

A unified treatment for non-asymptotic and asymptotic approaches to minimax signal detection*

Clément Marteau,
Institut Mathématiques de Toulouse, INSA de Toulouse,
Université de Toulouse,
135, avenue de Rangueil, 31 077 Toulouse Cedex 4, France.
Email: clement.marteau@math.univ-toulouse.fr

and

Theofanis Sapatinas,
Department of Mathematics and Statistics,
University of Cyprus,
P.O. Box 20537, CY 1678 Nicosia, Cyprus.
Email: fanis@ucy.ac.cy

May 8, 2021

Abstract

We are concerned with minimax signal detection. In this setting, we discuss non-asymptotic and asymptotic approaches through a unified treatment. In particular, we consider a Gaussian sequence model that contains classical models as special cases, such as, direct, well-posed inverse and ill-posed inverse problems. Working with certain ellipsoids in the space of squared-summable sequences of real numbers, with a ball of positive radius removed, we compare the construction of lower and upper bounds for the minimax separation radius (non-asymptotic approach) and the minimax separation rate (asymptotic approach) that have been proposed in the literature. Some additional contributions, bringing into light links between non-asymptotic and asymptotic approaches to minimax signal, are also presented. An example of a mildly ill-posed inverse problem is used for illustrative purposes. In particular, it is shown that tools used to derive ‘asymptotic’ results can be exploited to draw ‘non-asymptotic’ conclusions, and vice-versa.

AMS 2000 subject classifications: 62G05, 62K20

Keywords and phrases: Gaussian sequence models, ill-posed and well-posed inverse problems, minimax signal detection.

*This work is dedicated to the memory of our collaborator and friend Yuri I. Ingster who passed away on August 8th, 2012.

1 Introduction

We consider the following Gaussian sequence model (GSM),

$$Y_j = b_j \theta_j + \varepsilon \xi_j, \quad j \in \mathbb{N}, \quad (1.1)$$

where $\mathbb{N} = \{1, 2, \dots\}$ is the set of natural numbers, $b = \{b_j\}_{j \in \mathbb{N}} > 0$ is a known sequence, $\theta = \{\theta_j\}_{j \in \mathbb{N}} \in l^2(\mathbb{N})$ is the unknown signal of interest, $\xi = \{\xi_j\}_{j \in \mathbb{N}}$ is a sequence of independent standard Gaussian random variables, and $\varepsilon > 0$ is a known parameter (the noise level). The observations are given by the sequence $Y = \{Y_j\}_{j \in \mathbb{N}}$ from the GSM (1.1) and their joint law is denoted by \mathbb{P}_θ . Here, $l^2(\mathbb{N})$ denotes the space of squared-summable sequence of real numbers, i.e.,

$$l^2(\mathbb{N}) = \left\{ \theta \in \mathbb{R}^{\mathbb{N}} : \|\theta\|^2 := \sum_{j \in \mathbb{N}} \theta_j^2 < +\infty \right\}.$$

The GSM (1.1) arises in many well-known situations. For instance, consider the Gaussian white noise model (GWNM)

$$dX_\varepsilon(t) = Af(t)dt + \varepsilon dW(t), \quad t \in V, \quad (1.2)$$

where A is a known linear operator acting on a Hilbert space \mathcal{H}_1 with values on another Hilbert space \mathcal{H}_2 , $f(\cdot) \in \mathcal{H}_1$ is the unknown response function that one wants to detect or estimate, $W(\cdot)$ is a standard Wiener process on $V \subseteq \mathbb{R}$ and $\varepsilon > 0$ is a known parameter (the noise level). For the sake of simplicity, we only consider the case when A is injective (meaning that A has a trivial nullspace) and assume that $V = [0, 1]$, $\mathcal{H}_1 = L^2(V)$, $U \subseteq \mathbb{R}$ and $\mathcal{H}_2 = L^2(U)$.

- (direct problem) Let $A = I$ (the identity operator). Let $\{\phi_j\}_{j \in \mathbb{N}}$ be an orthonormal basis on $L^2(V)$. Transforming the GWNM (1.2) with $A = I$ into the Fourier domain, the GSM (1.1) arises with $Y_j = \int_0^1 \phi_j(t) dX_\varepsilon(t)$, $\theta_j = \int_0^1 \phi_j(t) f(t) dt$, $\xi_j = \int_0^1 \phi_j(t) dW(t)$ and $b_j = 1$, for all $j \in \mathbb{N}$.
- (well-posed inverse problem) Let A be a self-adjoint operator that admits an eigenvalue-eigenfunction decomposition $(b_j, \varphi_j)_{j \in \mathbb{N}}$, in the sense that

$$A\varphi_j = b_j \varphi_j, \quad j \in \mathbb{N},$$

where $b_j > b_0$, for some $b_0 > 0$, for all $j \in \mathbb{N}$. Thus, the GSM (1.1) arises with $Y_j = \int_0^1 \varphi_j(t) dX_\varepsilon(t)$, $\theta_j = \int_0^1 \varphi_j(t) f(t) dt$, $\xi_j = \int_0^1 \varphi_j(t) dW(t)$ and $b_j > b_0 > 0$, for all $j \in \mathbb{N}$. In this case, the GWNM (1.2) corresponds to a so-called well-posed inverse problem. Possible examples of such decompositions arise with, e.g., differential or Sturm-Liouville operators.

