



A limit theorem for processing signals that have unknown probability distributions and unknown probabilities of presence in white Gaussian noise

Dominique Pastor

► To cite this version:

Dominique Pastor. A limit theorem for processing signals that have unknown probability distributions and unknown probabilities of presence in white Gaussian noise. [Research Report] Dépt. Signal et Communications (Institut Mines-Télécom-Télécom Bretagne-UEB); Traitement Algorithmique et Matériel de la Communication, de l'Information et de la Connaissance (Institut Mines-Télécom-Télécom Bretagne-UEB). 2006, pp.37. hal-02316489

HAL Id: hal-02316489

<https://hal.science/hal-02316489>

Submitted on 15 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Collection
des Rapports
de Recherche
de l'ENST Bretagne



RR-2006007-SC

*A limit theorem for processing
signals that have unknown
probability distributions and
unknown probabilities
of presence in white Gaussian
noise*

*Un théorème limite pour
le traitement de signaux dont
les distributions de probabilité et
les probabilités de présence sont
inconnues dans un bruit blanc
Gaussien*

*Rapport Interne
GET / ENST Bretagne*

2006

Dominique PASTOR

A limit theorem for processing signals that
have unknown probability distributions and
unknown probabilities of presence
in white Gaussian noise

Un théorème limite pour le traitement de
signaux dont les distributions de probabilité et
les probabilités de présence sont inconnues
dans un bruit blanc Gaussien

Rapport Interne GET / ENST Bretagne

Dominique Pastor

Ecole Nationale Supérieure des Télécommunications de Bretagne,
CNRS TAMCIC (UMR 2872),
Technopôle de Brest Iroise, CS 83818
29238 BREST Cedex,
FRANCE
e-mail: dominique.pastor@enst-bretagne.fr

Foreword

This research report is an extended version of two previous reports, namely, “A limit theorem for sequences of independent random vectors with unknown distributions and its applications to non-parametric detection”, RR-2004001-SC, Collection des Rapports de Recherche de l’ENST Bretagne, November 2004, written by D. Pastor and R. Gay, and “On the detection of signals with unknown distributions and priors in white Gaussian noise”, RR-2006001-SC, Collection des Rapports de Recherche de l’ENST Bretagne, January 2006, written by D. Pastor.

In this new version, the proof of the main theoretical result is improved, more application fields and experimental results are given. Moreover, some typing faults have been corrected.

Avant-propos

Ce rapport de recherche est une version étendue de deux rapports antérieurs, à savoir, “A limit theorem for sequences of independent random vectors with unknown distributions and its applications to non-parametric detection”, RR-2004001-SC, Collection des Rapports de Recherche de l’ENST Bretagne, novembre 2004, écrit par D. Pastor et R. Gay, et “On the detection of signals with unknown distributions and priors in white Gaussian noise”, RR-2006001-SC, Collection des Rapports de Recherche de l’ENST Bretagne, janvier 2006, écrit par D. Pastor.

Dans cette nouvelle version, la démonstration du résultat théorique est améliorée, les applications ne se limitent pas à la détection de signaux et de nouveaux résultats expérimentaux sont présentés, quelques fautes de frappe ont été corrigées.

Abstract

Consider observations where signals are randomly present or absent in independent additive white Gaussian noise. We address the case where the probability distributions and the probabilities of presence of the signals are unknown. If the probabilities of presence of the signals are all less than or equal to some upper-bound in $[0, 1)$, a limit theorem established in this paper states that the noise standard deviation is the only positive real number that satisfies a specific convergence criterion when both the number of observations and the minimum amplitude of the signals tend to infinity. From this theorem, we derive an estimate of the noise standard deviation. This estimate is called the Essential Supremum Estimate (ESE). This estimate serves to adjust the Essential Supremum Test (EST), a binary hypothesis test introduced in this paper for the detection of signals with norms lower-bounded by some known real value and probabilities of presence less than or equal to one half.

We then present experimental results in the case of practical importance where the signals are independent two-dimensional random vectors uniformly distributed on a circle centred at the origin with known radius. Such random signals model modulated sinusoidal carriers whose amplitudes equal the radius of the circle and whose phases are uniformly distributed in $[0, 2\pi]$. We assume that the signals have their probabilities of presence less than or equal to one half. The empirical bias and the empirical Mean-Square Error of the ESE are computed; we also calculate the Binary Error Rate of the EST via a Monte-Carlo simulation when the EST is used to achieve the non-coherent detection of the carriers. Experimental results suggest that the asymptotic conditions of the limit theorem are not so constraining and can certainly be significantly relaxed in practice.

Keywords

Binary hypothesis testing, decision, estimation, likelihood theory, multivariate normal distribution, thresholding, signal processing, statistical test.

Résumé

Soient des observations où des signaux sont aléatoirement présents ou absents dans un bruit blanc Gaussien additif et indépendant. Nous Considérons le cas où les distributions de probabilité et les probabilités de présence des signaux sont inconnues. Si les probabilités de présence sont inférieures ou égales à une valeur de $[0, 1)$, un théorème limite établi dans ce papier énonce que l'écart-type du bruit est la seule valeur positive qui satisfait un critère spécifique de convergence quand le nombre d'observations et l'amplitude minimale des signaux tendent vers l'infini. Sur la base de ce théorème, nous proposons une estimation de l'écart-type du bruit. Cette estimée est utilisée pour ajuster un test d'hypothèse d'hypothèse binaire pour la détection de signaux dont les probabilités de présence sont inférieures ou égales à un demi et dont les normes admettent une borne inférieure connue.

Nous présentons ensuite des résultats expérimentaux dans le cas, important en pratique, où les signaux sont indépendants, de dimension 2 et uniformément répartis sur un cercle centré à l'origine et de rayon connu. De tels signaux peuvent être considérés comme des modèles pour des porteuses sinusoïdales modulées dont les amplitudes sont égales au rayon du cercle et dont les phases sont uniformément réparties sur $[0, 2\pi]$. On suppose que les probabilités de présence des signaux sont inférieures ou égales à $1/2$. Le biais et l'écart quadratique empiriques de l'estimateur sont calculés. Par une simulation de Monte-Carlo, nous évaluons aussi le Taux d'erreur Binaire de notre test lorsque celui-ci sert à détecter les porteuses. Les résultats expérimentaux suggèrent que les conditions asymptotiques du théorème limite ne sont pas si contraignantes dans la pratique.

Mots-clés

Décision, distribution multidimensionnelle Gaussienne, estimation, rapport de vraisemblance, test d'hypothèses, test statistique, traitement du signal.

Contents

1	Introduction	6
2	A limit theorem for sequences of independent random vectors that have unknown probability distributions	8
2.1	Notations and terminology	8
2.2	The theoretical result	10
3	Application to the estimation of the noise standard deviation and the detection of the signals	12
3.1	The Essential Supremum Estimate	12
3.2	The case of signals less present than absent: a thresholding function, a test for the signal detection and a search interval .	13
3.2.1	A thresholding function	13
3.2.2	The Essential Supremum Test	14
3.2.3	A search interval	15
4	Application to modulated sinusoidal carriers in AWGN: empirical bias and empirical MSE of the ESE, non-coherent detection of the carriers via the EST	16
4.1	Empirical bias and MSE of the ESE when observations are noisy modulated sinusoidal carriers	17
4.2	Application of the EST to the non-coherent detection of the carriers	17
5	Perspectives and extensions	20
6	Acknowledgement	24
	Appendix	24
A	Preliminary results	25
B	The case of large sample sizes	26
C	The case of large and small amplitudes	27
D	Combining large sample sizes and large or small amplitudes .	31
E	Uniqueness of the solution to equation	34

1 Introduction

In signal processing applications, we often meet sequences of d -dimensional real random observations that satisfy the following assumptions: each observation is either the sum of some random signal and independent noise or noise alone; noise is assumed to be Gaussian distributed with null mean and covariance matrix equal to $\sigma_0^2 \mathbf{I}_d$ where $\mathbf{I}_d$ is the $d \times d$ identity matrix. Summarizing these assumptions, we say that noise is “white and Gaussian” with standard deviation σ_0 and that each observation results from some signal randomly present or absent in independent and Additive White Gaussian Noise (AWGN).

In many applications, very little is known about the signals or most of their describing parameters ([3, Section I]), so that the probability distributions of the signals are partially or definitely unknown. This issue is met in Electronic (Warfare) Support Measure (ESM) systems faced with non-cooperant communications. This is also the case with passive sensors such as sonar systems that receive signals resulting from noise generated by motors and hull vibrations transmitted through a fluctuating environment. Another example is that of an echo received by a radar system from a target since this echo results from some kind of convolution between a known transmitted pulse and an unknown environment.

In situations such as those described above, the detection of the signals cannot be achieved via standard likelihood theory because the usual Bayes, minimax and Neyman-Pearson criteria require full knowledge of the probability distributions of these signals. Alternative solutions are then Wald’s tests [14], which have best constant (resp. average) power on a family of surfaces for making a decision between simple (resp. composite) hypotheses, non-parametric and robust detection ([10, section III.E]), as well as Generalized Likelihood Ratio Tests ([4]). Constant False Alarm Rate (CFAR) systems ([7]), standardly used in radar processing for detecting targets with a specified false alarm rate, typically derive from such alternative approaches.

In addition to our lack of prior knowledge about the signals, the noise standard deviation is sometimes unknown and must be estimated in order to process the observations. In general, the estimation is achieved by resorting to the physics of the problem. Two examples thereof are the following ones.

A radar target is detected when the observation is above a threshold height. In order to guarantee a specified false alarm probability, the threshold height must be chosen as a function of the noise standard deviation. When the latter can fluctuate, it is estimated so as to adjust the threshold height. By resorting to the physics of the radar, standard CFAR systems select observations that are considered as signal-free and called “reference

cells”; the noise standard deviation is then computed on the basis of these reference cells. The smaller the number of reference cells that contain radar echoes and not noise only, the better the estimate. Selecting the reference cells is no easy task ([7]).

A second example is the denoising of speech signals corrupted by independent AWGN since most standard filtering techniques must be adjusted with an estimate of the noise standard deviation when the latter is unknown. A standard solution involves estimating the noise standard deviation on the basis of periods of time considered as signal-free by a Voice Activity Detector, which, of course, can make errors. The noise standard deviation can also be estimated by computing the smallest eigenvalues of the noisy speech autocorrelation matrix but the computation of the eigenvalues may prove unstable depending on the matrix size.

Motivated by practical issues such as those described above, this paper introduces a theoretical result and, on the basis of this result, addresses the estimation of the noise standard deviation and the detection of the signals when the probability distributions of the signals are unknown.

More specifically, the theoretical result, namely theorem 2.1, is established in section 2. It concerns any sequence of mutually independent observations that are d -dimensional real random vectors where signals with unknown probability distributions are randomly present or absent with unknown probabilities of presence in independent AWGN. Theorem 2.1 does not assume that the available observations are identically distributed. It derives from a corollary of Kolmogorov’s classical strong limit theorem but not from usual generalizations of the central limit theorem such as the Lindeberg and the Lyapunov theorems.

On the basis of theorem 2.1, an estimate of the noise standard deviation is introduced in section 3. This estimate is called the Essential Supremum Estimate (ESE) for the crucial role played by the essential supremum norm in theorem 2.1. In the same section, and as a continuation of the results previously published in [8], we address the particular case where the amplitudes of the signals are above or equal to some known $\alpha \in [0, \infty)$ and the signals are less present than absent in the sense that their probabilities of presence are less than or equal to one half. Under these assumptions, we propose a test for the detection of these signals in AWGN with unknown standard deviation. This test, the Essential Supremum Test (EST), derives from the ESE and [8, Theorem VII.1].

