invariance principle [12–14] and design tests invariant to nuisance parameters. Such tests often relate to the GLRT for the natural invariance these tests exhibit by involving maximum likelihood estimates of nuisance parameters [16–19]. For instance, [15–19], amongst others, address the case of a noise covariance matrix with known form and design tests optimal within a restricted class of tests invariant to nuisance parameters. Subspace adaptive detectors take the scale-invariance of the detection problem into account, when the noise matrix covariance is unknown and auxiliary data are available [20–30].

For reasons similar to those of Section 1, a random signal model might be preferred in practice. In this respect, the sequel addresses the detection of a random signal with unknown distribution in independent additive Gaussian noise. This problem is cast in the RDT framework and the contribution of RDT and UBCCP tests to signal detection is discussed. We begin with the case of a known noise covariance matrix. In subsection 6.2, we consider the case where the Gaussian noise components are independent with same unknown standard deviation and the detection is performed via an estimate-and-plug-in detector, based on a noise reference used to
estimate this unknown noise standard deviation. We postpone to further study the case of any unknown noise covariance matrix.

### 6.1 Detection in noise with known covariance matrix

Let $\Xi$ be some $d$-dimensional real random signal whose distribution is unknown and such that $\Xi \neq 0$ (a-s). We assume that $\Xi$ is independent of noise $X \sim N(0, C)$ with $C > 0$. We assume that $C$ is known. The problem of detecting $\Xi$ is usually described as the binary hypothesis testing (see [7, 9, 31, 32]) where the null hypothesis $\mathcal{H}_0$ is that only noise is present and the alternative hypothesis $\mathcal{H}_1$ is that the observation is the sum of signal and noise. We always can assume the existence of some non-negative real value $r_0$, possibly equal to the trivial lower bound 0 for the norm, such that $\|\Xi\| > r_0$ (a-s). Denoting the observation by $Y$, the problem of detecting $\Xi$ in noise $X$ can then be summarized by

$$
\begin{cases}
\mathcal{H}_0 : Y \sim N(0, C), \\
\mathcal{H}_1 : Y = \Xi + X, X \sim N(0, C), P[\|\Xi\| > r_0] = 1.
\end{cases}
$$

The performance of a given test $T$ is usually measured via the false alarm and detection probabilities. The false alarm probability is the probability $P_{FA}[T] = P[T(X) = 1]$ of erroneously accepting the null hypothesis $\mathcal{H}_1$ when the observation is noise only. The detection probability is the probability $P_{D}[T] = P[T(\Xi + X) = 1]$ of correctly accepting the alternative hypothesis $\mathcal{H}_1$.

We now cast the detection problem (12) in the RDT theoretical framework of Section 4. First, we assume the existence of a random variable $\epsilon$, independent of $\Xi$ and $X$, defined on the same probability space as $\Xi$ and $X$, valued in $[0, 1]$ and such that $Y = \epsilon \Xi + X$. The signal is present (resp. absent) whenever $\epsilon = 1$ (resp. $\epsilon = 0$). Given any test $T$, the value of the random variable $T(Y) = T \circ Y$ is the index of the accepted hypothesis, whereas the value of $\epsilon$ is the index of the true hypothesis. Although the introduction of the indicator variable $\epsilon$ induces that of the signal prior probabilities of presence $P[\epsilon = 1]$ and absence $P[\epsilon = 0]$, in contrast to the standard Neymann-Pearson approach, the role of these priors is very limited but convenient to consider (12) as an RDT problem. For the problem to be meaningful, we assume that $P[\epsilon = 1] \in (0, 1)$. Second, set $\Theta = \epsilon \Xi$. For any non-negative real value $\tau$ such that $0 \leq \tau \leq r_0$, we have $P[\|\Theta\| \leq \tau] = P[\epsilon = 0]$ and $P[\|\Theta\| > \tau] = P[\epsilon = 1]$. It follows that the detection problem (12) is equivalent to the RDT problem (5) from above $\tau$, with $\delta_0 = 0$ and $\tau \in [0, r_0]$. According to theorem 1, any thresholding test $T_{\lambda_f}(\tau)$ given by (3) with $\theta_0 = 0$ and threshold height $\lambda_f(\tau)$ is UBCCP with level $\gamma \in (0, 1)$ for this RDT problem. We now calculate and comment the performance of this test in terms of false alarm and detection probabilities.

The false alarm probability of any thresholding test defined by (3) with $\theta_0 = 0$ — and we hereafter call such a test a thresholding test on the Mahalanobis norm — is

$$P_{FA}[T_\eta] = P[\|X\| > \eta] = 1 - \mathcal{R}(0, \eta) \leq 1 - \mathcal{R}(\tau, \eta),$$

which follows from (4) and the increasingness of $1 - \mathcal{R}(\cdot, \eta)$ induced by Lemma 1. Since the detection problem (12) is an RDT problem, (13) also derives from Section 4. In fact, because $\epsilon, \Xi$ and $X$ are independent, $P_{FA}[T_\eta] = P[T_\eta(Y) = 1 \mid \|\Theta\| \leq \tau] \leq a^0(T_\eta)$. We then derive (13) from statements (ii) and (iii) of Lemma 7 given in Appendix V, since $P[\Theta\|^{-1} = P[\epsilon = 0] \delta_0 + P[\epsilon = 1] P[\|\Xi\|^{-1} \leq \tau] and \|\Xi\| > \tau$ (a-s), where $\delta_4$ is
the Dirac measure at a given real number \(x\). According to the definition of \(\lambda_T(\tau)\), it follows from (13) that

\[
P_{\text{FA}}[\mathcal{J}_{\lambda_T(\tau)}] = 1 - \mathcal{R}(0, \lambda_T(\tau)) \leq \gamma. \tag{14}\]

The detection probability \(P_{\text{D}}[\mathcal{J}_\gamma] = P[\|X + \Xi\| > \eta]\) is equal to \(P[\mathcal{J}_\eta = 1]|\Theta| > \eta] = \beta^\Theta_0(\mathcal{J}_\gamma)\), because of the independence of \(\epsilon, \Xi\) and \(X\) again. Since \(\|\Xi\| > \tau_0\) (a-s), it follows from the expression of \(P[\Theta]\), statement (iv) of Lemma 7 and the decreasingness of \(\mathcal{R}(\cdot, \eta)\) guaranteed by Lemma 1 that

\[
P_{\text{D}}[\mathcal{J}_\eta] = \int_{[\tau_0, \infty)} (1 - \mathcal{R}(\rho, \eta))P\|\Xi\|^{-1}(d\rho) \geq 1 - \mathcal{R}(\tau_0, \eta). \tag{15}\]