- (ill-posed inverse problem) In most cases of interest, however, A is a compact operator (see, e.g., Chapter 2 of [7]). In particular, it admits a singular value decomposition (SVD) $(b_j, \psi_j, \varphi_j)_{j \in \mathbb{N}}$, in the sense that

$$A\varphi_j = b_j \psi_j, \quad A^* \psi_j = b_j \varphi_j, \quad j \in \mathbb{N},$$

where A^* denotes the adjoint operator of A – note that $(b_j^2)_{j \in \mathbb{N}}$ and $(\varphi_j)_{j \in \mathbb{N}}$ are, respectively, the eigenvalues and the eigenfunctions of A^*A . Thus, the GSM (1.1) arises with $Y_j = \int_0^1 \psi_j(t) dX_\varepsilon(t)$, $\theta_j = \int_0^1 \varphi_j(t) f(t) dt$, $\xi_j = \int_0^1 \psi_j(t) dW(t)$ and $b_j > 0$ (since A is injective), for all $j \in \mathbb{N}$. In this case, the GWNM (1.2) corresponds to a so-called ill-posed inverse problem since the inversion of A^*A is not bounded. Possible examples of such decompositions arise with, e.g., convolution or Radon-transform operators. The effect of the ill-posedness of the model is clearly seen in the decay of (the singular values) b_j towards 0 as $j \rightarrow +\infty$. As $j \rightarrow +\infty$, $b_j \theta_j$ gets weaker and it is then more difficult to estimate or detect the sequence $\theta = \{\theta_j\}_{j \in \mathbb{N}}$.

From the above discussion, it is evident that one can undertake statistical inference based on observations either from the GSM (1.1) or from the (equivalent) GWNM (1.2). Estimation in either models has received much attention over the last decades, providing also optimality results (in the minimax sense) over various loss functions and sequence/function spaces. Many methods have been considered including kernel, local polynomial, spline, projection and wavelet methods (see, e.g., [27], [10], [25], [6], [19]).

On the other hand, signal detection has received less attention. Minimax signal detection in the GSM (1.1) with $b_j = 1$ for all $j \in \mathbb{N}$ has been studied in [8] and in detail in the seminal work of [13], [14] and [15] (see also [18]). This work uses an asymptotic framework, that is, the noise level $\varepsilon > 0$ is allowed to *converge* to zero. A corresponding non-asymptotic framework, that is, for any *fixed* value of the noise level $\varepsilon > 0$, has been studied in [1] and [2]. Non-asymptotic and asymptotic studies for minimax signal detection in the GSM (1.1) with $b_j > 0$ for all $j \in \mathbb{N}$ have been recently considered in [17] and [22], respectively, in order to study minimax signal detection in ill-posed inverse problems. Despite the fact that the considered minimax signal detection problem is the same in the aforementioned studies, the final aims and the methodologies involved sometimes differ.

Bearing in mind the different issues and tasks involved, our aim below is to provide a unified treatment for non-asymptotic and asymptotic approaches to minimax signal detection in the GSM (1.1). In particular, we look for common ground between them that will enhance our understanding of these two existing minimax signal detection paradigms. This paper is organized as follows. Section 2 considers minimax signal detection from both non-asymptotic and asymptotic point of views. Section 3 discusses the construction of upper and lower bounds of the minimax separation radius (non-asymptotic approach) and the minimax separation rate (asymptotic approach) in a unified treatment, and points out several similarities. Section 4 brings into light hitherto unknown links between non-asymptotic and asymptotic approaches to minimax signal detection. It also contains an example of a mildly ill-posed inverse problem for illustrative purposes. In particular, it is shown that tools used to derive ‘asymptotic’ results can be exploited to draw ‘non-asymptotic’ conclusions, and vice-versa. Finally, Section 5 draws some concluding remarks and provides an avenue for future research.

Throughout the paper, we set the following notations. For all $x, y \in \mathbb{R}$, $\delta_x(y) = 1$ if $x = y$ and $\delta_x(y) = 0$ if $x \neq y$. Also, $x \wedge y := \inf\{x, y\}$ and $(x)_+ := \max\{0, x\}$. Given two collections $(c_\varepsilon)_{\varepsilon>0}$ and $(d_\varepsilon)_{\varepsilon>0}$ of real numbers, $c_\varepsilon \sim d_\varepsilon$ means that there exist $0 < \kappa_0 \leq \kappa_1 < \infty$ such that $\kappa_0 \leq c_\varepsilon/d_\varepsilon \leq \kappa_1$ for all $\varepsilon > 0$. In the same spirit, given two sequences $(c_j)_{j \in \mathbb{N}}$ and $(d_j)_{j \in \mathbb{N}}$ of real numbers, $c_j \asymp d_j$ means that there exist $0 < \kappa_0 \leq \kappa_1 < \infty$ such that $\kappa_0 \leq c_j/d_j \leq \kappa_1$ for all $j \in \mathbb{N}$. Finally, the abbreviation $o_\varepsilon(1)$ (resp. $\mathcal{O}_\varepsilon(1)$) will refer to a collection tending to 0 (resp. bounded) as ε tends to 0. When the dependence is not explicitly required on the noise level $\varepsilon > 0$, it will be simply denoted by $o(1)$ (resp. $\mathcal{O}(1)$).