Section 4 presents the empirical bias and the empirical Mean-Square Error (MSE) of the ESE, as well as the Binary Error Rate (BER) of the EST when the observations, and thus the signals and noise, are two-dimensional random vectors and the signals are uniformly distributed on the circle centred at

the origin with radius α . This model is relevant in many signal processing applications because the two components of each signal can be regarded as the in-phase and quadrature components of a modulated sinusoidal carrier whose amplitude is α and whose phase is uniformly distributed in $[0, 2\pi]$.

Section 5 concludes this paper by providing several suggestions regarding theoretical extensions and various applications, especially in radar processing, electronic warfare and speech processing.

2 A limit theorem for sequences of independent random vectors that have unknown probability distributions

2.1 Notations and terminology

Every random vector and every random variable encountered hereafter is assumed to be defined on the same probability space $(\Omega, \mathcal{B}, P)$ and for every $\omega \in \Omega$ by setting this random vector or variable to 0 on any negligible subset where it could be undefined. As usual, if a property $\mathcal{P}$ holds true almost surely, we write $\mathcal{P}$ (a-s).

As above, d stands for some fixed natural number and every random vector considered below is d -dimensional. The set of all the sequences of d -dimensional random vectors defined on Ω is denoted by $\mathcal{S}$.

In what follows, $\|\cdot\|$ is the standard Euclidean norm in $\mathbb{R}^d$. For every given random vector $Y : \Omega \rightarrow \mathbb{R}^d$ and any $\tau \in \mathbb{R}$, the notation $I(\|Y\| \leq \tau)$ stands for the indicator function of the event $\{\|Y\| \leq \tau\}$.

Given some positive real number σ_0 , we say that a sequence $X = (X_k)_{k \in \mathbb{N}}$ of d -dimensional real random vectors is a *d-dimensional white Gaussian noise* (WGN) with standard deviation σ_0 if the random vectors X_k , $k \in \mathbb{N}$, are mutually independent and identically Gaussian distributed with null mean vector and covariance matrix $\sigma_0^2 \mathbf{I}_d$. For the sake of brevity, we will henceforth write *independent* instead of *mutually independent*.

The *minimum amplitude* of an element $\Lambda = (\Lambda_k)_{k \in \mathbb{N}}$ of $\mathcal{S}$ is defined as the supremum $\mathbf{a}(\Lambda)$ of the set of those $\alpha \in [0, \infty]$ such that, for every natural number k , $\|\Lambda_k\|$ is larger than or equal to α (a-s):

$$\mathbf{a}(\Lambda) = \sup \{ \alpha \in [0, \infty] : \forall k \in \mathbb{N}, \|\Lambda_k\| \geq \alpha \text{ (a-s)} \}. \quad (1)$$

For every given $\Lambda = (\Lambda_k)_{k \in \mathbb{N}}$, the reader will easily verify that $\mathbf{a}(\Lambda)$ is finite, that $\|\Lambda_k\| \geq \mathbf{a}(\Lambda)$ for every $k \in \mathbb{N}$ and that, given $\alpha \in [0, \infty]$, $\mathbf{a}(\Lambda) \geq \alpha$ if and only if, for every $k \in \mathbb{N}$, $\|\Lambda_k\| \geq \alpha$ (a-s).

If f is some map of $\mathcal{S}$ into $\mathbb{R}$, we will say that the limit of f is $\ell \in \mathbb{R}$ when $\mathbf{a}(\Lambda)$ tends to ∞ and write that $\lim_{\mathbf{a}(\Lambda) \rightarrow \infty} f(\Lambda) = \ell$ if, for any positive real value η , there exists some $\alpha_0 \in (0, \infty)$ such that, for every $\alpha \geq \alpha_0$ and every $\Lambda \in \mathcal{S}$ such that $\mathbf{a}(\Lambda) \geq \alpha$, we have $|f(\Lambda) - \ell| \leq \eta$.

Given some non-negative real number a , let $L^a(\Omega, \mathbb{R}^d)$ stand for the set of those d -dimensional real random vectors $Y : \Omega \rightarrow \mathbb{R}^d$ for which $\mathbb{E}[\|Y\|^a] < \infty$. We hereafter deal with the set of those elements $\Lambda = (\Lambda_k)_{k \in \mathbb{N}}$ of $\mathcal{S}$ such that $\Lambda_k \in L^a(\Omega, \mathbb{R}^d)$ for every $k \in \mathbb{N}$ and $\sup_{k \in \mathbb{N}} \mathbb{E}[\|\Lambda_k\|^a]$ is finite. A suitable and standard notation for this set is $\ell^\infty(\mathbb{N}, L^a(\Omega, \mathbb{R}^d))$ because, although $\ell^\infty(\mathbb{N}, L^0(\Omega, \mathbb{R}^d))$ is trivially $\mathcal{S}$ and $L^a(\Omega, \mathbb{R}^d)$ is a Banach space only if $1 \leq a < \infty$, $L^a(\Omega, \mathbb{R}^d)$ remains a complete metric space even when $0 < a < 1$ ([5]).

Now, let $L^\infty(\Omega, \mathbb{R}^d)$ be the Banach space whose elements are the real random vectors $Y : \Omega \rightarrow \mathbb{R}^d$ whose essential supremum norm $\|Y\|_\infty$ is finite. Consider the set $\ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$ of those sequences $\Lambda = (\Lambda_k)_{k \in \mathbb{N}}$ of $\mathcal{S}$ such that $\Lambda_k \in L^\infty(\Omega, \mathbb{R}^d)$ for every $k \in \mathbb{N}$ and $\sup_{k \in \mathbb{N}} \|\Lambda_k\|_\infty$ is finite. It is a Banach space once it is endowed with the norm $\|\cdot\|_\infty$ defined for any element $\Lambda = (\Lambda_k)_{k \in \mathbb{N}}$ of $\mathcal{S}$ by $\|\Lambda\|_\infty = \sup_{k \in \mathbb{N}} \|\Lambda_k\|_\infty$. For every given $k \in \mathbb{N}$, $\|\Lambda_k\|$ is then less than or equal to $\|\Lambda\|_\infty$ (a-s). If $\alpha \in [0, \infty]$, $\|\Lambda\|_\infty \leq \alpha$ if and only if, for every $k \in \mathbb{N}$, $\|\Lambda_k\| \leq \alpha$ (a-s). For every $k \in \mathbb{N}$, $\mathbf{a}(\Lambda) \leq \|\Lambda_k\| \leq \|\Lambda\|_\infty$ (a-s).

If f is a map of $\ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$ into $\mathbb{R}$, we say that the limit of f is $\ell \in \mathbb{R}$ when $\|\Lambda\|_\infty$ tends to 0 and write that $\lim_{\|\Lambda\|_\infty \rightarrow 0} f(\Lambda) = \ell$ if, given any positive real value η , there exists some $\alpha_0 \in (0, \infty)$ such that, for every real value $\alpha \leq \alpha_0$ and every $\Lambda \in \ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$ such that $\|\Lambda\|_\infty \leq \alpha$, we have $|f(\Lambda) - \ell| \leq \eta$.

Given two (possibly extended) real numbers a and b such that $0 \leq a \leq b \leq \infty$, note that

$$\ell^\infty(\mathbb{N}, L^b(\Omega, \mathbb{R}^d)) \subset \ell^\infty(\mathbb{N}, L^a(\Omega, \mathbb{R}^d)). \quad (2)$$

We call a **thresholding function** any non-decreasing continuous and positive real function $\theta : [0, \infty) \rightarrow (0, \infty)$ such that $\theta(\rho) = C\rho + \gamma(\rho)$ where $0 < C < 1$, $\gamma(\rho)$ is positive for sufficiently large values of ρ and $\lim_{\rho \rightarrow \infty} \gamma(\rho) = 0$; given any $q \in [0, \infty)$, Υ_q will stand for the map of $[0, \infty)$ into $[0, \infty)$ defined for every $x \in [0, \infty)$ by

$$\Upsilon_q(x) = \int_0^x t^{q+d-1} e^{-t^2/2} dt. \quad (3)$$

2.2 The theoretical result

Let $\varepsilon = (\varepsilon_k)_{k \in \mathbb{N}}$ be a sequence of random variables valued in $\{0, 1\}$, $\Lambda = (\Lambda_k)_{k \in \mathbb{N}}$ stand for some sequence of d-dimensional real random vectors and $X = (X_k)_{k \in \mathbb{N}}$ be some d-dimensional WGN with standard deviation σ_0 . Suppose that, for every $k \in \mathbb{N}$, Λ_k , X_k and ε_k are independent.

Theorem 2.1 stated below concerns the sequence $U = \varepsilon\Lambda + X$, that is, the sequence $U = (U_k)_{k \in \mathbb{N}}$ such that, for every $k \in \mathbb{N}$, $U_k = \varepsilon_k\Lambda_k + X_k$. In the theorem statement, the random vectors U_k , $k \in \mathbb{N}$, are assumed to be independent. Before stating the result, it is worth mentioning the following.

The sequence U basically models a sequence of independent observations where random signals are either present or absent in independent AWGN: noise is modelled by the sequence X ; given $k \in \mathbb{N}$, Λ_k stands for some possible random signal and ε_k models the possible occurrence of Λ_k . The two possible values for ε_k form a hypothesis pair according to which the observation is distributed: the null hypothesis is that $\varepsilon_k = 0$ and the alternative one is $\varepsilon_k = 1$. Even though this model introduces *a priori* probabilities of presence and absence, the importance of these probabilities will be significantly reduced by theorem 2.1 since this result assumes that these probabilities are upper-bounded by a positive real number less than one.

We now state the following theorem whose proof is given in appendix .

Theorem 2.1 *Let $U = (U_k)_{k \in \mathbb{N}}$ be some element of $\mathcal{S}$ such that $U = \varepsilon\Lambda + X$ where $\Lambda = (\Lambda_k)_{k \in \mathbb{N}}$, $X = (X_k)_{k \in \mathbb{N}}$ and $\varepsilon = (\varepsilon_k)_{k \in \mathbb{N}}$ are an element of $\mathcal{S}$, some d-dimensional WGN with standard deviation σ_0 and a sequence of random variables valued in $\{0, 1\}$ respectively.*

Assume that

(Ass. 1) *for every $k \in \mathbb{N}$, Λ_k , X_k and ε_k are independent;*

(Ass. 2) *the random vectors U_k , $k \in \mathbb{N}$, are independent;*

(Ass. 3) *the set of priors $\{P(\{\varepsilon_k = 1\}) : k \in \mathbb{N}\}$ has an upper bound $p \in [0, 1)$ and the random variables ε_k , $k \in \mathbb{N}$, are independent;*

(Ass. 4) *there exists some $\nu \in (0, \infty]$ such that $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$.*

Let r and s be any two non-negative real numbers such that $0 \leq s < r \leq \nu/2$. Given some natural number m and any pair (σ, T) of positive real

numbers, define the random variable $\Delta_m(\sigma, T)$ by

$$\Delta_m(\sigma, T) = \left| \frac{\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \sigma T)}{\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \sigma T)} - \sigma^{r-s} \frac{\Upsilon_r(T)}{\Upsilon_s(T)} \right|. \quad (4)$$

Then, given any thresholding function θ ,

(i) σ_0 is the unique positive real number σ such that, for every $\beta_0 \in (0, 1]$,

$$\lim_{\mathbf{a}(\Lambda) \rightarrow \infty} \left\| \limsup_m \Delta_m(\sigma, \beta \theta(\mathbf{a}(\Lambda)/\sigma)) \right\|_\infty = 0 \quad (5)$$

uniformly in $\beta \in [\beta_0, 1]$;

(ii) if, in addition, $\nu = \infty$, σ_0 is such that, for every $\beta_0 \in (0, 1]$,

$$\lim_{\|\Lambda\|_\infty \rightarrow 0} \left\| \limsup_m \Delta_m(\sigma_0, \beta \theta(\|\Lambda\|_\infty/\sigma_0)) \right\|_\infty = 0 \quad (6)$$

uniformly in $\beta \in [\beta_0, 1]$.