Therefore, the detection probability of \(\mathcal{J}_{\lambda_T(\tau)}\) is lower bounded by

\[
P_{\text{D}}[\mathcal{J}_{\lambda_T(\tau)}] \geq 1 - \mathcal{R}(\tau_0, \lambda_T(\tau)). \tag{16}\]

It is usual to characterize the performance of a family of tests with levels in \((0, 1)\) by the receiver operator characteristic (ROC) curve of this family of tests. Each point \(M\) of the ROC curve of this family is obtained for a given level \(\gamma\), the abscissa of \(M\) being the false alarm probability of the test with level \(\gamma\) and the ordinate of \(M\) being the detection probability of this same test. For a given tolerance \(\tau\), we thus consider the ROC curve of the family of UBCCP tests \(\Xi = \{\mathcal{J}_{\lambda_T(\tau)} : \gamma \in (0, 1)\}\). This ROC curve is the set of points \(\mathcal{C}[\mathcal{J}_{\lambda_T(\tau)}] = \{P_{\text{FA}}[\mathcal{J}_{\lambda_T(\tau)}], P_{\text{D}}[\mathcal{J}_{\lambda_T(\tau)}] : \gamma \in (0, 1)\}\). We can also consider the set of points \(\hat{\mathcal{C}}[\mathcal{J}_{\lambda_T(\tau)}] = \{1 - \mathcal{R}(0, \lambda_T(\tau)), 1 - \mathcal{R}(\tau_0, \lambda_T(\tau)) : \gamma \in (0, 1)\}\). This curve is hereafter called the lower ROC curve since, according to (14) and (16), it lies below the ROC curve. When \(\tau\) ranges in \((0, \tau_0)\), the families of UBCCP tests \(\Xi\) have all the same ROC and lower ROC curves. This simply follows from the fact that, given \(\mathcal{J}_{\lambda_T(\tau)}\) with \(\tau \in [0, \tau_0]\) and any \(\tau' \in [0, \tau_0]\), statement (iii) of Lemma 2 guarantees the existence of a unique \(\gamma' \in (0, 1)\) such that \(\lambda_T(\tau) = \lambda_T(\tau')\). The difference in performance between the UBCCP tests \(\mathcal{J}_{\lambda_T(\tau)}\) when \(\tau\) ranges in \([0, \tau_0]\) can however be exhibited by observing, for a given level \(\gamma \in (0, 1)\), the false alarm probability and the lower bound for the detection probability when \(\tau\) varies. In fact, when \(\tau\) increases to \(\tau_0\), the false alarm probability and the lower bound for the detection probability of \(\mathcal{J}_{\lambda_T(\tau)}\) both decrease and the former tends from above to \(1 - \mathcal{R}(0, \lambda_T(\tau_0))\), whereas the latter tends from above to \(1 - \mathcal{R}(\tau_0, \lambda_T(\tau))\). In contrast, when \(\tau\) decreases to 0, the false alarm probability and the lower bound for the detection probability of \(\mathcal{J}_{\lambda_T(\tau)}\) both increase, the former tending from below to the specified level, whereas the latter tends from below to \(1 - \mathcal{R}(\tau_0, \lambda_T(\tau_0))\). This behaviour straightforwardly derives from the properties of \(\mathcal{R}\) and \(\lambda_T\) and is coherent with the fact that the UBCCP tests have same ROC curve. For instance, figure 1 displays the false alarm probabilities and the lower bounds for the detection probabilities of several UBCCP tests \(\mathcal{J}_{\lambda_T(\tau)}\) when \(d = 12\), the signal norm lower bound is \(\tau_0 = 7\) and \(C = \mathbf{I}_d\).

Summarizing, unless the application requires a false alarm probability actually lesser than the specified level \(\gamma\), the most appropriate UBCCP test for detecting the signal is \(\mathcal{J}_{\lambda_T(\tau_0)}\) obtained by considering the trivial lower bound for the signal Mahalanobis norm. Indeed, the detection probability lower bound yielded by this test is the largest possible one, whereas the false alarm of this test remains equal to the specified level. In the sequel, \(\mathcal{J}_{\lambda_T(\tau_0)}\) is called Wald’s test with size \(\gamma\) because, according to Theorem 3, it is the UBCP test with level \(\gamma\) for testing the mean of a Gaussian
6.2 Detection in noise with unknown standard deviation

We now consider the case where Gaussian noise has independent components with same unknown standard deviation. Therefore, the Mahalanobis norm is now the standard Euclidean norm in $\mathbb{R}^d$. We begin with an easy remark. Suppose a unitary noise standard deviation and that a measurement of this value, say $\sigma$, has been provided by some device. Let us consider that this measurement is deterministic. When Wald’s test $\mathcal{T}^{(e)}(\gamma)$ with size $\gamma$ is adjusted with $\sigma$, we obtain the thresholding test $\mathcal{T}^{(e)}(\gamma)$ with threshold height $\sigma r^{(e)}(\gamma)$. The false alarm probability of this test is $P_{FA}[\mathcal{T}^{(e)}(\gamma)] = 1 - R(0, \sigma r^{(e)}(\gamma))$. If $\sigma < 1$, the strict increasingness of $R(0, \cdot)$ and the

distribution with covariance matrix $C$. The next section exhibits a situation where the flexibility on the actual size of the UBCCP tests proves helpful and Wald’s test becomes inappropriate.

6.2 Detection in noise with unknown standard deviation

We now consider the case where Gaussian noise has independent components with same unknown standard deviation. Therefore, the Mahalanobis norm is now the standard Euclidean norm in $\mathbb{R}^d$. We begin with an easy remark. Suppose a unitary noise standard deviation and that a measurement of this value, say $\sigma$, has been provided by some device. Let us consider that this measurement is deterministic. When Wald’s test $\mathcal{T}^{(e)}(\gamma)$ with size $\gamma$ is adjusted with $\sigma$, we obtain the thresholding test $\mathcal{T}^{(e)}(\gamma)$ with threshold height $\sigma r^{(e)}(\gamma)$. The false alarm probability of this test is $P_{FA}[\mathcal{T}^{(e)}(\gamma)] = 1 - R(0, \sigma r^{(e)}(\gamma))$. If $\sigma < 1$, the strict increasingness of $R(0, \cdot)$ and the
6.2 Detection in noise with unknown standard deviation

definition of \( \lambda_\gamma(0) \) imply that \( \text{P}_{\text{FA}}[\mathcal{T}_{\lambda_\gamma(0)}] > \gamma \), which is undesirable. The conclusion of the previous section may thus fail in practical cases where an estimate of the noise standard deviation is injected into the expression of the test. The use of a UBCCP test with non-null tolerance can therefore be expected to avoid this unwanted behaviour because, as emphasized in the previous section, a non-null tolerance lowers the size of the UBCCP test for detecting the signal.