2 Minimax signal detection

Statistical estimation is concerned with a quantitative question. Instead, we address below a qualitative question: given observations from the GSM (1.1), our aim is to compare the underlying (unknown) signal $\theta \in l^2(\mathbb{N})$ to a (known) benchmark signal θ_0 , i.e., to test

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta - \theta_0 \in \mathcal{F}, \quad (2.1)$$

for some given θ_0 and a given subspace \mathcal{F} . The statistical setting (2.1) is known as goodness-of-fit testing when $\theta_0 \neq 0$ and as signal detection when $\theta_0 = 0$.

Remark 2.1 Given observations from the GWNM (1.2), the test (2.1) is related to the test

$$H_0 : f = f_0 \text{ versus } H_1 : f - f_0 \in \tilde{\mathcal{F}}, \quad (2.2)$$

for a given benchmark function f_0 and a given subspace $\tilde{\mathcal{F}}$. In most cases, $\tilde{\mathcal{F}}$ contains functions $f \in L^2(V)$ that admit a Fourier series expansion with Fourier coefficients θ belonging to \mathcal{F} (see, e.g., [18], Section 3.2). In these cases, the problems (2.1) and (2.2) are equivalent.

The choice of the set \mathcal{F} is important. Indeed, it should be rich enough in order to contain the true θ . At the same time, if it is too rich, it will not be possible to control the performances of a given test due to the complexity of the problem. The common approach for such problems is to impose both a *regularity* condition (which characterizes the smoothness of the underlying signal) and an *energy* condition (which measures the amount of the underlying signal).

Concerning the regularity condition, we will work with certain ellipsoids in $l^2(\mathbb{N})$. In particular, we assume that $\theta \in \mathcal{E}_a(R)$, the set $\mathcal{E}_a(R)$ being defined as

$$\mathcal{E}_a(R) = \left\{ \theta \in l^2(\mathbb{N}), \sum_{j \in \mathbb{N}} a_j^2 \theta_j^2 \leq R \right\},$$

where $a = (a_j)_{j \in \mathbb{N}}$ denotes a non-decreasing sequence of positive real numbers with $a_j \rightarrow +\infty$ as $j \rightarrow +\infty$, and $R > 0$ is a constant. The set $\mathcal{E}_a(R)$ can be seen as a condition on the decay of θ . The cases where a increases very fast correspond to θ with a small amount of non-zero coefficients. In such a case, the corresponding signal can be considered as being ‘smooth’.

Without loss of generality, in what follows, we set $R = 1$. In order to simplify the notation, we will avoid the dependency to this term in all related quantities. In particular, we will write \mathcal{E}_a instead of $\mathcal{E}_a(1)$.

Regarding the energy condition, it will be measured in the $l^2(\mathbb{N})$ -norm. In particular, given $r_\varepsilon > 0$ (called the radius), which is allowed to depend on the noise level $\varepsilon > 0$, we will consider $\theta \in \mathcal{E}_a$ such that $\|\theta\| > r_\varepsilon$. Given a smoothness sequence a and a radius $r_\varepsilon > 0$, the set \mathcal{F} can thus be defined as

$$\mathcal{F} := \Theta_a(r_\varepsilon) = \{\theta \in \mathcal{E}_a, \|\theta\| \geq r_\varepsilon\}. \quad (2.3)$$

Since θ_0 and $b_j > 0$, $j \in \mathbb{N}$, are known, and assuming that $\theta_0 \in \mathcal{E}_a$, without loss of generality, given observations from the GSM (1.1), we restrict ourselves to the hypothesis testing setting (2.1) with $\theta_0 = 0$ (i.e., signal detection).

In summary, given observations from the GSM (1.1), we will be dealing with the following signal detection problem

$$H_0 : \theta = 0 \text{ versus } H_1 : \theta \in \Theta_a(r_\varepsilon), \quad (2.4)$$

where $\Theta_a(r_\varepsilon)$ is defined in (2.3). The sequence a being fixed, the main issue for the problem (2.4) is then to characterize the values of $r_\varepsilon > 0$ for which both hypotheses H_0 (called the null hypothesis) and H_1 (called the alternative hypothesis) are ‘separable’ (in a sense which will be made precise later on).