Throughout the rest of the paper, the notations introduced in the foregoing statement are used with the same meaning and always under assumptions **(Ass. 1-4)**.

For $\tau \in [0, \infty)$, the ratio $\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \tau) / \sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \tau)$ is defined everywhere for the following reason. Let $x_1, x_2, \dots, x_m$ be m real numbers. If there exists at least one natural number $k \in \{1, \dots, m\}$ such that $x_k \neq 0$, the finiteness of the ratio $\sum_{k=1}^m |x_k|^r / \sum_{k=1}^m |x_k|^s$ is trivial. Since $r > s$, the definition of this ratio is then extended by continuity by setting $\sum_{k=1}^m |x_k|^r / \sum_{k=1}^m |x_k|^s = 0$ if $(x_1, \dots, x_m) = (0, \dots, 0)$.

Theorem 2.1 concerns the positive solutions of Eq. (5) only because $\sigma = 0$ trivially satisfies Eq. (5) regardless of the specific convergence involved. Straightforwardly, Eq. (5) is also satisfied for all $\sigma \in [0, \infty)$ when $r = s \geq 0$. This explains why it is assumed that $r > s \geq 0$. Finally, statement **(i)** is of more interest than statement **(ii)**. First, the former specifies the uniqueness of the solution in σ to Eq. (5) whereas this uniqueness is not stated by the latter and remains an open question; second, the assumption $\nu = \infty$ is strong since, according to Eq. (2), $\ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$ is a subset of every $\ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$, $\nu \in [0, \infty)$.

3 Application to the estimation of the noise standard deviation and the detection of the signals

3.1 The Essential Supremum Estimate

Basically, Eq. (5) means the following. Let $\nu \in (0, \infty)$, $\beta_0 \in (0, 1]$ and consider some positive real value η . Then, there exists some positive real number α_0 with the following property: for every real number α larger than or equal to α_0 , every element Λ of $L^\nu(\Omega, \mathbb{R}^d)$ whose minimum amplitude is larger than or equal to α , and every $\beta \in [\beta_0, 1]$, there exists a negligible subset Ω_β of Ω such that, for every $\omega \in \Omega \setminus \Omega_\beta$, $\Delta_m(\sigma_0, \beta\theta(\alpha/\sigma_0))(\omega) \leq \eta$ if the sample size m is large enough.

Consider now a finite subset $\mathcal{I}$ of $[\beta_0, 1]$, a real number $\alpha \geq \alpha_0$ and an element Λ of $L^\nu(\Omega, \mathbb{R}^d)$ whose minimum amplitude is larger than or equal to α . According to the foregoing, every $\beta \in \mathcal{I}$ can be assigned to a negligible subset $\Omega_\beta \subset \Omega$ outside which $\Delta_m(\sigma_0, \beta\theta(\alpha/\sigma_0))$ is less than or equal to η when m is large enough. The finite union of these negligible subsets is also a negligible subset of Ω and, outside this union, $\max_{\beta \in \mathcal{I}} \Delta_m(\sigma_0, \beta\theta(\alpha/\sigma_0))$ is less than or equal to η and, thus, small.

The foregoing suggests proceeding as follows to estimate σ_0 , given m observations $U_k = \varepsilon_k \Lambda_k + X_k$ where $\Lambda \in L^\nu(\Omega, \mathbb{R}^d)$ with $\mathbf{a}(\Lambda) \geq \alpha$: choose a thresholding function θ , two values r and s such that $0 \leq s < r \leq \nu/2$, a finite subset $\mathcal{I}$ of $(0, 1]$ and a search interval included in $(0, \infty)$; estimate σ_0 by a possibly local minimum $\hat{\sigma}_0$ of the discrete cost $\max_{\beta \in \mathcal{I}} \Delta_m(\sigma, \beta\theta(\alpha/\sigma))$ when σ ranges over the search interval; the value $\hat{\sigma}_0$ is henceforth called the Essential Supremum Estimate (ESE) of the noise standard deviation. In the rest of this paper, the notation $\hat{\sigma}_0$ will always stand for this estimate.

A great number of such estimates exist, depending on our choice for the thresholding function θ , the set $\mathcal{I}$, the values r and s and the search interval. Henceforth, the ESE is computed with the following choices for r , s and $\mathcal{I}$.

In most applications of importance, the signals have “finite energy” in the sense that the second order moments $\mathbb{E}[\|\Lambda_k\|^2]$, $k \in \mathbb{N}$, are finite. With the notations used so far, this corresponds to the assumption that $\nu = 2$. Since the values r and s must verify the inequalities $0 \leq s < r \leq 1$, a rather natural choice is $s = 0$ and $r = 1$.

We set $\mathcal{I} = \{\beta_\ell = \ell/L : \ell = 1, 2, \dots, L\}$ where L is some natural number. On the basis of preliminary experimental results such as those of section 4, we take $L = m$ because it seems a reasonable trade-off between computational load and accuracy.

Under the “finite energy” assumption and with the choice proposed above for $\mathcal{I}$, the ESE considered henceforth is a possibly local minimum $\hat{\sigma}_0$ of

$$\sup_{\ell \in \{1, \dots, L\}} \left\{ \left| \frac{\sum_{k=1}^m \|U_k\| I(\|U_k\| \leq \beta_\ell \sigma \theta(\alpha/\sigma))}{\sum_{k=1}^m I(\|U_k\| \leq \beta_\ell \sigma \theta(\alpha/\sigma))} - \sigma \frac{\int_0^{\beta_\ell \theta(\alpha/\sigma)} t^d e^{-t^2/2} dt}{\int_0^{\beta_\ell \theta(\alpha/\sigma)} t^{d-1} e^{-t^2/2} dt} \right| \right\} \quad (7)$$

when σ ranges over a search interval to specify. Any minimization routine for scalar bounded non-linear functions is suitable for the computation of $\hat{\sigma}_0$. For instance, the experimental results presented in the sequel were obtained with the MATLAB routine `fminbnd.m` based on parabolic interpolation ([11]). In the next section, we propound a thresholding function and a search interval under the assumption that the signals are less present than absent in the sense that their probabilities of presence are less than or equal to one half.

3.2 The case of signals less present than absent: a thresholding function, a test for the signal detection and a search interval

3.2.1 A thresholding function

For every given $\rho \in [0, \infty)$, let $\xi(\rho)$ be the unique positive solution for x in the equation

$${}_0F_1(d/2; \rho^2 x^2/4) = e^{\rho^2/2}, \quad (8)$$

where ${}_0F_1$ is the generalized hypergeometric function ([6, p. 275]).

The map ξ is a thresholding function with $C = 1/2$. This follows from the three following facts: first, the increasingness of the generalized hypergeometric function ${}_0F_1(d/2; \cdot)$ and that of the function of ρ in the right-hand side of Eq. (8) imply the increasingness of ξ ; second, ξ is a positive map because $\xi(0) = \sqrt{d}$; third, $\xi(\rho) = (\rho/2) + (\log 2/\rho)(1 + \psi_1(\rho))$ if $d = 1$ and $\xi(\rho) = (\rho/2) + (d-1)(\log \rho/\rho)(1 + \psi_d(\rho))$ otherwise, with $\lim_{\rho \rightarrow +\infty} \psi_d(\rho) = 0$ for all $d \geq 1$ ([8, Lemma VI.1]).

When the signals are less present than absent, the use of this thresholding function for the computation of the ESE is particularly relevant. We now explain why.

Given a thresholding function θ and under the assumptions of theorem 2.1, it follows from lemma C.1 that the probability $P(\{\|\Lambda_k + X_k\| \leq \sigma_0 \theta(a(\Lambda)/\sigma_0)\})$ tends to 0 uniformly in k when $a(\Lambda)$ tends to infinity and

that the faster the convergence of $P(\{\|\Lambda_k + X_k\| \leq \sigma_0 \theta(\mathbf{a}(\Lambda)/\sigma_0)\})$ to 0, the better.

Now, assume that the signals are less present than absent. We keep this assumption in the rest of this section. Given $h \in [0, \infty)$, let the *thresholding test* $\mathcal{T}_h$ with *threshold height* h stand for the map $I_{[h, \infty)}(\|\cdot\|)$, that is, the map such that $\mathcal{T}_h(u) = 1$ if $\|u\| \geq h$ and $\mathcal{T}_h(u) = 0$ otherwise. Then, according to [8, Theorem VII-1], for every $k \in \mathbb{N}$, the probability of error of the test $\mathcal{T}_{\sigma_0 \xi(\mathbf{a}(\Lambda)/\sigma_0)}$ for the decision about the value of ε_k is less than or equal to $V(\mathbf{a}(\Lambda)/\sigma_0)$ where, for every $\rho \in [0, \infty)$,

$$V(\rho) = \frac{e^{-\rho^2/2}}{2^{d/2}\Gamma(d/2)} \int_0^{\xi(\rho)} e^{-t^2/2} t^{d-1} {}_0F_1(d/2; \rho^2 t^2/4) dt \\ + \frac{1}{2} \left[1 - \frac{1}{2^{d/2-1}\Gamma(d/2)} \int_0^{\xi(\rho)} e^{-t^2/2} t^{d-1} dt \right], \quad (9)$$

and where Γ is, as usual, the standard (complete) gamma function. The function $V(\rho)$ decreases very rapidly when ρ increases. Therefore, when $\mathbf{a}(\Lambda)$ increases, the probability of error of the test $\mathcal{T}_{\sigma_0 \xi(\mathbf{a}(\Lambda)/\sigma_0)}$ and thus the probability $P(\{\|\Lambda_k + X_k\| \leq \sigma_0 \xi(\mathbf{a}(\Lambda)/\sigma_0)\})$ tend to 0 uniformly in k and, above all, very rapidly. This is what we need. Thus, we recommend and use the thresholding function ξ to compute the ESE when signals are less present than absent.

3.2.2 The Essential Supremum Test

Given $\alpha \in [0, \infty)$, the test $I(\|\cdot\| \geq \hat{\sigma}_0 \xi(\alpha/\hat{\sigma}_0))$ assigns, to a given d -dimensional observation, the value 1 if the norm of the observation is above or equal to $\hat{\sigma}_0 \xi(\alpha/\hat{\sigma}_0)$ and 0 otherwise. This test is henceforth called the *Essential Supremum Test* (EST). Strictly speaking, it is not a thresholding test since the threshold height to which the norm of a given observation is compared is the random variable $\hat{\sigma}_0 \xi(\alpha/\hat{\sigma}_0)$. However, with a slight abuse of language, we will denote it by $\mathcal{T}_{\hat{\sigma}_0 \xi(\alpha/\hat{\sigma}_0)}$.

Assume that the signals are less present than absent and have norms larger than or equal to α . If α and the sample size m are large enough, the ESE $\hat{\sigma}_0$ should be a good estimate of σ_0 ; therefore, we can expect that the performance of the EST approaches that of the test $\mathcal{T}_{\sigma_0 \xi(\alpha/\sigma_0)}$ and, thus, that the probability of error of the EST does not significantly exceed $V(\alpha/\sigma_0)$. If the probabilities of presence of the signals all equal one half and the signals are uniformly distributed on αS^{d-1} , which is the sphere centred at the origin with radius α in $\mathbb{R}^d$, the probability of error of the test $\mathcal{T}_{\sigma_0 \xi(\alpha/\sigma_0)}$ is $V(\alpha/\sigma_0)$ for the detection of any of these signals (see [8, Theorem VII.1]). Hence, in

this particular case, the probability of error of the EST should become close to $V(\alpha/\sigma_0)$ when m and α are large enough.

The computation of the probability of error of the EST is still an open issue and the foregoing claims remain intuitive only. However, the experimental results presented in section 4 agree with them.