Instead of further detailing the example above, let us tackle the more general situation where an estimate-and-plug-in detector is used to detect the signal. An estimate-and-plug-in detector involves performing an estimate \( \hat{\sigma} \) of the noise standard deviation thanks to a noise reference and using \( \hat{\sigma} \) instead of the true value in the expression of a test designed for the nominal case of a known standard deviation [7, Chapter 9, p. 337]. We assume that \( \hat{\sigma}, X \) and \( \Xi \) are independent. Without loss of generality because of the scale invariance of the problem, we suppose that the noise standard deviation is 1 again. Given \( \tau \in [0, \tau_0] \), the thresholding test with threshold height \( \lambda_\gamma(\tau) \) on the Mahalanobis norm is UBCCP for testing \( \|\Theta\| \leq \tau \) with \( \text{P}_{\|\Theta\| \leq \tau}(0) \in (0,1) \). By replacing the actual value of the noise standard deviation by its estimate \( \hat{\sigma} \) in the expression of this UBCCP test, we do not obtain a test in the sense given above but a UBCCP estimate-and-plug-in detector — in short, UBCCP detector —, which is henceforth denoted by \( \mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y) \). The UBCCP detector decides that the signal is present if \( \|y\| > \hat{\sigma} \lambda_\gamma(\tau/\hat{\sigma}) \) and that the signal is absent, otherwise. Once again, the handling of equality in this decision does not matter for the absolute continuity of \( \|Y\| \) with respect to Lebesgue’s measure in \( \mathbb{R} \). The index of the hypothesis accepted by \( \mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y) \) is thus the value of \( \mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y) \). Because of the independence of \( \hat{\sigma} \) and \( X \), it follows from (14) that

\[
P_{\text{FA}}[\mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y)] = \int (1 - \mathcal{R}(0, \sigma \lambda_\gamma(\tau/\sigma))) \text{P}_{\|\Theta\| \leq \tau}^{-1}(d\sigma).
\] (17)

The detection probability of \( \mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y) \) is unknown. However, thanks to the independence of \( \hat{\sigma} \) and \( X \), we have \( \text{P}_{\|\Theta\| \leq \tau}[\mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y)] = \text{P}_{\|\Theta\| \leq \tau}[\mathcal{T}_{\lambda_\gamma(\tau/\sigma)}] \text{P}_{\|\Theta\| \leq \tau}^{-1}(d\sigma) \). It then follows from (15) that the detection probability of the UBCCP detector is lower bounded by :

\[
\text{P}_{\|\Theta\| \leq \tau}[\mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y)] \geq \int (1 - \mathcal{R}(\tau_0, \sigma \lambda_\gamma(\tau/\sigma))) \text{P}_{\|\Theta\| \leq \tau}^{-1}(d\sigma).
\] (18)

As in Section 6.1, the lower ROC curve of the UBCCP detector \( \mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y) \) is defined as the set of points:

\[
\hat{\Gamma}[\mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y)] = \left\{ \left\{ \text{P}_{\|\Theta\| \leq \tau}[\mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y)], \int (1 - \mathcal{R}(\tau_0, \sigma \lambda_\gamma(\tau/\sigma))) \text{P}_{\|\Theta\| \leq \tau}^{-1}(d\sigma) : \gamma \in (0,1) \right\} \right\}.
\]

On the one hand, the larger the right hand side (rhs) in (18), the larger the detection probability of \( \mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y) \). The largest possible value for the rhs in (18) is

\[
\int (1 - \mathcal{R}(\tau_0, \sigma \lambda_\gamma(0/\sigma))) \text{P}_{\|\Theta\| \leq \tau}^{-1}(d\sigma).
\]

This value is the detection probability lower bound of the Wald estimate-and-plug-in detector — in short, Wald detector — \( \mathcal{T}_{\lambda_\gamma(0/\hat{\sigma})}(Y) \), which derives from Wald’s test \( \mathcal{T}_{\lambda_\gamma(0/\hat{\sigma})} \) by replacing the known unitary standard deviation by \( \hat{\sigma} \). Therefore, a suitable tolerance \( \tau \) should be as small as possible so as to guarantee a detection probability lower bound close to that yielded by the Wald detector. However, a too small value for \( \tau \) may not be appropriate for the following reason. According to the properties of \( \mathcal{R} \) and \( \lambda_\gamma \) stated in Section 3, \( \text{P}_{\text{FA}}[\mathcal{T}_{\lambda_\gamma(\tau/\hat{\sigma})}(Y)] \)
is a continuous and decreasing function of $\tau$ and thus, since $\tau \in [0, \tau_0]$, we have $\text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_0(Y)] \leq \text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)] \leq \text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_0(Y)]$. The upper bound in this inequality is the false alarm probability of the Wald detector $\mathcal{A}_{\hat{\lambda}_T}\tau_0(Y)$. If the false alarm probability $\text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_0(Y)]$ of the Wald detector is above $\gamma$, $\tau$ should therefore not be chosen too close to 0. It follows from the above remarks that the UBCCP detector for a given level $\gamma$ should be the adjusted-UBCCP (A-UBCCP) detector

$$\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)$$

with adjusted tolerance $\tau^* = \arg\min_{\tau \in [0, \tau_0]} \text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)] - \gamma$. From the continuity and strict decreasingness of $\text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)]$ with $\tau$,

$$\tau^* = \begin{cases} \tau_0 & \text{if } \text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)] \geq \gamma, \\ 0 & \text{if } \text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_0(Y)] < \gamma \end{cases}$$

and $\tau^*$ is the unique solution in $\tau$ to $\text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)] = \gamma$ if $\text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_0(Y)] < \gamma < \text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_0(Y)]$. In the last case, $\tau^*$ guarantees a false alarm probability of the A-UBCCP detector $\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)$ equal to the specified level $\gamma$ and

$$\tau^* = \min \{ \tau \in [0, \tau_0] : \text{P}_T[\mathcal{A}_{\hat{\lambda}_T}\tau_\gamma(Y)] < \gamma \}.$$  

If the estimate $\hat{\delta}$ is good enough, the A-UBCCP detector is expected to perform closely to Wald's test based on prior knowledge of the noise standard deviation.