In the following, a (non-randomized) test $\Psi := \Psi(Y)$ will be defined as a measurable function of the observation $Y = (Y_j)_{j \in \mathbb{N}}$ from the GSM (1.1) having values in the set $\{0, 1\}$. By convention, H_0 is rejected if $\Psi = 1$ and H_0 is not rejected if $\Psi = 0$. Then, given a test Ψ , we can investigate

- the first kind error probability defined as

$$\alpha_\varepsilon(\Psi) := \mathbb{P}_0(\Psi = 1), \quad (2.5)$$

which measures the probability to reject H_0 when H_0 is true (i.e., $\theta = 0$); it is often constrained as being bounded by a prescribed level $\alpha \in]0, 1[$, and

- the maximal second kind error probability defined as

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi) := \sup_{\theta \in \Theta_a(r_\varepsilon)} \mathbb{P}_\theta(\Psi = 0), \quad (2.6)$$

which measures the worst possible probability not to reject H_0 when H_0 is not true (i.e., when $\theta \in \Theta_a(r_\varepsilon)$); one would like to ensure that it is (asymptotically) bounded by a prescribed level $\beta \in]0, 1[$.

For simplicity in our exposition, we will restrict ourselves to α -level tests, the value of $\alpha \in]0, 1[$ being fixed.

Definition 2.1 *A test Ψ_α is called an α -level test if*

$$\alpha_\varepsilon(\Psi_\alpha) \leq \alpha.$$

Given the trivial test $\Psi_\alpha := \alpha \in]0, 1[$, which does not depend on any observation, and extending the definition of a (non-randomized) test to a randomized test¹, it is easily seen that

$$\inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \beta_\varepsilon(\Theta_a(r_\varepsilon), \tilde{\Psi}_\alpha) \in [0, 1 - \alpha], \quad \text{for all } \alpha \in]0, 1[$$

(see, e.g., [18], pp. 10-11).

Definition 2.2 *A minimax hypothesis testing problem*

$$H_0 : \theta = 0 \quad \text{versus} \quad H_1 : \theta \in \mathcal{G},$$

for some set \mathcal{G} (with $0 \notin \mathcal{G}$), is called trivial if

$$\inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \beta_\varepsilon(\mathcal{F}, \tilde{\Psi}_\alpha) = 1 - \alpha \quad \text{for all } \alpha \in]0, 1[,$$

and is called asymptotical trivial if

$$\inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \beta_\varepsilon(\mathcal{F}, \tilde{\Psi}_\alpha) = 1 - \alpha + o_\varepsilon(1) \quad \text{for all } \alpha \in]0, 1[.$$

The regularity and energy conditions imposed above, when taken together, i.e., when \mathcal{F} is given by (2.3), result (provided the radius $r_\varepsilon > 0$ is ‘well-chosen’) in a non-trivial or an asymptotical non-trivial minimax signal detection problem (2.4). This means, in particular, that both hypotheses H_0 and H_1 are, in some sense, separable in such a framework. Two different point of views, the so-called non-asymptotic and asymptotic minimax signal detection approaches, are at hand, that have been respectively developed in, e.g., [1], [2], [22] and [17]. We elaborate on both approaches in the subsequent sections.

¹a measurable function $\Psi := \Psi(Y)$ of the observation $Y = (Y_j)_{j \in \mathbb{N}}$ from the GSM (1.1) with values in the interval $[0, 1]$: the null hypothesis is rejected with probability $\Psi(Y)$ and it is not rejected with probability $1 - \Psi(Y)$. In this case, $\alpha_\varepsilon(\Psi) := \mathbb{E}_0(\Psi(Y))$ and $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi) := \sup_{\theta \in \Theta_a(r_\varepsilon)} \mathbb{E}_\theta(1 - \Psi(Y))$.

2.1 The non-asymptotic approach

Let $\alpha, \beta \in]0, 1[$ be given, and let Ψ_α be an α -level test.

Definition 2.3 *The separation radius of the α -level test Ψ_α over the class \mathcal{E}_a is defined as*

$$r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta) := \inf \{r_\varepsilon > 0 : \beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha) \leq \beta\},$$

where the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha)$ is defined in (2.6).

In some sense, the separation radius $r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta)$ corresponds to the smallest possible value of the available signal $\|\theta\|$ for which H_0 and H_1 can be ‘separated’ by the α -level test Ψ_α with prescribed first and maximal second kind error probabilities, α and β , respectively.

Definition 2.4 *The minimax separation radius $\tilde{r}_\varepsilon := \tilde{r}_\varepsilon(\mathcal{E}_a, \alpha, \beta) > 0$ over the class \mathcal{E}_a is defined as*

$$\tilde{r}_\varepsilon := \inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} r_\varepsilon(\mathcal{E}_a, \tilde{\Psi}_\alpha, \beta). \quad (2.7)$$

The minimax separation radius \tilde{r}_ε corresponds to the smallest radius $r_\varepsilon > 0$ such that there exists some α -level test $\tilde{\Psi}_\alpha$ for which the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \tilde{\Psi}_\alpha)$ is not greater than β .

It is worth mentioning that Definitions 2.3 and 2.4 are valid for any *fixed* $\varepsilon > 0$ (i.e., it is not required that $\varepsilon \rightarrow 0$). The performances of any given test Ψ_α is easy to handle in the sense that the first kind error probability $\alpha_\varepsilon(\Psi_\alpha)$ is bounded by α (i.e., Ψ_α is an α -level test), and that the dependence of the minimax separation radius \tilde{r}_ε with respect to given α and β can be precisely described.