3.2.3 A search interval

We conclude this section by presenting a search interval for the computation of the ESE. This search interval is defined so that the EST satisfies two constraints. In what follows, $U_{[k]}$, $k = 1, 2, \dots, m$, stands for the sequence of observations $U_1, U_2, \dots, U_m$, sorted by increasing norm.

The left endpoint of the search interval is chosen so as to guarantee that the EST accepts a minimum number of times the null hypothesis, i.e. the hypothesis that no signal is present. More specifically, let Q be some positive real number less than or equal to $1 - \frac{m}{4(m/2-1)^2}$. According to Bienaymé-Chebyshev's inequality, the probability that the number of observations due to noise alone is above $k_{\min} = m/2 - hm$ is larger than or equal to Q if $h = 1/\sqrt{4m(1-Q)}$. The lower bound of the search interval is then set to $\sigma_{\min} = \|U_{[k_{\min}]}\|/\sqrt{d}$. For every non-negative real value ρ , $\xi(\rho) \geq \sqrt{d}$ ([8]). Therefore, an ESE above or equal to $\sigma_{\min}$ guarantees a minimum of $k_{\min}$ acceptances of the null hypothesis by the EST, whatever the value of the minimum amplitude of the signals.

The right endpoint is computed so as to guarantee that the EST accepts at least once the alternative hypothesis, i.e. the hypothesis that some signal is present. Hence, the upper bound of the search interval is fixed to $\sigma_{\max} = \|U_{[m]}\|/\sqrt{d}$. An estimate $\hat{\sigma}_0$ larger than $\sigma_{\max}$ would imply that $\|U_k\| \leq \hat{\sigma}_0 \xi(\alpha/\hat{\sigma}_0)$ for every $k = 1, 2, \dots, m$, since $\xi(\rho) \geq \sqrt{d}$ for all $\rho \in [0, \infty)$. The outcome of the EST could then be that no signal is present whereas the full absence of signals amongst m observations is hardly probable when m is large, provided that the probabilities of presence are not too small. The experimental results presented in the sequel were obtained by choosing $Q = \lfloor (1 - \frac{m}{4(m/2-1)^2}) \times 1000 \rfloor / 1000$ where $\lfloor x \rfloor$ rounds the real value x to the nearest integer smaller than or equal to x .

4 Application to modulated sinusoidal carriers in AWGN: empirical bias and empirical MSE of the ESE, non-coherent detection of the carriers via the EST

The statistical properties of the ESE and the probabilities of error of the first and second kind of the EST are not known yet. Hence, in this section, we will restrict ourselves to some experimental results concerning the following case study.

The observations considered below, and thus the signals and noise, are two-dimensional random vectors ($d = 2$), or equivalently, complex random variables. We assume that the norms of the random signals all equal some known value α , and that each random signal is, in fact, uniformly distributed on the circle αS^1 centred at the origin with radius α . Therefore, the two components of each signal can be regarded as the in-phase and quadrature components of some modulated sinusoidal carrier with amplitude α and phase uniformly distributed in $[0, 2\pi]$. This model is relevant in many applications.

In what follows, we address the case where the probabilities of presence of the carriers all equal some unknown $p \in (0, 1/2]$. Consequently, the carriers are less present than absent and we make use of the results of section 3.2. In particular, note that the expression of the thresholding function ξ introduced in paragraph 3.2.1 simplifies in the two-dimensional case considered in this section: according to [1, Eq. 9.6.47, p. 377], for every $x \in [0, \infty)$,

$$I_0(x) = {}_0F_1(1; x^2/4), \quad (10)$$

where I_0 is the zeroth-order modified Bessel function of the first kind. Therefore, by setting $d = 2$ in Eq. (8), we obtain, in the two-dimensional case and for every given $\rho \in [0, \infty)$

$$\xi(\rho) = I_0^{-1}(e^{\rho^2/2})/\rho. \quad (11)$$

In the next subsection, we present the absolute value $|\text{Bias}|$ of the empirical bias and the empirical MSE of the ESE when this estimate is obtained by minimizing Eq. (7) when the search interval is computed according to paragraph 3.2.3 and the thresholding function is given by Eq. (11).

In subsection 4.2, the EST of paragraph 3.2.2 is used to make a decision about the value of every given ε_k , $k = 1, 2, \dots, m$. This amounts to achieving the so-called “non-coherent detection of the modulated sinusoidal carriers” ([10, Example III.B.5, p. 65]), the modulated carriers being represented by the random signals Λ_k , $k = 1, 2, \dots, m$. This detection problem is standard and arises in many applications.

4.1 Empirical bias and MSE of the ESE when observations are noisy modulated sinusoidal carriers

Given N independent copies of the sequence $U_1, U_2, \dots, U_m$, where the observations $U_1, U_2, \dots, U_m$ obey the assumptions introduced above, we can obtain N samples of the ESE and, then, calculate the empirical bias and the empirical MSE of the ESE. In this respect, tables 1, 2 and 3 display the absolute value of the empirical bias and the empirical MSE on the basis of $N = 500$ samples of the ESE, when the signals have the same amplitude α and the same probability of presence p equal to 0.1, 0.3 and 0.5, respectively. These results comply with theorem 2.1 since the larger the amplitude α , the smaller the bias and the MSE of the estimate. These experimental results also show that the ESE quite rapidly becomes a reasonably good estimate of σ_0 with increasing values for α and m .

α		0.5	1	1.5	2	2.5	3	3.5	4
m = 100	Bias	0.212	0.119	0.095	0.063	0.044	0.039	0.038	0.048
	MSE	0.166	0.095	0.068	0.038	0.025	0.020	0.015	0.012
m = 200	Bias	0.136	0.114	0.048	0.012	0.011	0.028	0.019	0.021
	MSE	0.088	0.056	0.041	0.037	0.026	0.014	0.011	0.008
m = 400	Bias	0.101	0.056	0.023	0.019	0.016	0.003	0.001	0.009
	MSE	0.051	0.034	0.019	0.011	0.007	0.005	0.004	0.003
m = 800	Bias	0.060	0.044	0.022	0.001	0.013	0.001	0.003	0.013
	MSE	0.020	0.015	0.009	0.004	0.004	0.003	0.002	0.002

Table 1: Absolute value of the empirical bias and empirical MSE of the ESE when signals are uniformly distributed on the circle centred at the origin with radius α and have a probability of presence equal to 0.1.

4.2 Application of the EST to the non-coherent detection of the carriers

Since we assume that each signal is uniformly distributed on αS^1 and that the probabilities of presence of the signals all equal some $p \in (0, 1/2]$, the probability of error of the test $\mathcal{T}_{\sigma_0 \xi(\alpha/\sigma_0)}$ is less than or equal to $V(\alpha/\sigma_0)$ for making a decision about the value of ε_k , given $k \in \mathbb{N}$ (see paragraph 3.2.2).

The following fact, which can easily be derived from [10, Example II.E.1, p. 33, & Section III.B, p. 69] and is a straightforward consequence of [8, Theorem VII.1] also holds true. If $P(\{\varepsilon_k = 1\}) = 1/2$ and since Λ_k is

α		0.5	1	1.5	2	2.5	3	3.5	4
m = 100	Bias	0.148	0.102	0.019	0.022	0.030	0.015	0.013	0.024
	MSE	0.140	0.113	0.077	0.057	0.043	0.031	0.021	0.016
m = 200	Bias	0.118	0.160	0.019	0.044	0.031	0.022	0.018	0.017
	MSE	0.087	0.056	0.041	0.038	0.026	0.014	0.011	0.008
m = 400	Bias	0.054	0.016	0.044	0.056	0.056	0.051	0.022	0.011
	MSE	0.042	0.028	0.023	0.019	0.013	0.010	0.008	0.004
m = 800	Bias	0.047	0.005	0.039	0.084	0.080	0.042	0.019	0.008
	MSE	0.022	0.013	0.011	0.013	0.013	0.006	0.003	0.002

Table 2: Absolute value of the empirical bias and empirical MSE of the ESE when signals are uniformly distributed on the circle centred at the origin with radius α and have a probability of presence equal to 0.3.

α		0.5	1	1.5	2	2.5	3	3.5	4
m = 100	Bias	0.142	0.780	0.071	0.118	0.178	0.130	0.050	0.023
	MSE	0.140	0.113	0.077	0.057	0.043	0.031	0.021	0.016
m = 200	Bias	0.105	0.006	0.117	0.192	0.199	0.114	0.062	0.017
	MSE	0.078	0.067	0.066	0.093	0.102	0.051	0.025	0.015
m = 400	Bias	0.067	0.031	0.144	0.203	0.204	0.149	0.071	0.017
	MSE	0.051	0.039	0.052	0.073	0.076	0.042	0.017	0.006
m = 800	Bias	0.025	0.062	0.165	0.208	0.222	0.146	0.072	0.040
	MSE	0.017	0.021	0.044	0.058	0.060	0.031	0.010	0.005

Table 3: Absolute value of the empirical bias and empirical MSE of the ESE when signals are uniformly distributed on the circle centred at the origin with radius α and have a probability of presence equal to 0.5.

uniformly distributed on αS^1 , the test $\mathcal{T}_{\sigma_0 \xi(\alpha/\sigma_0)}$, where ξ is given by Eq. (11), is the Bayesian likelihood ratio test guaranteeing the least probability of error among all the possible tests for deciding about the value of ε_k . The probability of error of this test is then

$$V(\rho) = \frac{1}{2}e^{-\rho^2/2} \int_0^{\xi(\rho)} te^{-t^2/2} I_0(\rho t) dt + \frac{1}{2}e^{-\xi(\rho)^2/2}, \quad (12)$$

which follows from Eqs. (9) with $d = 2$, (10) and (11).

Since the empirical bias and the empirical MSE of the ESE presented above suggest that the ESE is a reasonably good estimate of σ_0 when α and m are large enough, we can reasonably expect that, for large values of α and m , the EST behaves as described in section 3.2.2: if $P(\{\varepsilon_k = 1\}) \leq 1/2$, the probability of error of the EST should not exceed $V(\alpha/\sigma_0)$ and become rather close to $V(\alpha/\sigma_0)$ when $P(\{\varepsilon_k = 1\}) = 1/2$.

In the absence of theoretical results about the probability of error of the EST, even under the assumptions made in this section, the foregoing claim remains intuitive only. This is why, for different values of α , m and p , we computed the BER of the EST via a Monte-Carlo simulation: for every $k \in \{1, 2, \dots, m\}$, the EST was used to make a decision about the value of ε_k and, since the m observations and the m successive decisions made by the EST are independent, we calculated the BER of the EST by counting the number of errors made by the EST. More specifically, the several Monte-Carlo simulations yielding the results presented below were achieved according to the following protocol, standardly employed to assess telecommunication systems.

Given α , m and p , independent trials of m observations each were carried out until two conditions were fulfilled. First, the total number N_e of errors made by the EST had to be above or equal to a specified number N . Second, a minimum number M of trials had to be achieved. The reason for this second condition is the following one. Inasmuch as the decision about the presence or the absence of signals is made on the observations used for estimating σ_0 , the accuracy of the estimate affects m decisions at one go. This effect is then reduced by performing a minimum number of trials. If j is the first trial number larger than or equal to M and for which the total number of errors N_e becomes larger than or equal to N , the BER is then defined as the ratio $N_e/(j \times m)$. The simulations presented below were performed with $\sigma_0 = 1$. The pre-specified number of errors was fixed to $N = 400$ and the minimum number of trials was set to $M = 150$.