To prolong the discussion, suppose that the noise reference is an $N$-dimensional random vector $W \sim N(0, \Sigma)$, independent of X and $\Xi$ so that $\hat{\delta}$ is the noise standard deviation maximum likelihood estimate (MLE) given by $\hat{\delta}_N = \|W\|/\sqrt{N}$. The A-UBCCP detector is then the MLE A-UBCCP detector $\mathcal{A}_{\hat{\delta}_N}\lambda_T(Y)$ where $\lambda_T$ is the adjusted tolerance calculated with $\hat{\delta} = \hat{\delta}_N$. The MLE strong consistency then implies that $\hat{\delta}_N$ tends to 1 (a.s) and $\lim_{\lambda_T \rightarrow \infty} \text{P}_T[\mathcal{A}_{\hat{\delta}_N}\lambda_T(Y)] = \text{P}_T[\mathcal{A}_{\lambda_T}\lambda_T(Y)] = 1 - \mathcal{R}(0, \lambda_T(Y))$. From the strict increasingness of $1 - \mathcal{R}(0, \lambda_T(Y))$ and the definition of $\lambda_T(Y)$ given by Lemma 2, we have $1 - \mathcal{R}(0, \lambda_T(Y)) < \gamma$ and $\text{P}_T[\mathcal{A}_{\hat{\delta}_N}\lambda_T(Y)] < \gamma$ for $N$ large enough. Therefore, for $N$ above some natural integer, $\tau^*_N$ is either 0 or such that $\text{P}_T[\mathcal{A}_{\hat{\delta}_N}\lambda_T(Y)] = \gamma$. Since the MLE strong consistency also implies that $\lim_{\lambda_T \rightarrow \infty} \text{P}_T[\mathcal{A}_{\hat{\delta}_N}\lambda_T(Y)] = \text{P}_T[\mathcal{A}_{\lambda_T}\lambda_T(Y)] = 1 - \mathcal{R}(0, \lambda_T(Y)) = \gamma$, we can conjecture that $\tau^*_N$ tends to 0. The proof of this conjecture is still an open issue. However, we can establish the existence of a subsequence of $(\tau^*_N)_{N=1,2,\ldots}$ that converges to $\tau$. Indeed, suppose the existence of some positive real value $\tau$ such that $\tau^*_N \rightarrow \tau$ for any large enough integer $N$. On the one hand, for any integer $N$ large enough, we would have $\text{P}_T[\mathcal{A}_{\hat{\delta}_N}\lambda_T(Y)] = \gamma$. On the other hand, for $N$ large enough again, it would follow from (17) and the decreasingness of $1 - \mathcal{R}(0, \cdot)$ that $\text{P}_T[\mathcal{A}_{\hat{\delta}_N}\lambda_T(Y)] \leq \int (1 - \mathcal{R}(0, \cdot)) \hat{\pi}_N^{-1}(d\sigma)$. Since the rhs in the inequality above tends to 1 by the MLE strong consistency, we would necessarily have $\gamma \leq 1 - \mathcal{R}(0, \lambda_T(Y))$, a contradiction since $\tau > 0, 1 - \mathcal{R}(0, \lambda_T(Y))$ is strictly increasing and $\gamma = 1 - \mathcal{R}(0, \lambda_T(Y))$ by Lemmas 1 and 2. For a sequence of tolerances that converges to 0, the lower ROC curve of the MLE A-UBCCP detector thus approaches that obtained when the noise standard deviation is known and the detection is performed by Wald's test. This is illustrated by figures 2 and 3. Figure 2 displays the false alarm probability of the MLE Wald detector — that is, the Wald detector adjusted with $\hat{\delta}_N$ — is above the specified level, whereas the MLE A-UBCCP detector guarantees the specified level. Figure 3 compares the lower ROC curve of MLE
A-UBCCP detector to that of Wald’s test based on prior knowledge of the noise standard deviation. The false alarm probabilities and the detection probability lower bounds were numerically calculated by quadrature Gaussian integration and standard MATLAB routines of the toolbox stats. In particular, \( R(\rho, \eta) = F_{\chi^2_d}(\rho^2, \eta^2) \) where \( F_{\chi^2_d}(\rho^2) \) is the cumulative distribution function of the non-central \( \chi^2 \) distribution with \( d \) degrees of freedom and non-central parameter \( \rho^2 \).

\[ \begin{align*}
\text{Figure 2: The signal and noise have dimension } d &= 12, \text{ the signal norm lower bound is } \tau_0 = 7 \text{ and the noise covariance matrix is } I_d. \text{ This figure displays the false alarm probability of the MLE Wald detector as a function of } \log_{10}(\gamma), \text{ when } N = 100, 150, 300, 500 \text{ to compute } \hat{\sigma}_N \text{ and the specified level } \gamma \text{ ranges in (0, 1). The false alarm probabilities of the MLE Wald detector are above the diagonal and, thus, above the specified level } \gamma, \text{ which is undesirable. In contrast, the MLE A-UBCCP detector guarantees the specified level.}
\end{align*} \]

7 Conclusion and perspectives

We have introduced the RDT problem in presence of additive independent Gaussian noise, that is, the problem of testing whether the Mahalanobis distance between some random signal \( \Theta \) with unknown distribution and a deterministic known model \( \theta_0 \) is below or above some non-negative tolerance \( \tau \) when we observe \( Y = \Theta + X \) where \( X \sim N(0, C) \) is independent of \( \Theta \) and \( C > 0 \). RDT also involves DDT, that is the problem of testing the Mahalanobis distance between \( \theta \) and \( \theta_0 \) with respect to \( \tau \), when we observe \( Y \sim N(\theta_0, C) \). The testing of the mean of a Gaussian distribution is the particular case where \( \tau = 0 \).

By introducing 1) the notion of power conditioned to the ellipsoids characterizing the RDT invariance and 2) the notion of tests with constant conditional power on family \( \mathcal{F} \) of these ellipsoids, thresholding tests on the Mahalanobis distance to \( \theta_0 \)
Figure 3: The signal and noise have dimension $d = 12$, the signal norm lower bound is $\tau_0 = 7$ and the noise covariance matrix is $I_d$. The figure displays the lower ROC curves of the MLE A-UBCCP detector with adjusted tolerance $\tau^*_N$, when $N = 100, 150, 300, 500$ to compute $\hat{\sigma}_N$. Whatever the tested value of $N$, $P_{FA}[T_{\hat{\sigma}_N, \lambda, \gamma}(Y)] < \gamma < P_{FA}[T_{\hat{\sigma}_N, \lambda, \gamma}(0)]$ so that $\tau^*_N$ guarantees that the false alarm probability of the MLE A-UBCCP detector equals the specified level. The use of an estimate of the noise standard deviation impacts the detection performance. Indeed, the lower ROC curves of the MLE A-UBCCP detector are below the lower ROC curve of Wald’s test obtained when the noise standard deviation is known. However, this performance loss in detection reduces as $N$ increases, whereas the false alarm probability of the MLE A-UBCCP detector remains equal to the specified level.
the deletion) of new (resp. old) obstacles, as well as the management of “tracked obstacles in a thresholded proximity of measurement” [34]. Fault-detection and structural health monitoring (SHM) could also be natural applications. “Because the stress level in any element will never be exactly zero, one must establish a threshold stress level for proper damage diagnosis” [35] and the introduction of a tolerance, aimed at bracketing possible fluctuations other than noise around the signal nominal model, could therefore be considered regardless of the fluctuation distributions. Fault-detection, robust to system uncertainties and external noise, remains a challenging task [36–38] and could possibly benefit from RDT, in combination or as alternative to usual approaches [39].