In practice, given an α -level test Ψ_α , it might be appropriate to compare its separation radius $r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta)$ to the minimax separation radius \tilde{r}_ε . Hence, the following definition is in order (see, e.g., [1], [22]).

Definition 2.5 *A α -level test Ψ_α is said to be powerful over the class \mathcal{E}_a if there exists a constant $\mathcal{C} \geq 1$ such that, for all $\varepsilon > 0$,*

$$\beta_\varepsilon(\Theta_a(\mathcal{C}\tilde{r}_\varepsilon), \Psi_\alpha) \leq \beta,$$

or, equivalently,

$$r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta) \leq \mathcal{C}\tilde{r}_\varepsilon,$$

for any given $\beta \in]0, 1[$.

According to Definition 2.5, for every $\varepsilon > 0$, the separation radius $r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta)$ of a powerful α -level test Ψ_α is of the order (up to a constant) of the minimax separation radius \tilde{r}_ε . In some sense, a powerful test appears to be *rate-optimal*.

We present below a general strategy for obtaining the minimax separation radius \tilde{r}_ε (that implicitly also produces a powerful α -level test Ψ_α). Given an ellipsoid \mathcal{E}_a , one has to find a radius $r_\varepsilon^* > 0$ such that

$$\text{(Lower bound)} \quad \tilde{r}_\varepsilon \geq r_\varepsilon^*,$$

and to construct a specific α -level test Ψ_α for which

$$\text{(Upper bound)} \quad r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta) \leq \mathcal{C}r_\varepsilon^*,$$

for some (explicitly obtained) constant $\mathcal{C} \geq 1$. It can be then easily seen that

$$r_\varepsilon^* \leq \tilde{r}_\varepsilon \leq \mathcal{C}r_\varepsilon^*.$$

More precisely,

Lower bound: It is enough to bound from below the following quantity

$$\inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \beta_\varepsilon(\Theta_a(r_\varepsilon), \tilde{\Psi}_\alpha),$$

for some radius $r_\varepsilon := r_\varepsilon^* > 0$. Indeed, if

$$\inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \beta_\varepsilon(\Theta_a(r_\varepsilon^*), \tilde{\Psi}_\alpha) \geq \beta, \quad (2.8)$$

for some $r_\varepsilon^* > 0$, then

$$\tilde{r}_\varepsilon \geq r_\varepsilon^*.$$

Upper bound: We first construct an α -level test Ψ_α . Then, we are looking for a radius $r_\varepsilon > 0$ such that, uniformly over all $\theta \in \mathcal{E}_a$,

$$\|\theta\| > r_\varepsilon \quad \Rightarrow \quad \mathbb{P}_\theta(\Psi_\alpha = 0) \leq \beta.$$

It is then evident that

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha) \leq \beta \quad \text{implying that} \quad r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta) \leq r_\varepsilon. \quad (2.9)$$

Finally, if $r_\varepsilon \leq \mathcal{C}r_\varepsilon^*$ for some $\mathcal{C} \geq 1$, it then follows immediately that

$$r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta) \leq \mathcal{C}r_\varepsilon^*.$$

(Note that the α -level test Ψ_α constructed above is powerful according to Definition 2.5.)

Figure 1 illustrates the areas where, according to Definitions 2.3–2.5, minimax signal detection can, or cannot, be possible.

We stress at this point that the quantity

$$\inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \beta_\varepsilon(\Theta_a(r_\varepsilon), \tilde{\Psi}_\alpha)$$

needed to bound from below in the above discussion is precisely the minimax second kind error probability to be introduced in the asymptotic approach that we elaborate in the following section.

2.2 The asymptotic approach

Let $\alpha \in]0, 1[$ be fixed and let $r_\varepsilon > 0$ be a given radius.

Definition 2.6 *The minimax second kind error probability is defined as*

$$\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon)) := \inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \beta_\varepsilon(\Theta_a(r_\varepsilon), \tilde{\Psi}_\alpha).$$

Given a radius $r_\varepsilon > 0$, the minimax second kind error probability $\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon))$ characterizes the minimax testing performances over all α -level tests $\tilde{\Psi}_\alpha$ for signal detection problem (2.4). In other words, it corresponds to the lowest maximal second kind error probability over the set $\Theta_a(r_\varepsilon)$. In particular, one would like to identify the different possible values of the radius $r_\varepsilon > 0$ such that the minimax second kind error probability $\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon))$ tends to 0 or to a constant or to 1, as ε tends to 0.