When the signals have the same amplitude α and the same probability of presence $p \leq 1/2$, the BER's of the EST for different values of α , p and

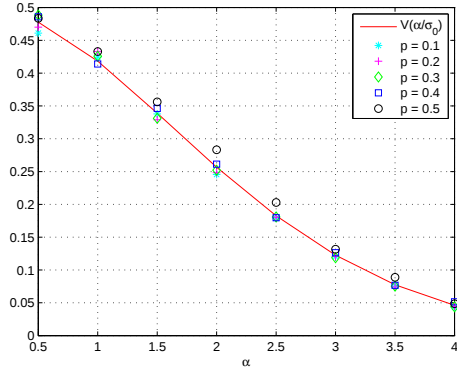
m are those of figure 1. These results show that, under the experimental conditions considered in this section, the EST behaves as expected: when $p = 1/2$, the larger α and m , the smaller the difference between the BER of the EST and $V(\alpha/\sigma_0)$; yet for $p = 1/2$, for every tested sample size, the difference between the BER of the EST and $V(\alpha/\sigma_0)$ becomes rapidly small when the value of α increases; when p is less than one half, the BER of the EST remains less than or equal to $V(\alpha/\sigma_0)$.

Now, assume that the amplitudes of the signals all equal $\kappa\alpha$ where κ is some real number larger than 1 and that the probabilities of presence of these signals all equal some $p \in (0, 1/2]$. The expected behaviour of the EST (see section 3.2.2) is that the probability of error of the EST remains less than or equal to $V(\alpha/\sigma_0)$ if α and m are large enough. Monte-Carlo simulations of the same type as those described above were carried out for different values of α and p . When $\kappa = 1.1220$, that is, when the actual norms of the signals are one dB larger than the value α with respect to which the EST is adjusted, the results obtained are then those of figure 2: the BER of the EST remains, as expected, less than or equal to $V(\alpha/\sigma_0)$ for every α , every m and every p .

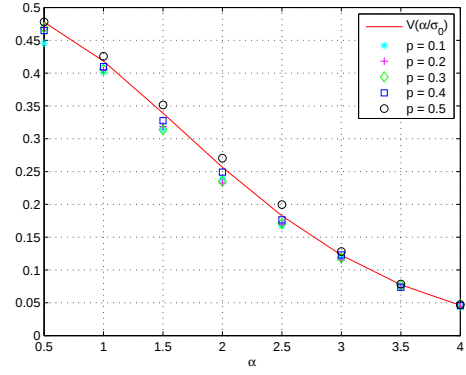
5 Perspectives and extensions

This paper has presented a limit theorem, namely theorem 2.1, that concerns independent observations deriving from the random presence or absence of signals whose probability distributions and prior probabilities of presence are unknown in independent AWGN. On the basis of this theorem, we have introduced the Essential Supremum Estimate (ESE), an estimate of the noise standard deviation. The Essential Supremum Test (EST) is a binary hypothesis test also introduced in this paper for the detection of signals that are less present than absent and whose norms are equal to or above some known $\alpha \in [0, \infty)$. The ESE and the EST have been tested in the particular case, relevant in many applications, where the signals are two-dimensional random vectors representing modulated sinusoidal carriers whose amplitudes all equal some known $\alpha \in [0, \infty)$ and whose phases are uniformly distributed in $[0, 2\pi]$. Experimental results obtained when the signals are less present than absent suggest that, under the assumptions made for the signals, the convergence stated by theorem 2.1 is fast.

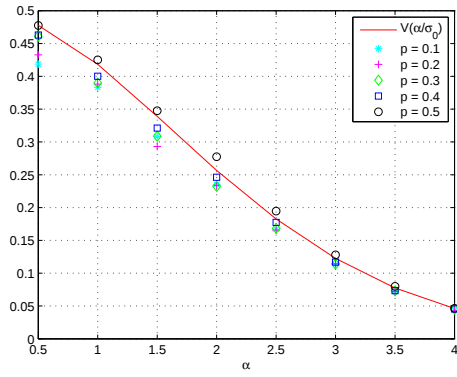
This is the reason why complementary theoretical studies in progress are aimed at analysing to what extent the asymptotic conditions of theorem 2.1 can actually be relaxed, as well as studying the statistical properties of the ESE and the performance of the EST. In particular, the influence of L , r , s and the thresholding function must be addressed.



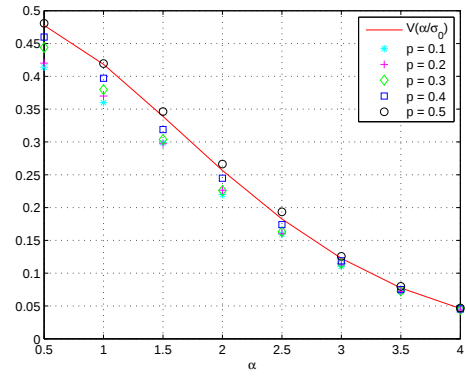
(a) $m = 100$



(b) $m = 200$

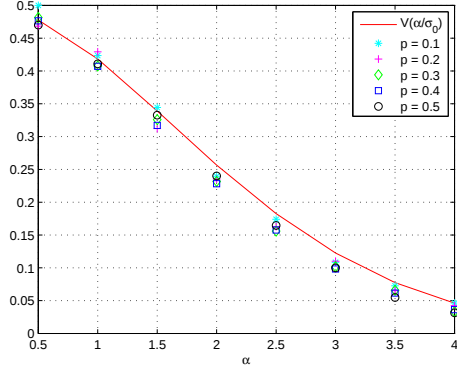


(c) $m = 400$

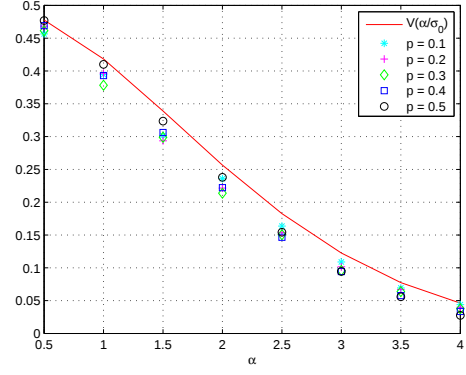


(d) $m = 800$

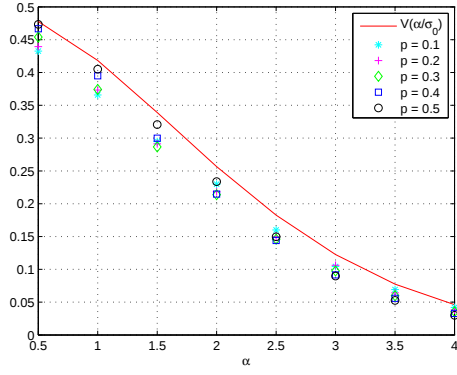
Figure 1: BER of the EST for different values of m and α in comparison to $V(\alpha/\sigma_0)$



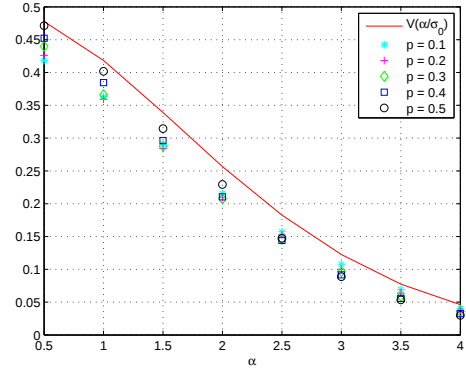
(a) $m = 100$



(b) $m = 200$



(c) $m = 400$



(d) $m = 800$

Figure 2: BER of the EST versus $V(\alpha/\sigma_0)$ for different values for m and α and when the amplitudes of the signals all equal $\kappa\alpha$ with $\kappa = 1.1220$, that is when the amplitudes of the signals are one dB larger than expected.

In fact, for signals less present than absent, the author's feeling is that theorem 2.1 and the ESE can be exploited to derive an estimate that requires no lower bound for the norms of the signals. Attempts of this type are presented in [2] and [9]. In these papers, methods for denoising speech signals corrupted by independent AWGN are adjusted by an estimate of the noise standard deviation. This estimate is performed on the basis of the ESE obtained by setting $\alpha = 0$, a trivial lower bound for the amplitudes of the signals. The performance of the denoising so adjusted is very close to that achieved by using the exact value of the noise standard deviation.

In order to state results that extend theorem 2.1 and give better insight into the behaviour of the ESE and the EST, we are strongly interested in establishing links between theorem 2.1, [8, Theorem VII.1] and Wald's fundamental paper [14].

Although theorem 2.1 is stated for signals whose probabilities of presence are upper-bounded by $p \in [0, 1)$, the EST and the experiments of section 4 concern the particular case $p = 1/2$ and the thresholding function ξ . Can the thresholding function ξ be used even when the upper bound for the probabilities of presence exceeds one half? From a more general point of view, we are looking for thresholding functions other than those proposed in the paper and, to deal with situations where signals may have probabilities of presence above one half, we plan to analyse the behaviour of the ESE adjusted with such thresholding functions so as to derive extensions of the EST.

As far as practical applications of the contents of this paper are concerned, speech denoising has already been mentioned as a promising field of application of the foregoing results (see [2] and [9]). A rather natural application of the approach presented in this paper is the design of Constant False Alarm Rate (CFAR) systems for the detection of radar targets. Our intention is to study to what extent theorem 2.1, the ESE and the EST are complementary to standard results and algorithms such as those described in [7]. ESM is also a possibly relevant field of application where the ESE and the EST could contribute to the interception of non-cooperant communications. Proximity sensing aimed, for instance, at informing a robot that it is approaching an object or that something is near it, can also be regarded as a potential application area where our results could apply. Finally, another possibly relevant and exciting field of theoretical and practical application of the material presented in this paper is the design and analysis of distributed detection systems (see [13]), where the absence of prior knowledge about the statistics of the observations is one of the main issues to overcome.

6 Acknowledgement

The author is very grateful to the reviewers for their insightful remarks that helped to significantly improve this paper. In particular, their suggestions strongly encourage the author to extend his approach in relation to Wald's theory and to address applications such as distributed detection and proximity sensing.

The author is also very thankful to Janet Ormrod for her help in improving the English and the organization of the paper.

Many thanks to Roger Gay and Albert 'Bert' Groenenboom for their keen suggestions and remarks.

Appendix

The notations introduced in theorem 2.1 are kept with the same meaning throughout this proof.

After stating some preliminary results in the next subsection, our first purpose is to prove that the noise standard deviation satisfies Eqs. (5) and (6). This will be achieved in three steps, namely subsections B, C and D.

More specifically, in subsection B, the behaviour of $\sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq T)$ is studied when m is large, T is a positive real value and q is any non-negative real number. A corollary of Kolmogorov's classical strong limit theorem shows the almost sure convergence of the ratio

$$\frac{\sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq T)}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq T) \right]}$$

to 1. From this almost sure convergence, we derive that of the ratio

$$\frac{\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq T) / \sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq T)}{\sum_{k=1}^m \mathbb{E} [\|U_k\|^r I(\|U_k\| \leq T)] / \sum_{k=1}^m \mathbb{E} [\|U_k\|^s I(\|U_k\| \leq T)]}$$

when $0 \leq r < s \leq \nu/2$ and $\nu \in (0, \infty]$. In subsection C, we then analyse the behaviour of the denominator of the ratio above when T is $\beta\theta(\alpha)$, $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$, $\mathbf{a}(\Lambda) \geq \alpha$ if $\nu \neq \infty$ and $\|\Lambda\|_\infty \leq \alpha$ otherwise. We will then see that this ratio can be approximated by $\Upsilon_r(\beta\theta(\alpha))/\Upsilon_s(\beta\theta(\alpha))$. In subsection D, the two types of convergence studied previously will be combined to conclude that σ_0 actually satisfies Eqs. (5) and (6).

In subsection E, we complete the proof of statement **(i)** by showing that the noise standard deviation is the sole solution in σ to Eq. (5).

From now on, we assume that $\sigma_0 = 1$ and carry out the proof in this case. The reader will easily verify that this simplification infers no loss of generality.