UBCCP tests, as UBCP ones, are alternative to likelihood ratio tests, which may perform poorly over the complementary of the class for which they are optimal [40, Sec. 3.1, p. 1160]. In connection with [40–42], it should therefore be relevant to elaborate on RDT problems involving nuisance parameters and large sample sizes. It could also be relevant to accommodate the theory of UBCCP tests with [10, Theorem VII.1], where the problem of detecting a random signal in AWGN is considered under the assumption that the signal norm is almost surely above some known positive value, as in Section 6.1 above, but where the criterion to optimize is the error probability instead of the power constrained by a specified level. A complete study of the MLE A-UBCCP, involving the case of a non signal-free reference, could also impact the design of constant false alarm rate (CFAR) systems in radar processing [43, 44] and, lately, in ultra wideband (UWB) receivers [45]. In this respect, to estimate the noise standard deviation from a non signal-free reference, RDT could perhaps be combined to standard results in robust statistics [46–49], as well as to [50, 51] that propose estimators in presence of random signals that have unknown distributions but obey sparsity hypotheses.

Another theoretical perspective is the following one. The invariance of the RDT problem is characterized by the orbits of the group \( G \) that leaves invariant the distribution of the observation in absence of distortion. This group directly derives from the group \( G \) of transforms leaving the noise distribution unchanged. The RDT problem and the UBCCP tests are then specified by the Mahalanobis distance to \( \theta_0 \), which is nothing else but a translated version of the Mahalanobis norm, a maximal invariant of \( G \). Therefore, group \( G \) and one of its maximal invariants have been sufficient to exhibit a large class of tests to solve event testing problems on \( \Theta \), without prior knowledge on the signal distribution. We aim at extending this approach to other types of additive independent noise, such as Generalized Gaussian or Gaussian mixture distributed noise. More generally, assuming that \( Y \) depends on some random signal \( \Theta \) with unknown distribution, we seek distribution families for \( Y \) conditioned to \( \Theta = \theta \), other than the Gaussian one, that are preserved by groups whose maximal invariants can be used to test events concerning \( \Theta \).

Acknowledgement

The author is very grateful to Lionel Fillatre for profitable discussions on invariance and Wald’s test and to his Ph. D. student Quang-Thang Nguyen, for his re-reading and remarks. The author is also very thankful to the reviewers and the editor, whose very sound comments and suggestions were crucial to improve the contents and the overall organization of this paper.
Appendix I
Proof of Lemma 1

This improvement of [10, Lemma IV.2] is proved similarly by refining some arguments. Let $\rho$ and $\rho'$ be two real numbers such that $0 \leq \rho < \rho' < \infty$. Let $\theta$ and $\theta'$ be two collinear vectors of $\mathbb{R}^d$ such that $\|\theta\| = \rho$ and $\|\theta'\| = \rho'$. According to (4), $\mathcal{R}(\rho, \eta) = \int_{B(\theta, \eta)} f(x) \, dx$ and $\mathcal{R}(\rho', \eta) = \int_{B(\theta', \eta)} f(x) \, dx$ where $f$ is the probability density function of $X$ and $B(\theta, \eta)$ (resp. $B(\theta', \eta)$) is the closed ball, in $\mathbb{R}^d$, centred at $\theta$ (resp. $\theta'$) with radius $\eta$. We have $\mathcal{R}(\rho, \eta) - \mathcal{R}(\rho', \eta) = \int_{B(\theta, \eta) \setminus B(\theta', \eta)} (f(x) - f(\theta + \theta' - x)) \, dx$. Let $(e_1, e_2, \ldots, e_d)$ be an orthonormal basis of $\mathbb{R}^d$ such that $\theta = \rho e_1$ and $\theta' = \rho' e_1$. We have $\|\theta + \theta' - x\|^2 - \|x\|^2 = (\rho + \rho')(\rho + \rho' - 2x_1)$ for any $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. If $x \in B(\theta, \eta) \setminus B(\theta', \eta)$, then $\|x - \theta\|^2 > \|x - \theta'\|$, which implies that $(\rho' - \rho)(\rho + \rho' - 2x_1) > 0$ and, thus, that $\rho + \rho' - 2x_1 > 0$ since $\rho' > \rho$. Therefore, $\|\theta + \theta' - x\| > \|x\|$. Since $f$ decreases strictly with the norm of its argument, it follows that $f(x) - f(\theta + \theta' - x) > 0$ so that $\mathcal{R}(\rho, \eta) > \mathcal{R}(\rho', \eta)$ and the proof is complete.

Appendix II
Proof of Lemma 2

[Existence and uniqueness of $\lambda_\gamma(\rho)$] $\mathcal{R}(\rho, \cdot)$ is a one-to-one mapping from $[0, \infty)$ into $[0, 1]$. Hence, the existence and the uniqueness of $\lambda_\gamma(\rho)$ for $\gamma \in (0, 1]$.

[Strict increasingness of $\lambda_\gamma$] Let $\rho$ and $\rho'$ be two non-negative real number such that $\rho < \rho'$. According to Lemma 1, $\mathcal{R}(\rho', \lambda_\gamma(\rho)) < \mathcal{R}(\rho, \lambda_\gamma(\rho))$. The rhs in this inequality equals $1 - \gamma$ and, thus, $\mathcal{R}(\rho', \lambda_\gamma(\rho'))$. The result then follows from the strict increasingness of $\mathcal{R}(\rho', \cdot)$.