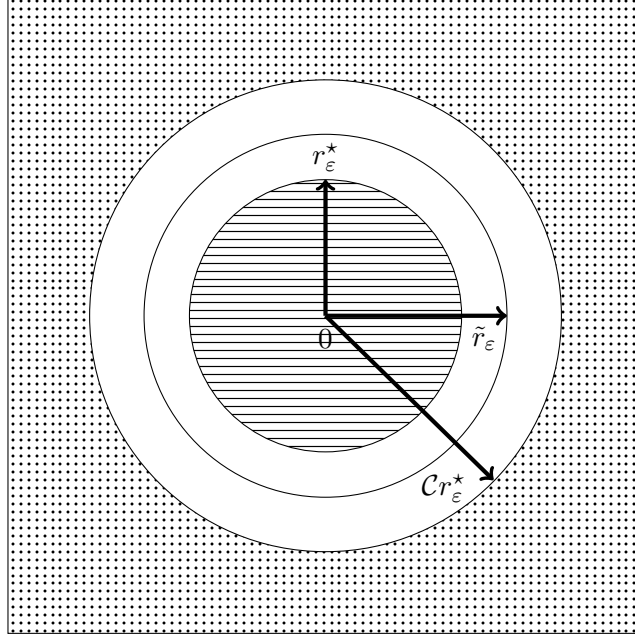


Figure 1: According to the lower bound (2.8), minimax signal detection is not possible inside the circle with center 0 and radius r_ε^* . According to the upper bound (2.9), for any given $\alpha \in]0, 1[$, the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha)$ of an α -level test Ψ_α can be controlled by a prescribed level $\beta \in]0, 1 - \alpha[$ outside the circle with center 0 and radius $C r_\varepsilon^*$, for some $C \geq 1$. The ‘optimal frontier’ is determined by the circle with center 0 and radius \tilde{r}_ε (i.e., the minimax separation radius).

Definition 2.7 The term $\bar{r}_\varepsilon := \bar{r}_\varepsilon(\mathcal{E}_a, \alpha) > 0$ is called the minimax separation rate if, for any given $r_\varepsilon > 0$,

$$\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon)) = 1 - \alpha + o_\varepsilon(1) \quad \text{if} \quad \frac{r_\varepsilon}{\bar{r}_\varepsilon} \rightarrow 0 \quad \text{as} \quad \varepsilon \rightarrow 0,$$

and

$$\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon)) = o_\varepsilon(1) \quad \text{if} \quad \frac{r_\varepsilon}{\bar{r}_\varepsilon} \rightarrow +\infty \quad \text{as} \quad \varepsilon \rightarrow 0.$$

The minimax separation rate \bar{r}_ε identifies, in some sense, the frontiers between detectable and undetectable signals. In other words, it means that, for small $\varepsilon > 0$, one can detect *all* $\theta \in \Theta_a(r_\varepsilon)$ for which the ratio $r_\varepsilon/\bar{r}_\varepsilon$ is large. On the other hand, if, for small $\varepsilon > 0$, the ratio $r_\varepsilon/\bar{r}_\varepsilon$ is small, it is then *impossible* to distinguish H_0 from H_1 with small maximal second kind error probability $\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon))$.

In practice, given an α -level test Ψ_α , it might be useful, for small $\varepsilon > 0$, to compare its maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha)$ to the minimax second kind error probability $\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon))$. Hence, the following definition is appropriate.

Definition 2.8 An α -level test Ψ_α is said to be

(i) asymptotical minimax consistent if, for any given $r_\varepsilon > 0$,

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha) = o_\varepsilon(1) \quad \text{if} \quad \frac{r_\varepsilon}{\bar{r}_\varepsilon} \rightarrow +\infty \quad \text{as} \quad \varepsilon \rightarrow 0.$$

(ii) asymptotical minimax if, for any given $r_\varepsilon > 0$,

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha) = \beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon)) + o_\varepsilon(1).$$

Regarding Definition 2.8, given an α -level test Ψ_α , item (i) provides a weak condition in the sense that, for small $\varepsilon > 0$, one can detect all $\theta \in \Theta_a(r_\varepsilon)$ for which the ratio $r_\varepsilon/\bar{r}_\varepsilon$ is large. On the other hand, item (ii) refers to a strong condition in the sense that one needs to asymptotically attain the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$.

In this setting, the point of view is asymptotic. The performance of any testing procedure is investigated as ε tends to 0. Nevertheless, such a point of view allows, sometimes, to provide a precise description of the asymptotic value for the minimax separation rate \bar{r}_ε . In particular, one can, in some cases, determine sharp asymptotics of Gaussian type for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$.

Definition 2.9 *The minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ is said to possess a sharp asymptotic of Gaussian type if it has an asymptotic Gaussian shape, i.e., if there exists a function $\nu(r_\varepsilon) \in]-\infty, \Phi^{-1}(1-\alpha)]$ (that should be determined later on) such that*

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) = \Phi(\nu(r_\varepsilon)) + o_\varepsilon(1),$$

where Φ denotes the distribution function of the standard Gaussian distribution.

Sharp asymptotics of Gaussian type for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ have been observed in particular settings (see e.g., [18] and references therein).

We present below a general strategy for obtaining the minimax separation rate \bar{r}_ε and sharp asymptotics of Gaussian type for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$. Given an ellipsoid \mathcal{E}_a , this amounts to investigate the construction of both lower and upper bounds on $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$.