A Preliminary results

For every $k \in \mathbb{N}$, X_k and Λ_k are assumed to be independent and the distribution of X_k is absolutely continuous with respect to Lebesgue's measure in $\mathbb{R}^d$. Therefore, given $k \in \mathbb{N}$, the distribution of the random vector $\Lambda_k + X_k$ is absolutely continuous with respect to Lebesgue's measure in $\mathbb{R}^n$. Its density, denoted hereafter by $f_{\Lambda_k + X_k}$, is given by

$$f_{\Lambda_k + X_k}(u) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-\|u - \lambda\|^2/2} P_{\Lambda_k}(d\lambda), u \in \mathbb{R}^d \quad (13)$$

where the probability distribution P_{Λ_k} is the positive measure such that $P_{\Lambda_k}(E) = P(\Lambda_k^{-1}(E))$ for any Borel set E of $\mathbb{R}^d$.

For every $k \in \mathbb{N}$, the distribution of $\|\Lambda_k + X_k\|$ is absolutely continuous with respect to Lebesgue's measure. The density $f_{\|\Lambda_k + X_k\|}$ of this distribution is given, for any non-negative real number x , by

$$f_{\|\Lambda_k + X_k\|}(x) = \frac{2\pi^{d/2}}{\Gamma(d/2)} x^{d-1} \int_{S^{d-1}} f_{\Lambda_k + X_k}(xu) \Sigma(du), \quad (14)$$

where Σ stands for the rotation-invariant positive Borel measure on the sphere $S^{d-1} = \{u \in \mathbb{R}^d : \|u\| = 1\}$ for which $\Sigma(S^{d-1}) = 1$ and Γ is the usual complete Gamma function. It then follows from Eqs. (13) and (14), and Fubini's theorem, that, for any non-negative real number x ,

$$f_{\|\Lambda_k + X_k\|}(x) = \frac{e^{-x^2/2} x^{d-1}}{2^{d/2-1} \Gamma(d/2)} \int_{\mathbb{R}^d} e^{-\|\lambda\|^2/2} \left(\int_{S^{d-1}} e^{x(\lambda|u)} \Sigma(du) \right) P_{\Lambda_k}(d\lambda), \quad (15)$$

where $(u|v) = \sum_{i=1}^d u_i v_i$, $u, v \in \mathbb{R}^d$, is the usual scalar product in $\mathbb{R}^d$. It is then known ([8, Lemma A.1]) that

$$\int_{S^{d-1}} e^{x(\lambda|u)} \Sigma(du) = {}_0F_1(d/2; x^2 \|\lambda\|^2/4),$$

where ${}_0F_1$ is the generalized hypergeometric function ([6, p. 275]). We thus derive from Eq. (15) that

$$f_{\|\Lambda_k + X_k\|}(x) = \frac{e^{-x^2/2} x^{d-1}}{2^{d/2-1} \Gamma(n/2)} \int_{\mathbb{R}^d} e^{-\|\lambda\|^2/2} {}_0F_1(d/2; x^2 \|\lambda\|^2/4) P_{\Lambda_k}(d\lambda). \quad (16)$$

For any non-negative real numbers q and T , the following equalities are trivial but will prove very useful in the sequel. To begin with, we have

$$\mathbb{E} [\|\Lambda_k + X_k\|^q I(\|\Lambda_k + X_k\| \leq T)] = \int_0^T x^q f_{\|\Lambda_k + X_k\|}(x) dx \quad (17)$$

and

$$\mathbb{E} [\|X_k\|^q I(\|X_k\| \leq T)] = \int_0^T x^q f_d(x) dx, \quad (18)$$

where

$$f_d(x) = \frac{1}{2^{d/2-1} \Gamma(d/2)} e^{-x^2/2} x^{d-1}, x \geq 0. \quad (19)$$

is the probability density function of the square root of any centred chi-2 distribution with d degrees of freedom.

With p_k and q_k henceforth standing for $P(\{\varepsilon_k = 1\})$ and $1 - p_k$ respectively, we also can write that

$$\begin{aligned} \mathbb{E} [\|U_k\|^q I(\|U_k\| \leq T)] = \\ p_k \mathbb{E} [\|\Lambda_k + X_k\|^q I(\|\Lambda_k + X_k\| \leq T)] + q_k \int_0^T x^q f_d(x) dx. \end{aligned} \quad (20)$$

B The case of large sample sizes

The behaviour of $\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq T) / \sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq T)$ is given in lemma B.2 when m is large and T is any positive real value. The starting point is the following corollary ([12, chapter 5, p 108, corollary 3]) of Kolmogorov's classical strong limit theorem.

Lemma [Rao] B.1 *If $Y_1, Y_2, \dots$ is a sequence of independent random variables with non-negative means $\alpha_1, \alpha_2, \dots$, and variances $\sigma_1^2, \sigma_2^2, \dots$, such that $\sum_{m=1}^{+\infty} \alpha_m = +\infty$ (i) and $\sum_{m=1}^{+\infty} \sigma_m^2 / a_m^2 < +\infty$ (ii), where $a_m = \sum_{k=1}^m \alpha_k$, then $\lim_m (1/a_m) \sum_{k=1}^m Y_m = 1$ (a-s).*

Proposition B.2 *Under assumptions (Ass. 2), (Ass. 3) and (Ass. 4), for every non-negative real number q less than or equal to $\nu/2$ and every positive real number T ,*

$$\lim_{m \rightarrow +\infty} \sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq T) / \mathbb{E} \left[\sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq T) \right] = 1 \quad (\text{a-s}).$$

PROOF: The proof involves establishing that

(i') the sequence $(\text{Var}(\|U_k\|^q I(\|U_k\| \leq T)))_{k \in \mathbb{N}}$ is upper-bounded;

(ii') the sequence $(\mathbb{E}[\|U_k\|^q I(\|U_k\| \leq T)])_{k \in \mathbb{N}}$ is bounded away from 0.

Actually, the two conditions (i') and (ii') are sufficient to guarantee that the sequence $(\|U_k\|^q I(\|U_k\| \leq T))_{k \in \mathbb{N}}$ of independent variables fulfils conditions (i) and (ii) of lemma B.1 and, thereby, to conclude.

Clearly, $X \in \ell^\infty(\mathbb{N}, L^{2q}(\Omega, \mathbb{R}^d))$; on the other hand, it follows from Eq. (2) that $\ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ is a subset of $\ell^\infty(\mathbb{N}, L^{2q}(\Omega, \mathbb{R}^d))$. Therefore, $\Lambda + X$ is an element of $\ell^\infty(\mathbb{N}, L^{2q}(\Omega, \mathbb{R}^d))$. Condition (i') then holds true as a consequence of the inequality $\text{Var}(\|U_k\|^q I(\|U_k\| \leq T)) \leq \mathbb{E}[\|\Lambda_k + X_k\|^{2q}] + \mathbb{E}[\|X_k\|^{2q}]$, which straightforwardly follows from Eq. (20).

Condition (ii') derives from the non-vanishing behaviour of the probabilities of absence q_k , $k \in \mathbb{N}$, the fact that T is positive and the easy inequality

$$\mathbb{E}[\|U_k\|^q I(\|U_k\| \leq T)] \geq (1 - p) \int_0^T x^q f_d(x) dx,$$

which also follows from Eq. (20). ■

The subsequent result straightforwardly derives from the foregoing.

Corollary B.3 *Under assumptions (Ass. 2), (Ass. 3) and (Ass. 4), for every pair (r, s) of real numbers such that $0 \leq s < r \leq \nu/2$ and every positive real number T ,*

$$\lim_{m \rightarrow \infty} \left(\frac{\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq T)}{\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq T)} \times \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq T) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq T) \right]} \right) = 1 \text{ (a-s).}$$

C The case of large and small amplitudes

The thresholding function θ is of most importance in the present section where, given $\beta \in [0, 1]$, we derive the behaviour of the ratio

$$\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta \theta(\mathbf{a}(\Lambda))) \right] / \mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta \theta(\mathbf{a}(\Lambda))) \right]$$

when $\mathbf{a}(\Lambda)$ tends to ∞ and that of

$$\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta \theta(\|\Lambda\|_\infty)) \right] / \mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta \theta(\|\Lambda\|_\infty)) \right]$$

when $\|\Lambda\|_\infty$ tends to 0. The starting point of the analysis is the following result.

Lemma C.1 Under assumption **(Ass. 1)**, for every given $q \in [0, \infty)$,

(i) $\lim_{\mathbf{a}(\Lambda) \rightarrow \infty} \mathbb{E} [\|\Lambda_k + X_k\|^q I(\|\Lambda_k + X_k\| \leq \beta \theta(\mathbf{a}(\Lambda)))] = 0$, uniformly in k and $\beta \in [0, 1]$.

(ii) If $\Lambda \in \ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$, then, in addition to statement (i),

$$\lim_{\|\Lambda\|_\infty \rightarrow 0} \left(\mathbb{E} [\|\Lambda_k + X_k\|^q I(\|\Lambda_k + X_k\| \leq \beta \theta(\|\Lambda\|_\infty))] - \mathbb{E} [\|X_k\|^q I(\|X_k\| \leq \beta \theta(\|\Lambda\|_\infty))] \right) = 0,$$

uniformly in k and $\beta \in [0, 1]$.

PROOF: We start by proving statement (i). For the sake of simplifying notations, we put $\rho = \mathbf{a}(\Lambda)$. Given any $\beta \in [0, 1]$, we easily derive from Eq. (17) that $\mathbb{E} [\|\Lambda_k + X_k\|^q I(\|\Lambda_k + X_k\| \leq \beta \theta(\rho))]$ is less than or equal to $\theta(\rho)^q P(\{\|\Lambda_k + X_k\| \leq \theta(\rho)\})$. Our proof involves showing that the latter quantity tends to 0 uniformly with k when ρ tends to ∞ .

We derive from the properties of θ that, for ρ large enough, $C\rho < \theta(\rho) \leq \rho$. According to Eq. (13) and since, for any natural number k , $\|\Lambda_k\| \geq \rho$ (a-s), the probability that $\Lambda_k + X_k$ lies inside the ball of radius $\theta(\rho)$ can be written as follows:

$$\begin{aligned} & P(\{\|\Lambda_k + X_k\| \leq \theta(\rho)\}) \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\|x\| \leq C\rho} \int_{\|\lambda\| \geq \rho} e^{-\|x-\lambda\|^2/2} P_{\Lambda_k}(d\lambda) dx, \\ & \quad + \frac{1}{(2\pi)^{d/2}} \int_{C\rho < \|x\| \leq \theta(\rho)} \int_{\|\lambda\| \geq \rho} e^{-\|x-\lambda\|^2/2} P_{\Lambda_k}(d\lambda) dx. \end{aligned} \quad (21)$$

Because $\|x - \lambda\| \geq \rho(1 - C)$ for $\|x\| \leq C\rho$ and $\|\lambda\| \geq \rho$, the first term in the right-hand side of Eq. (21) is bounded by a constant multiplied with $\rho^d e^{-\frac{(1-C)^2}{2}\rho^2}$.

For $C\rho < \|x\| \leq \theta(\rho)$ and $\|\lambda\| \geq \rho$, $\|x - \lambda\|$ is larger than or equal to $\rho(1 - C) - \gamma(\rho)$. Hence, by taking into account the volume of the shell $\{x \in \mathbb{R}^d : C\rho < \|x\| \leq \theta(\rho)\}$, the second term in the right-hand side of Eq. (21) is less than a constant times $\left(1 + \frac{\gamma(\rho)}{C\rho}\right)^d - 1 \rho^n e^{-((1-C)-\gamma(\rho)/\rho)^2(\rho^2/2)}$.