[Continuity of $\lambda_\gamma$] Given $\rho_0 \in [0, \infty)$, the strict increasingness of $\lambda_\gamma$ implies the existence of a limit $\lambda_\gamma(\rho_0) \in [0, \infty)$ when $\rho$ tends to $\rho_0$ from below and the existence of a limit $\lambda_\gamma(\rho_0) \in [0, \infty)$ when $\rho$ tends to $\rho_0$ from above. Since $\mathcal{R}$ is continuous in the plane and $\mathcal{R}(\rho, \lambda_\gamma(\rho)) = 1 - \gamma$ for every $\rho \in [0, \infty)$, $\mathcal{R}(\rho_0, \lambda_\gamma(\rho_0)) = \mathcal{R}(\rho_0, \lambda_\gamma(\rho_0)) = 1 - \gamma$. Since $\mathcal{R}(\rho_0, \cdot)$ is one-to-one, $\lambda_\gamma(\rho_0) = \lambda_\gamma(\rho_0) = \lambda_\gamma(\rho_0)$ and $\lambda_\gamma$ is continuous.

[Strict decreasingness of $\gamma \mapsto \lambda_\gamma(\rho)$] Let $\rho$ be some element of $[0, \infty)$. Let us consider two elements $\gamma$ and $\gamma'$ of $(0, 1]$. We have $1 - \mathcal{R}(\rho, \lambda_\gamma(\rho)) = \gamma$ and $1 - \mathcal{R}(\rho, \lambda_{\gamma'}(\rho)) = \gamma'$. If $\gamma < \gamma'$, we thus have $\mathcal{R}(\rho, \lambda_\gamma(\rho)) > \mathcal{R}(\rho, \lambda_{\gamma'}(\rho))$, which implies that $\lambda_{\gamma'}(\rho) > \lambda_\gamma(\rho)$ since $\mathcal{R}(\rho, \cdot)$ is strictly increasing.

[Continuity of $\gamma \mapsto \lambda_\gamma(\rho)$] The proof is similar to that of the continuity of $\lambda_\gamma$ and left to the reader.

Appendix III
No UMP test for RDT in the random case

Let $\theta$ and $\theta'$ be any two elements of $\mathbb{R}^d$ such that $\|\theta - \theta_0\| \leq \tau$ and $\|\theta - \theta_0\| > \tau$. Let $\varepsilon$ be any Bernoulli distributed random variable independent of $X$, valued in $\{0, 1\}$ and such that $P[\varepsilon = 1] \in (0, 1)$. We construct the random vector $\Theta = (1 - \varepsilon)\theta + \varepsilon\theta' \in \Theta_r$. If a UMP test $\mathcal{F}$ existed within $\mathcal{K}^\Theta_r$, Lemma 3 implies that this test would be most powerful with level $\gamma$ to test $E[Y] = \theta$ against $E[Y] = \theta'$, given observation $Y = \Theta + X$. The existence of such a most powerful test, independent of the arbitrarily chosen $\theta$
and \( \theta' \), would contradict the Neyman-Pearson Lemma [12, Theorem 3.2.1, Sec. 3.2, p. 60].

**Appendix IV**

**Proof of Proposition 1**

We begin with the direct implication. If \( \mathcal{I} \) has constant power on every \( \mathcal{Y}_\rho \in \mathcal{F} \), we define the map \( \mathcal{R} \) of \([0,\infty)\) into \([0,1]\) such that \( \mathcal{R}(\rho) = \mathbb{P}[\mathcal{I}(\theta + X) = 1] \) for any \( \theta \in \mathcal{Y}_\rho \).

Let \( \Theta \) be some element of \( \Theta_r \) and \( B \) be any Borel set of \( \mathbb{R} \). From the independence of \( \Theta \) and \( X \), it follows from the definition of \( \mathcal{R} \) that, for any \( \theta \in \mathbb{R}^d \), \( \mathbb{P}[\mathcal{I}(Y) = 1, \|\Theta - \theta_0\| \in B \mid \Theta = \theta] = I_B(\|\theta - \theta_0\|)\mathcal{R}(\|\theta - \theta_0\|) \)

where \( I_B \) is the indicator function of \( B \): \( I_B(x) = 1 \) if \( x \in B \) and \( I_B(x) = 0 \), otherwise.

By the standard change-of-variable formula [52, Theorem 16.13], we now have

\[
\int I_B(\|\theta - \theta_0\|)\mathcal{R}(\|\theta - \theta_0\|)\Theta^{-1}(d\theta) = \int_B \mathcal{R}(\rho)\mathcal{P}[\Theta - \theta_0]^{-1}(d\rho).
\]

Therefore, the equalities above imply that

\[
\mathbb{P}[\mathcal{I}(Y) = 1, \|\Theta - \theta_0\| \in B] = \int_B \mathcal{R}(\rho)\mathcal{P}[\Theta - \theta_0]^{-1}(d\rho).
\]

On the other hand, \( \mathbb{P}[\mathcal{I}(Y) = 1, \|\Theta - \theta_0\| \in B] = \int_B \mathbb{P}[\mathcal{I}(Y) = 1 \mid \|\Theta - \theta_0\| = \rho]\mathcal{P}[\Theta - \theta_0]^{-1}(d\rho) \).

From the foregoing and the definition of a conditional probability, we obtain

\[
\mathbb{P}[\mathcal{I}(Y) = 1 \mid \|\Theta - \theta_0\| = \rho] = \mathcal{R}(\rho), \quad \mathcal{P}[\Theta - \theta_0]^{-1} - (a-s).
\]

Conversely, let test \( \mathcal{I} \) be such that, for any \( \theta \in \Theta \) \( r \), \( \mathcal{I} \) has constant power conditioned to \( \Theta \in \mathcal{Y}_\rho \) for \( \mathbb{P}[\Theta - \theta_0]^{-1} \) almost every \( \rho > \tau \). If \( \rho \leq \tau \) (resp. \( \rho > \tau \)), consider some \( \rho' > \tau \) (resp. \( \rho' \leq \tau \)) and construct \( \Theta = (1 - \varepsilon)\theta + \rho'\theta' \) where \( \theta \in \mathcal{Y}_\rho \), \( \theta' \in \mathcal{Y}_{\rho'} \) and the random variable \( \varepsilon \) is independent of \( X \) and valued in \([0,1]\) with \( \mathbb{P}[\varepsilon = 1] \in (0,1) \).

According to Lemma 4, \( \beta_{\rho'}(\mathcal{I} \mid \Theta \in \mathcal{Y}_\rho) = \beta_{\rho}(\mathcal{I}) \). Since \( \rho \) belongs to any support of \( \mathbb{P}[\Theta - \theta_0]^{-1} \), \( \mathcal{I} \) has constant power conditioned to \( \Theta \in \mathcal{Y}_\rho \) and thus constant power of every \( \mathcal{Y}_\rho \in \mathcal{F} \) since \( \rho \leq \tau \) (resp. \( \rho > \tau \)) has been chosen arbitrarily.
Appendix V
Proof of Theorem 1

We begin with two preliminary lemmas concerning thresholding tests.