Lower bound: Find a radius $r_{\varepsilon,1} > 0$ such that, for any given $r_\varepsilon > 0$,

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq 1 - \alpha + o_\varepsilon(1) \quad \text{if} \quad \frac{r_\varepsilon}{r_{\varepsilon,1}} \rightarrow 0 \quad \text{as} \quad \varepsilon \rightarrow 0.$$

If possible, one may also want to determine the shape of $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$, i.e., to find a function $\nu_1(r_\varepsilon) \in]-\infty, \Phi^{-1}(1-\alpha)]$ such that, for any given $r_\varepsilon > 0$,

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq \Phi(\nu_1(r_\varepsilon)) + o_\varepsilon(1).$$

Upper bound: Given an α -level test Ψ_α , find a radius $r_{\varepsilon,2} > 0$ such that, for any given $r_\varepsilon > 0$,

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha) = o_\varepsilon(1) \quad \text{if} \quad \frac{r_\varepsilon}{r_{\varepsilon,2}} \rightarrow +\infty \quad \text{as} \quad \varepsilon \rightarrow 0.$$

Additionally, one may again want to determine the shape of $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha)$, i.e., to find a function $\nu_2(r_\varepsilon) \in]-\infty, \Phi^{-1}(1-\alpha)]$ such that, for any given $r_\varepsilon > 0$,

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha) \leq \Phi(\nu_2(r_\varepsilon)) + o_\varepsilon(1).$$

If the α -level test Ψ_α is such that $r_{\varepsilon,1}/r_{\varepsilon,2} = \mathcal{O}_\varepsilon(1)$, then, obviously, $\bar{r}_\varepsilon/r_{\varepsilon,1} = \mathcal{O}_\varepsilon(1)$. It means that, according to Definition 2.7, either $r_{\varepsilon,1}$ or $r_{\varepsilon,2}$ correspond to the minimax separation rate \bar{r}_ε . Furthermore, in the case when $\nu_1(r_\varepsilon)/\nu_2(r_\varepsilon) = 1 + o_\varepsilon(1)$, then, according to Definition 2.9, we get sharp asymptotics of Gaussian type for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$, with $\nu(\cdot) = \nu_1(\cdot)$.

Figure 2 illustrates the areas where, according to Definition 2.7, minimax signal detection can, or cannot, be possible. It also illustrates, according to Definition 2.9, the area where sharp asymptotics of Gaussian type for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ are feasible.

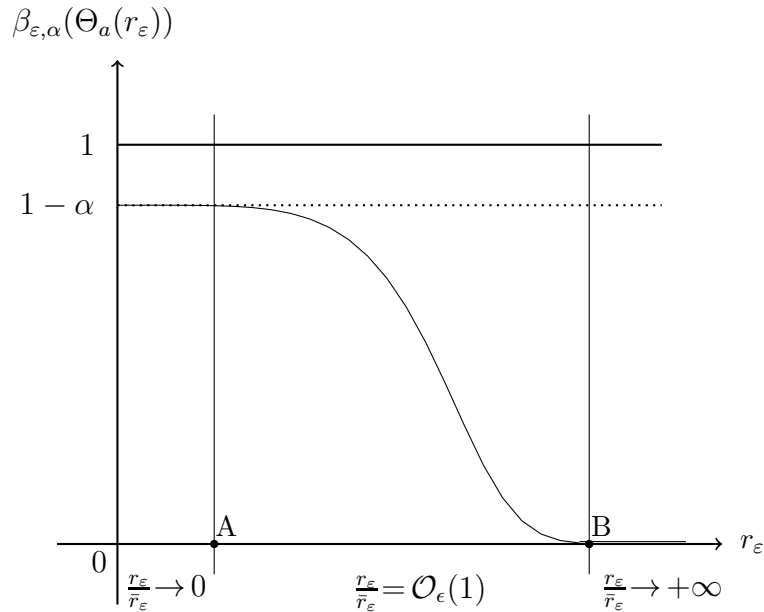


Figure 2: The interval $[0, A]$ (resp. $[B, +\infty[$) illustrates the area where $\frac{r_\epsilon}{\bar{r}_\epsilon} \rightarrow 0$ (resp. $\frac{r_\epsilon}{\bar{r}_\epsilon} \rightarrow +\infty$) as $\epsilon \rightarrow 0$, i.e., where the minimax second kind error probability $\beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon))$ satisfies $\beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon)) = 1 - \alpha + o_\epsilon(1)$ (resp. $\beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon)) = o_\epsilon(1)$) (see Definition 2.7). The interval $[A, B]$ determines the frontier for the minimax separation rate \bar{r}_ϵ . In particular, inside this area, sharp asymptotics of Gaussian type (solid curve) for $\beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon))$ are feasible, i.e., there exists a function $\nu(r_\epsilon) \in]-\infty, \Phi^{-1}(1 - \alpha)]$ such that $\beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon)) = \Phi(\nu(r_\epsilon)) + o_\epsilon(1)$ (see Definition 2.9).

2.3 A brief motivation

Although the considered minimax signal detection problem (2.4) is the same for both approaches (non-asymptotic and asymptotic), the way the optimality of the considered testing procedures is measured differs.

In the non-asymptotic setting, the statistician sets in advance some prescribed values $\alpha, \beta \in]0, 1[$. Then, the goal is to find ‘optimal’ (non-asymptotic) separation conditions for H_0 and H_1 that allow a precise (non-asymptotic) control of the first kind error probability and maximal second kind error probability by α and β , respectively. On the other hand, in the asymptotic setting, the aim is slightly different. Given any $r_\epsilon > 0$, the goal is to measure the best possible associated maximal second kind error probability of an (asymptotical) α -level test and to (asymptotically) determine whether it tends to $1 - \alpha$ or to 0, as the noise level ϵ tends to 0.