In addition, $\theta(\rho) \sim C\rho$ when ρ tends to $+\infty$. It thus follows from the foregoing that, for every non-negative real number q , $\theta(\rho)^q P(\{\|\Lambda_k + X_k\| < \theta(\rho)\})$ tends to 0 uniformly with k when $\rho = \mathbf{a}(\Lambda)$ tends to ∞ , which completes the proof of statement (i).

We now turn our attention to the proof of statement **(ii)**. Still for notational sake, we set $\rho = \|\Lambda\|_\infty$. An easy consequence of Eq. (17) is that, for any $\beta \in [0, 1]$, the absolute value of the difference

$$\mathbb{E} \left[\|\Lambda_k + X_k\|^q I(\|\Lambda_k + X_k\| \leq \beta\theta(\rho)) \right] - \mathbb{E} \left[\|X_k\|^q I(\|X_k\| \leq \beta\theta(\rho)) \right]$$

is less than or equal to $\theta(\rho)^q \int_0^{\theta(\rho)} |f_{\|\Lambda_k + X_k\|}(x) - f_d(x)| dx$. Our intention is then to prove that the latter quantity tends to 0 uniformly in k when ρ tends to 0. The hypergeometric function ${}_0F_1(d/2; \cdot)$ is continuous, increasing and ${}_0F_1(d/2; 0) = 1$. Hence, for $\|\lambda\| \leq \rho$ and $x \in [0, \theta(\rho)]$,

$$1 \leq {}_0F_1(d/2; x^2\|\lambda\|^2/4) \leq {}_0F_1(d/2; \rho^2\theta(\rho)^2/4).$$

From the inequality above, Eqs. (16) and (19), and the fact that for every $k \in \mathbb{N}$, $\|\Lambda_k\|$ is less than or equal to ρ (a-s), it follows that, for all $x \in [0, \theta(\rho)]$,

$$\frac{e^{-x^2/2} x^{d-1}}{2^{d/2-1} \Gamma(d/2)} \left(e^{-\rho^2/2} - 1 \right) \leq f_{\|\Lambda_k + X_k\|}(x) - f_d(x)$$

and

$$f_{\|\Lambda_k + X_k\|}(x) - f_d(x) \leq \frac{e^{-x^2/2} x^{d-1}}{2^{d/2-1} \Gamma(d/2)} \left({}_0F_1(d/2; \rho^2\theta(\rho)^2/4) - 1 \right).$$

The two functions of ρ that bracket $f_{\|\Lambda_k + X_k\|}(x) - f_d(x)$ in the inequalities above both trivially tend to 0 independently of k when ρ tends to 0. Since θ is continuous at the origin, we have

$$\lim_{\rho \rightarrow 0} I_{[0, \theta(\rho)]}(x) \theta(\rho)^q |f_{\|\Lambda_k + X_k\|}(x) - f_d(x)| = 0,$$

uniformly in k . It follows from Lebesgue's dominated convergence theorem that $\theta(\rho)^q \int_0^{\theta(\rho)} |f_{\|\Lambda_k + X_k\|}(x) - f_d(x)| dx$ tends to 0 uniformly in k when ρ tends to 0, which completes the proof. ■

Proposition C.2 *Under assumptions **(Ass. 1)**, **(Ass. 2)** and **(Ass. 3)**, for every given $\beta_0 \in (0, 1]$ and every given pair (r, s) of non-negative real numbers such that $0 \leq s < r$,*

$$\lim_{a(\Lambda) \rightarrow \infty} \left(\frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(a(\Lambda))) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(a(\Lambda))) \right]} - \frac{\Upsilon_r(\beta\theta(a(\Lambda)))}{\Upsilon_s(\beta\theta(a(\Lambda)))} \right) = 0 \quad (22)$$

uniformly in m and $\beta \in [\beta_0, 1]$.

Furthermore, if Λ belongs to $\ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$,

$$\lim_{\|\Lambda\|_\infty \rightarrow 0} \left(\frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta \theta(\|\Lambda\|_\infty)) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta \theta(\|\Lambda\|_\infty)) \right]} - \frac{\Upsilon_r(\beta \theta(\|\Lambda\|_\infty))}{\Upsilon_s(\beta \theta(\|\Lambda\|_\infty))} \right) = 0 \quad (23)$$

uniformly in m and $\beta \in [\beta_0, 1]$.

PROOF: We prove Eqs. (22) and (23) together. As above, let ρ stand for either $\mathbf{a}(\Lambda)$ or $\|\Lambda\|_\infty$. In the latter case, it is implicitly assumed that $\Lambda \in \ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$.

Let q be any non-negative real value, β_0 any element of $(0, 1]$ and β any element of $[\beta_0, 1]$. It follows from Eq. (20) that

$$\begin{aligned} & \mathbb{E} \left[\sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq \beta \theta(\rho)) \right] \\ &= \sum_{k=1}^m p_k \left(\mathbb{E} [\|\Lambda_k + X_k\|^q I(\|\Lambda_k + X_k\| \leq \beta \theta(\rho))] \right. \\ & \quad \left. - \delta_{\ell,0} \mathbb{E} [\|X_k\|^q I(\|X_k\| \leq \beta \theta(\rho))] \right) \\ & \quad + \left(\sum_{k=1}^m q_k + \delta_{\ell,0} \sum_{k=1}^m p_k \right) \int_0^{\beta \theta(\rho)} x^q f_d(x) dx, \end{aligned} \quad (24)$$

where $\ell = \infty$ if $\rho = \mathbf{a}(\Lambda)$ and $\ell = 0$ otherwise and $\delta_{a,b}$, $a, b \in \mathbb{R}$, stands for the usual Kronecker symbol: $\delta_{a,b} = 1$ if $a = b$ and $\delta_{a,b} = 0$ otherwise.

According to lemma C.1, if $\rho = \mathbf{a}(\Lambda)$ (resp. $\rho = \|\Lambda\|_\infty$), there exists some positive real number α_0 such that, for every $\alpha \geq \alpha_0$ (resp. $\alpha \leq \alpha_0$), every Λ such that $\mathbf{a}(\Lambda) \geq \alpha$ (resp. $\|\Lambda\|_\infty \leq \alpha$) and every $\beta \in [\beta_0, 1]$, each factor of p_k in the right-hand side of Eq. (24) is less than or equal to η . By taking into account that each p_k is smaller than or equal to $p < 1$ and that $\sum_{k=1}^m q_k + \delta_{\ell,0} \sum_{k=1}^m p_k$ remains positive, it follows that

$$\left| \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq \beta \theta(\rho)) \right]}{\sum_{k=1}^m q_k + \delta_{\ell,0} \sum_{k=1}^m p_k} - \int_0^{\beta \theta(\rho)} x^q f_d(x) dx \right| \leq \frac{p}{1-p} \eta$$

for ρ large enough (resp. small enough). The right-hand side of the latter inequality does not depend on m . Moreover,

$$\lim_{\rho \rightarrow \ell} \int_0^{\beta\theta(\rho)} x^q f_d(x) dx = \int_0^{\beta\theta(\ell)} x^q f_d(x) dx,$$

whether $\ell = 0$ or $\ell = \infty$ since $\theta(0) > 0$, $\theta(\infty) = \infty$ and $\beta \geq \beta_0 > 0$. Thereby, we obtain

$$\lim_{\rho \rightarrow \ell} \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^q I(\|U_k\| \leq \beta\theta(\rho)) \right]}{\sum_{k=1}^m q_k + \delta_{\ell,0} \sum_{k=1}^m p_k} = \int_0^{\beta\theta(\ell)} x^q f_d(x) dx \in (0, \infty),$$

uniformly in m and $\beta \in [\beta_0, 1]$. Given two real numbers $r > s \geq 0$, it straightforwardly follows that

$$\lim_{\rho \rightarrow \ell} \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\rho)) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\rho)) \right]} = \frac{\int_0^{\beta\theta(\ell)} x^r f_d(x) dx}{\int_0^{\beta\theta(\ell)} x^s f_d(x) dx},$$

which concludes the proof. ■

D Combining large sample sizes and large or small amplitudes

Corollary B.3 holds true for any threshold T and real values r and s such that $0 \leq s < r \leq \nu/2$. Proposition C.2 is valid for all non-negative real values r and s such that $s < r$ and thresholds deriving from the thresholding function θ . Hence, the following lemma D.1 particularizes corollary B.3 to such thresholds. The rest of the section then combines proposition C.2 and lemma D.1.

Lemma D.1 *Under assumptions (Ass. 1), (Ass. 2), (Ass. 3), (Ass. 4), let r and s be two real numbers such that $0 \leq s < r \leq \nu/2$. Then, for any $\beta_0 \in (0, 1]$, there exists some positive real number α_0 (resp. α_1) such that, for any α in $[\alpha_0, \infty)$ (resp. $(0, \alpha_1]$), any $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ (resp.*

$\Lambda \in \ell^\infty(\mathbb{N}, L^\infty(\Omega, \mathbb{R}^d))$) such that $\mathbf{a}(\Lambda) \geq 1$ (resp. $\|\Lambda\|_\infty \leq \alpha$) and any $\beta \in [\beta_0, 1]$,

$$\lim_{m \rightarrow +\infty} \left(\frac{\sum_{k=1}^m \|U_k\|^r I(\|U\| \leq \beta\theta(\rho))}{\sum_{k=1}^m \|U_k\|^s I(\|U\| \leq \beta\theta(\rho))} - \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\rho)) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\rho)) \right]} \right) = 0 \text{ (a-s)} \quad (25)$$

with $\rho = \mathbf{a}(\Lambda)$ (resp. $\rho = \|\Lambda\|_\infty$).

PROOF: By differentiating $\Upsilon_r(x)/\Upsilon_s(x)$ with respect to $x \geq 0$, it follows that Υ_r/Υ_s is a non-decreasing map of $[0, \infty]$ into $[0, \Upsilon_r(\infty)/\Upsilon_s(\infty)]$ with $\Upsilon_r(\infty)/\Upsilon_s(\infty) < \infty$.

On the other hand, since $\theta(x) \geq \theta(0) > 0$ for every non-negative real number x , a consequence of Eqs. (17) and (20) is that, for all $\beta > 0$,

$$\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right] / \mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right] > 0$$

Let β_0 be an element of $(0, 1]$ and η be some positive real number. It follows from proposition C.2 and the remarks above that there exists some positive real number α_0 such that, for any real number α larger than or equal to α_0 , any $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ such that $\mathbf{a}(\Lambda) \geq \alpha$, any $\beta \in [\beta_0, 1]$ and any $m \in \mathbb{N}$,

$$0 < \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]} \leq \frac{\Upsilon_r(\infty)}{\Upsilon_s(\infty)} + \eta.$$

If $\alpha \in [\alpha_0, \infty)$ and $\beta \in [\beta_0, 1]$, the latter inequality implies that, for every $\omega \in \Omega$ and every $m \in \mathbb{N}$,

$$\begin{aligned} & \left| \frac{\sum_{k=1}^m \|U_k(\omega)\|^r I_{[0, \beta\theta(\mathbf{a}(\Lambda))]}(\|U_k(\omega)\|)}{\sum_{k=1}^m \|U_k(\omega)\|^s I_{[0, \beta\theta(\mathbf{a}(\Lambda))]}(\|U_k(\omega)\|)} - \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]} \right| \\ & \leq \left(\frac{\Upsilon_r(\infty)}{\Upsilon_s(\infty)} + \eta \right) \\ & \quad \times \left| \frac{\sum_{k=1}^m \|U_k(\omega)\|^r I_{[0, \beta\theta(\mathbf{a}(\Lambda))]}(\|U_k(\omega)\|) \mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]}{\sum_{k=1}^m \|U_k(\omega)\|^s I_{[0, \beta\theta(\mathbf{a}(\Lambda))]}(\|U_k(\omega)\|) \mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]} - 1 \right|, \end{aligned}$$

where I_K stands for the indicator function of a given set K . According to corollary B.3, for every ω in some measurable subset with measure 1 of Ω , the second factor in the right-hand side of the inequality above can be made arbitrarily small for m large enough, which completes the proof for $\mathbf{a}(\cdot)$. The reader will then easily obtain the same result for $\|\cdot\|_\infty$ by mimicking the foregoing. ■

We now combine the almost everywhere convergence established by the latter lemma with the uniform convergence stated in proposition C.2 so as to complete the proof that $\sigma = 1$ satisfies Eqs. (5) and (6). In fact, we prove Eq. (5) only, because the proof of Eq. (6) can be achieved in exactly the same way as below by replacing $\mathbf{a}(\cdot)$ with $\|\cdot\|_\infty$.