Lemma 7 Let $\mathcal{T}_\eta$ be a thresholding test with threshold height $\eta \geq 0$ on the Mahalanobis distance to $\theta_0$.

(i) For any $\Theta \in \Theta_\tau$, $\mathcal{T}_\eta$ has constant power conditioned to $\Theta \in Y_\rho$ and $\beta_0(\mathcal{T}_\eta | \Theta \in Y_\rho) = 1 - \mathcal{R}(\rho, \eta)$ for $P(\Theta - \theta_0)^{-1}$ - almost every $\rho > 0$;

(ii) we have $P[\mathcal{T}_\eta(Y) = 1 | \Theta \leq \tau] = \frac{1}{P[Y]} \int_{(0,1]} \left( 1 - \mathcal{R}(\rho, \eta) \right) P(\Theta - \theta_0)^{-1}(d\rho)$;

(iii) $\mathcal{T}_\eta$ has size $\alpha^\Theta \left( \mathcal{T} \right) = 1 - \mathcal{R}(\tau, \eta)$ for the RDT problem (5);

(iv) $\mathcal{T}_\eta$ has power $\beta^\rho_\Theta(\mathcal{T}_\eta) = \frac{1}{P[Y]} \int_{(r,\infty)} \left( 1 - \mathcal{R}(\rho, \eta) \right) P(\Theta - \theta_0)^{-1}(d\rho)$ for the RDT problem (5).

Proof:
Proof of statement (i): Test $\mathcal{T}_\eta$ has constant power on every $Y_\rho$ in $\hat{X}$. Therefore, according to Proposition 1, for any $\Theta \in \Theta_\tau$, there exists some support $\mathcal{D}$ of $P(\Theta - \theta_0)^{-1}$ such that $\mathcal{T}_\eta$ has constant power conditioned to any $\Theta \in Y_\rho$ with $\rho \in \mathcal{D}$. Hence, $\beta_0(\mathcal{T}_\eta | \Theta \in Y_\rho) = \beta_0(\mathcal{T}_\eta)$ for any $\Theta \in Y_\rho$ and any $\rho \in \mathcal{D}$ and the value of $\beta_0(\mathcal{T}_\eta)$ derives from (1) and (4).

Proof of statement (ii): A straightforward application of statement (i) above and (8).

Proof of statement (iii): From statement (ii), (6) and the decreasingness of $\mathcal{R}(\tau, \eta)$ guaranteed by Lemma 1, we derive that $\alpha^\Theta(\mathcal{T}_\eta) \leq 1 - \mathcal{R}(\tau, \eta)$. For the reverse inequality, consider any $\rho \in [0, \tau)$. Set $\Theta = (1 - \epsilon)\theta + \epsilon \theta' \in \Theta_{\rho'}$ where $\theta \in Y_\rho$, $\theta' \in Y_{\rho'}$ with $\rho' > \tau$ and the random variable $\epsilon$ is independent of $X$, valued in $[0,1]$ with $P[\epsilon = 1] = 0.1$. We then have $P(\Theta - \theta_0)^{-1} = P[\epsilon = 1] \delta_\rho + P[\epsilon = 0] \delta_{\rho'}$ and it follows from statement (ii) that $P[\mathcal{T}_\eta(Y) = 1 | \parallel \Theta \parallel \leq \tau] = 1 - \mathcal{R}(\rho, \eta)$. By definition of $\alpha^\Theta(\mathcal{T}_\eta)$ given by (6), we have $1 - \mathcal{R}(\rho, \eta) \leq \alpha^\Theta(\mathcal{T}_\eta)$. Since $\rho$ is arbitrary in $[0, \tau)$, the continuity of $\mathcal{R}(\rho, \eta)$ implies that $\lim_{\rho \to \tau} \mathcal{R}(\rho, \eta) = \mathcal{R}(\tau, \eta)$. Therefore, 1 - $\mathcal{R}(\tau, \eta) \leq \alpha^\Theta(\mathcal{T}_\eta)$.

Proof of statement (iv): A direct application of statement (i) and (9).

In the sequel, $\rho S_d^{-1}$ stands for the sphere centred at the origin in $\mathbb{R}^d$ with radius $\rho > 0$.

Lemma 8 Let matrix $\Phi$ be defined as in sections 3.2 and 4.2. Given $\rho_0 < \rho_1$, consider any two random vectors $\Xi_0$ and $\Xi_1$ such that $\Phi(\Xi_0 - \theta_0)$ and $\Phi(\Xi_1 - \theta_0)$ are uniformly distributed on $\rho_0 S_d^{-1}$ and $\rho_1 S_d^{-1}$, respectively. Given any $Y \in \Xi_1$, any thresholding test with threshold height $\lambda_\gamma(\rho_0)$ on the Mahalanobis distance to $\theta_0$ is most powerful with size $\gamma$ for testing $\mathcal{H}_0: Y = \Xi_0 + X$ against $\mathcal{H}_1: Y = \Xi_1 + X$ with $X \sim N(0, \Sigma)$. The power of this test is $1 - \mathcal{R}(\rho_1, \lambda_\gamma(\rho_0))$.

Proof: Given any $y \in \mathbb{R}^d$, $\mathcal{L}(y) = f_{\Xi_i + X}(y)/f_{\Xi_i + X}(y)$ is the likelihood ratio for testing $\mathcal{H}_0$ against $\mathcal{H}_1$, where $f_{\Xi_i + X}$ is the probability density function (pdf) of $\Xi_i + X$, $i = 0, 1$. The Neyman-Pearson Lemma [12, Theorem 3.2.1, Sec. 3.2, p. 60] implies the existence of a most powerful test with size $\gamma$ for testing $\mathcal{H}_0$ against $\mathcal{H}_1$. This test accepts (resp. rejects) $\mathcal{H}_0$ if $\mathcal{L}(y) < \zeta$ (resp. $\mathcal{L}(y) > \zeta$), where $\zeta$ is such that $P[\mathcal{L}(\Xi_0 + X) > \zeta] = \gamma$. Under each hypothesis $\mathcal{H}_i$, $i = 0, 1$, $\Phi(Y - \theta_0) = \lambda_i + Z$ where $\lambda_i = \Phi(\Xi_i - \theta_0)$ is uniformly distributed on $\rho_i S_d^{-1}$ and $Z = \Phi(X) \sim \mathcal{N}(0, \Sigma)$. The pdf of $\Phi(Y - \theta_0)$ under each hypothesis being given by [10, Proposition V.1, p. 232], the
pdf’s \( f_{\Xi_0 + \mathbf{X}} \) and \( f_{\Xi_1 + \mathbf{X}} \) can then be calculated by using [53, Theorem 2.1.4, p. 57] and the standard change of variable [52, Theorem 17.2, p. 225]. Finally, by taking (2) into account, the reader will find that \( \mathcal{L}(y) = f_{\Xi_0 + \mathbf{X}}(y) / f_{\Xi_1 + \mathbf{X}}(y) \), where, for \( i = 0, 1 \), \( f_{\Xi_i + \mathbf{X}} \) is the pdf of the non-central \( \chi^2 \) distribution with \( d \) degrees of freedom and non-central parameter \( \rho_i^2 \). Since the family of non-central \( \chi^2 \) distributions with \( d \) degrees of freedom has monotone likelihood ratio with its non-central parameter \( \rho \), we can derive that \( P(\mathcal{T}_\iota'(\Xi_0 + X) = 1) \geq \gamma \). From (2) and [10, Proposition V.1, p. 232], we derive that \( P(\mathcal{T}_\iota'(\Xi_1 + X) = 1) = 1 - \mathcal{R}(\rho, \zeta) \) for \( i = 0, 1 \). Hence, the value of \( \zeta \) and the power of \( \mathcal{T}_\iota' \).

We now tackle the proof of Theorem 1. Let \( \mathcal{T}_\lambda \) be any thresholding test with threshold height \( \lambda^\ast = \lambda_T(\tau) \) on the Mahalanobis distance to \( \theta_0 \). It follows from statement (iii) of Lemma 7 that \( a_\Theta(\mathcal{T}_\lambda^\ast) = \gamma \). According to statement (i) of Lemma 7, given any \( \Theta \in \theta_1 \), there exists a support \( \mathcal{D} \) of \( P(\Theta - \theta_0)^{-1} \) such that \( P(\mathcal{T}_\lambda \mid \Theta = \theta_0) \) has constant power conditioned to \( \Theta = \theta_0 \) for any \( \rho \in \mathcal{D} \). Consider any \( \rho_0 \) such that \( \mathcal{T}_\lambda^\ast \) have constant power conditioned to \( \Theta = \theta_0 \). Choose two elements \( \Xi_0 \) and \( \Xi_1 \) of \( M(\Omega, \mathbb{R}^d) \) independent of \( \mathbf{X} \) such that \( \Xi_0 \in \mathcal{Y}_{\rho_0} \) (a-s) and \( \Xi_1 \in \mathcal{Y}_{\rho_1} \) (a-s) and take some random variable \( \varepsilon \) valued in \( (0, 1) \), independent of \( \Xi_0 \), \( \Xi_1 \) and \( \mathbf{X} \), with \( P(\varepsilon = 1) = 0, 1 \). We have \( P(\mathcal{T}(\Xi_0 + X) = 1) = P(\mathcal{T}(\Xi_1 + X) = 1) = \mathcal{R}(\rho, \zeta) \), where \( \Xi = (1 - \varepsilon) \Xi_0 + \varepsilon \Xi_1 \in \theta_1 \). Thereby, since \( \mathcal{T} \in \mathcal{K}_\rho^\iota \), we have \( P(\mathcal{T}(\Xi_0 + X) = 1) \leq \gamma \). According to Lemma 5, \( \beta_0(\Theta = \theta_0) = P(\mathcal{T}(\Xi_0 + X) = 1) \leq \gamma \). It follows from statement (i) of Lemma 7 that \( \mathcal{T} \) has level \( \gamma \) and power equal to \( \beta_0(\mathcal{T} \mid \Theta = \theta_0) \) for testing \( \mathcal{H}_0 : Y = \Xi_0 + X \) against \( \mathcal{H}_1 : Y = \Xi_1 + X \). This holds true for any \( \Xi_0 \in \mathcal{Y}_{\rho_0} \) (a-s) with \( i = 0, 1 \). We then choose \( \Xi_0 \) and \( \Xi_1 \) so that \( \Phi(\Xi_0 - \theta_0) \) and \( \Phi(\Xi_1 - \theta_0) \) are uniformly distributed on \( \rho_0 S^{d-1} \) and \( \rho_1 S^{d-1} \), respectively. According to Lemma 8, the thresholding test \( \mathcal{T}_{\lambda, \beta_0} \) with threshold height \( \lambda_T(\rho_0) \) on the Mahalanobis distance to \( \theta_0 \) is most powerful with size \( \gamma \) and power equal to \( 1 - \mathcal{R}(\rho_1, \lambda_T(\rho_0)) \) for testing \( \mathcal{H}_0 \) against \( \mathcal{H}_1 \). Therefore, \( 1 - \mathcal{R}(\rho_1, \lambda_T(\rho_0)) \geq \beta_0(\Theta = \theta_0) \). The rhs in the previous inequality tends to \( 1 - \mathcal{R}(\rho_1, \lambda^\ast) \) when \( \rho_0 \) tends to \( \tau \) by continuity of \( \mathcal{R}(\rho_1, \cdot) \) and \( \lambda_T(\cdot) \) (see Lemma 2). Since \( \mathcal{T}_{\lambda, \beta_0} \) has constant power conditioned to \( \Theta \in \mathcal{Y}_{\rho_1} \), statement (i) of Lemma 7 implies that \( 1 - \mathcal{R}(\rho_1, \lambda^\ast) = \beta_0(\mathcal{T}_{\lambda^\ast} \mid \Theta = \theta_0) \). Therefore, \( \beta_0(\mathcal{T}_{\lambda} \mid \Theta = \theta_0) \geq \beta_0(\mathcal{T} \mid \Theta = \theta_0) \) and \( \mathcal{T}_{\lambda} \) is UBCCP.

According to statement (iv) of Lemma 7,

\[
\beta_0^\iota(\mathcal{T}_{\lambda^\ast}) = \frac{1}{P(\Omega_1)} \int_{(\tau, \infty)} (1 - \mathcal{R}(\rho, \lambda^\ast)) P(\Theta - \theta_0)^{-1}(d\rho).
\]

For \( \rho > \tau \), \( 1 - \mathcal{R}(\rho, \lambda^\ast) \geq 1 - \mathcal{R}(\tau, \lambda^\ast) \). Thence, the unbiasedness of \( \mathcal{T}_{\lambda} \) since \( 1 - \mathcal{R}(\tau, \lambda^\ast) = \gamma \).
References