Let $\beta_0 \in (0, 1]$ and η be any positive real number. Let α_1 be a positive real value such that, for all $\alpha \in [\alpha_1, \infty)$, all $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ such that $\mathbf{a}(\Lambda) \geq \alpha$ and every $\beta \in [\beta_0, 1]$, Eq. (25) holds true with $\rho = \mathbf{a}(\Lambda)$.

We derive from proposition C.2 the existence of another positive real number α_2 such that, for any $\alpha \in [\alpha_2, \infty)$, any $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ with $\mathbf{a}(\Lambda) \geq \alpha$, any $\beta \in [\beta_0, 1]$ and any natural number m ,

$$\left| \frac{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]}{\mathbb{E} \left[\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda))) \right]} - \frac{\Upsilon_r(\beta\theta(\mathbf{a}(\Lambda)))}{\Upsilon_s(\beta\theta(\mathbf{a}(\Lambda)))} \right| \leq \eta. \quad (26)$$

For any α larger than or equal to $\alpha_0 = \max(\alpha_1, \alpha_2)$, any Λ in $\ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ such that $\mathbf{a}(\Lambda) \geq \alpha$ and any $\beta \in [\beta_0, 1]$, Eqs. (25) and (26) imply that

$$\limsup_m \left| \frac{\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda)))}{\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda)))} - \frac{\Upsilon_r(\beta\theta(\mathbf{a}(\Lambda)))}{\Upsilon_s(\beta\theta(\mathbf{a}(\Lambda)))} \right| \leq \eta \text{ (a-s),}$$

or equivalently, that

$$\left\| \limsup_m \left| \frac{\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda)))}{\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta\theta(\mathbf{a}(\Lambda)))} - \frac{\Upsilon_r(\beta\theta(\mathbf{a}(\Lambda)))}{\Upsilon_s(\beta\theta(\mathbf{a}(\Lambda)))} \right| \right\|_\infty \leq \eta. \quad (27)$$

Summarizing, for any given real number $\beta_0 \in (0, 1]$ and any positive real number η , we have found a positive real number α_0 such that, for any $\alpha \in [\alpha_0, \infty)$, any $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ that satisfies $\mathbf{a}(\Lambda) \geq \alpha$ and any $\beta \in [\beta_0, 1]$, Eq. (27) holds true. This means that $\sigma = 1$ satisfies Eq. (5).

E Uniqueness of the solution to equation

Assume the existence of two positive real numbers $\sigma_1 \geq \sigma_2 > 0$ that both satisfy Eq. (5). For notational sake, put $\rho_i = \mathbf{a}(\Lambda)/\sigma_i$, $i \in \{1, 2\}$ and set $\beta_i = \theta(\rho_i)/\rho_i$, $i = 1, 2$. For every positive real number x , $0 < \Upsilon_r(x)/\Upsilon_s(x) < \infty$. Therefore, for any given pair $(m, \omega) \in \mathbb{N} \times \Omega$, we can write that

$$\begin{aligned} & \left| \sigma_1^{r-s} \frac{\Upsilon_r(\beta_2\theta(\rho_1))}{\Upsilon_s(\beta_2\theta(\rho_1))} - \sigma_2^{r-s} \frac{\Upsilon_r(\beta_1\theta(\rho_2))}{\Upsilon_s(\beta_1\theta(\rho_2))} \right| \leq \\ & \left| \frac{\sum_{k=1}^m \|U_k\|^r I_{[0, \beta_2\sigma_1\theta(\rho_1)]}(\|U_k(\omega)\|)}{\sum_{k=1}^m \|U_k\|^s I_{[0, \beta_2\sigma_1\theta(\rho_1)]}(\|U_k(\omega)\|)} - \sigma_1^{r-s} \frac{\Upsilon_r(\beta_2\theta(\rho_1))}{\Upsilon_s(\beta_2\theta(\rho_1))} \right| \\ & + \left| \frac{\sum_{k=1}^m \|U_k\|^r I_{[0, \beta_2\sigma_1\theta(\rho_1)]}(\|U_k(\omega)\|)}{\sum_{k=1}^m \|U_k\|^s I_{[0, \beta_2\sigma_1\theta(\rho_1)]}(\|U_k(\omega)\|)} - \sigma_2^{r-s} \frac{\Upsilon_r(\beta_1\theta(\rho_2))}{\Upsilon_s(\beta_1\theta(\rho_2))} \right|. \end{aligned}$$

We derive from this inequality that

$$\begin{aligned} & \left| \sigma_1^{r-s} \frac{\Upsilon_r(\beta_2\theta(\rho_1))}{\Upsilon_s(\beta_2\theta(\rho_1))} - \sigma_2^{r-s} \frac{\Upsilon_r(\beta_1\theta(\rho_2))}{\Upsilon_s(\beta_1\theta(\rho_2))} \right| \leq \\ & \left\| \limsup_m \left| \frac{\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta_2\sigma_1\theta(\rho_1))}{\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta_2\sigma_1\theta(\rho_1))} - \sigma_1^{r-s} \frac{\Upsilon_r(\beta_2\theta(\rho_1))}{\Upsilon_s(\beta_2\theta(\rho_1))} \right| \right\|_\infty \\ & + \left\| \limsup_m \left| \frac{\sum_{k=1}^m \|U_k\|^r I(\|U_k\| \leq \beta_2\sigma_1\theta(\rho_1))}{\sum_{k=1}^m \|U_k\|^s I(\|U_k\| \leq \beta_2\sigma_1\theta(\rho_1))} - \sigma_2^{r-s} \frac{\Upsilon_r(\beta_1\theta(\rho_2))}{\Upsilon_s(\beta_1\theta(\rho_2))} \right| \right\|_\infty. \end{aligned}$$

Since $\beta_1\sigma_2\theta(\rho_2) = \beta_2\sigma_1\theta(\rho_1)$, it follows from the inequality above and the definition of Δ_m given by Eq. (4) that

$$\begin{aligned} & \left| \sigma_1^{r-s} \frac{\Upsilon_r(\beta_2\theta(\rho_1))}{\Upsilon_s(\beta_2\theta(\rho_1))} - \sigma_2^{r-s} \frac{\Upsilon_r(\beta_1\theta(\rho_2))}{\Upsilon_s(\beta_1\theta(\rho_2))} \right| \\ & \leq \left\| \limsup_m \Delta_m(\sigma_1, \beta_2\theta(\rho_1)) \right\|_\infty + \left\| \limsup_m \Delta_m(\sigma_2, \beta_1\theta(\rho_2)) \right\|_\infty. \quad (28) \end{aligned}$$

According to the asymptotic behaviour of θ , if $\mathbf{a}(\Lambda)$ is large enough, β_1 and β_2 are both elements of $(0, 1]$; moreover, σ_1 and σ_2 are both assumed to satisfy statement **(i)** of theorem 2.1. Therefore, for every given positive real number η , there exists some $\alpha_0 \in (0, \infty)$ such that, for every real number $\alpha \geq \alpha_0$ and every $\Lambda \in \ell^\infty(\mathbb{N}, L^\nu(\Omega, \mathbb{R}^d))$ that satisfies $\mathbf{a}(\Lambda) \geq \alpha$,

$$\left\| \limsup_m \Delta_m(\sigma_1, \beta_2\theta(\rho_1)) \right\|_\infty \leq \eta \text{ and } \left\| \limsup_m \Delta_m(\sigma_2, \beta_1\theta(\rho_2)) \right\|_\infty \leq \eta. \quad (29)$$

Writing ρ_i and β_i in function of α and σ_i , $i = 1, 2$, it then follows from Eqs. (28) and (29) that

$$\lim_{\rho \rightarrow \infty} \left(\sigma_1^{r-s} \frac{\Upsilon_r(\theta(\rho/\sigma_1)\theta(\rho/\sigma_2)/(\rho/\sigma_2))}{\Upsilon_s(\theta(\rho/\sigma_1)\theta(\rho/\sigma_2)/(\rho/\sigma_2))} - \sigma_2^{r-s} \frac{\Upsilon_r(\theta(\rho/\sigma_1)\theta(\rho/\sigma_2)/(\rho/\sigma_1))}{\Upsilon_s(\theta(\rho/\sigma_1)\theta(\rho/\sigma_2)/(\rho/\sigma_1))} \right) = 0.$$

According to the asymptotic behaviour of θ , the limits of the respective coefficients of σ_1^{r-s} and σ_2^{r-s} in the equality above are both equal to $\Upsilon_r(\infty)/\Upsilon_s(\infty)$ when ρ tends to ∞ . Since this common limit is positive and finite, it follows that $\sigma_1 = \sigma_2$ since $r > s$.

References

- [1] M. Abramowitz, I. Stegun, 1972. Handbook of Mathematical Functions. Ninth printing. Dover Publications Inc., New York.
- [2] A. Amehraye, D. Pastor, S. Ben Jebara, 2006. On the Application of Recent Results in Statistical Decision and Estimation Theory to Perceptual Filtering of Noisy Speech Signals, Proceedings of the second IEEE-EURASIP International Symposium on Control, Communications, and Signal Processing, ISCCSP'06, Marrakech, Morocco, 2006.
- [3] T. Kailath, H. V. Poor, 1998. Detection of Stochastic Processes. IEEE Transactions on Information Theory. 44, 2230-2259.
- [4] S. M. Kay, 1998. Fundamentals of Statistical Signal Processing. Volume II: Detection Theory. Prentice Hall, Upper Saddle River.
- [5] G. Köthe, 1969. Topological Vector Spaces-I. Springer-Verlag (Grundlehren der Mathematischen Wissenschaften, 159), Berlin .
- [6] N. N. Lebedev, 1965. Special Functions and their Applications. Prentice-Hall, Englewood Cliffs.
- [7] G. Minkler, J. Minkler, 1990. The Principles of Automatic Radar Detection in Clutter, CFAR. Magellan Book Company, Baltimore.
- [8] D. Pastor, R. Gay, B. Groenenboom, 2002. A Sharp Upper-Bound for the Probability of Error of the Likelihood Ratio Test for Detecting Signals in White Gaussian Noise. IEEE Transactions on Information Theory, 48, No. 1, pp 228-238.
- [9] D. Pastor, A. Amehraye, 2006. From non parametric statistics to speech denoising. Proceedings of the 3rd international symposium on Image/Video Communications over fixed and mobile networks, ISIVC'06, Hammamet, Tunisia.
- [10] H. V. Poor, 1994. An Introduction to Signal Detection and Estimation. 2nd Edition. Springer-Verlag, New York.
- [11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, 1992. Numerical recipes in C, The Art of Scientific Computing, second edition. University Press, Cambridge.
- [12] M. M. Rao, 1993. Conditional Measures and Applications. Marcel Dekker, New York.

- [13] P. K. Varshney, 1997. Distributed Detection and Data Fusion. Springer-Verlag New-York Inc., New-York.
- [14] A. Wald, 1943. Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. Transactions of the American Mathematical Society, Vol. 49, No. 3, pp. 426-482.

© ENST Bretagne. 2006

Ecole nationale supérieure des Télécommunications de Bretagne – Brest

Dépôt légal : 4^{ème} trimestre 2006

ISSN : 1255-2275