In order to study the signal detection problem (2.4), from a minimax point of view, different testing methodologies have been developed over the years that strongly depend on the two considered signal detection paradigms. We refer to, e.g., [1], [21], [22], for the non-asymptotic paradigm, and to, e.g., [13], [14], [15], [18], [17], for the asymptotic paradigm. Unsurprisingly, the results in these studies are coherent (i.e., the associated minimax separation radii \tilde{r}_ϵ and minimax separation rates \bar{r}_ϵ are *asymptotically* equivalent, as ϵ tends to 0). Indeed, one can formally prove (using the respective definitions) that $\tilde{r}_\epsilon/\bar{r}_\epsilon = O_\epsilon(1)$ as $\epsilon \rightarrow 0$.

In the sequel, we propose below a unified treatment for the study of the minimax separation radius \tilde{r}_ϵ (non-asymptotic approach) and the minimax separation rate \bar{r}_ϵ (asymptotic approach). We compare the construction of their lower and upper bounds and point out similarities in both

settings (Sections 3 and 4). In particular, tools constructed in the non-asymptotic paradigm can be used in order to draw conclusions in the asymptotic paradigm and vice-versa. In other words, one can perform asymptotic analysis for non-asymptotic testing procedures and investigate non-asymptotic performances for asymptotic testing procedures. This will be demonstrated later on, when explicit sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b = (b_j)_{j \in \mathbb{N}}$ are at hand (see Section 4.3).

3 Control of the lower and upper bounds

3.1 Control of the lower bounds

One of the main issues of minimax signal detection is to establish lower bounds for the minimax separation radius \tilde{r}_ε (non-asymptotic approach) and the minimax separation rate \bar{r}_ε (asymptotic approach). In both approaches, this amounts to determine the values of the available signal for which H_0 and H_1 cannot be separated with prescribed minimax second kind error probability $\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon))$.

More formally, we are interested to bound from below the minimax second kind error probability $\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon))$. In particular, an interesting question is to investigate the smallest possible value of the radius $r_\varepsilon > 0$ for which $\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon))$ can be, following the non-asymptotic or asymptotic approaches, (asymptotically) lower bounded by $\beta \in]0, 1 - \alpha[$ or tends to $1 - \alpha$.

A possible way to achieve this goal is to consider a (prior) probability measure π on the set associated with H_1 , i.e., a probability measure π on the set $\Theta_a(r_\varepsilon)$ (see, e.g., [1], [18]). Then, it is easily verified that

$$\begin{aligned}
\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon)) &\geq \inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \mathbb{P}_\pi(\tilde{\Psi}_\alpha = 0) \\
&= \inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \left[\mathbb{P}_0(\tilde{\Psi}_\alpha = 0) + \mathbb{P}_\pi(\tilde{\Psi}_\alpha = 0) - \mathbb{P}_0(\tilde{\Psi}_\alpha = 0) \right] \\
&\geq \inf_{\tilde{\Psi}_\alpha: \alpha_\varepsilon(\tilde{\Psi}_\alpha) \leq \alpha} \left[1 - \alpha - \left| \mathbb{P}_\pi(\tilde{\Psi}_\alpha = 0) - \mathbb{P}_0(\tilde{\Psi}_\alpha = 0) \right| \right] \\
&\geq 1 - \alpha - \sup_{A: \mathbb{P}_0(A) \leq \alpha} |\mathbb{P}_\pi(A) - \mathbb{P}_0(A)| \\
&\geq 1 - \alpha - \sup_{A \in \mathcal{A}} |\mathbb{P}_\pi(A) - \mathbb{P}_0(A)| \\
&= 1 - \alpha - V(\mathbb{P}_\pi, \mathbb{P}_0),
\end{aligned}$$

where

$$V(\mathbb{P}_\pi, \mathbb{P}_0) := \sup_{A \in \mathcal{A}} |\mathbb{P}_\pi(A) - \mathbb{P}_0(A)|$$

denotes the total variation norm between the two probability measures \mathbb{P}_0 and $\mathbb{P}_\pi = \int \mathbb{P}_\theta d\pi(\theta)$, and \mathcal{A} denotes the σ -field of the underlying probability space. Assuming that \mathbb{P}_π is absolutely continuous with respect to \mathbb{P}_0 , using first the Scheffé Theorem (see, e.g., [25], Lemma 2.1) and then the Cauchy-Schwarz inequality, it can be seen that

$$\begin{aligned}
V(\mathbb{P}_\pi, \mathbb{P}_0) &:= \sup_{A \in \mathcal{A}} |\mathbb{P}_\pi(A) - \mathbb{P}_0(A)| \\
&= \frac{1}{2} \int \left| \frac{d\mathbb{P}_\pi}{dy}(y) - \frac{d\mathbb{P}_0}{dy}(y) \right| dy \\
&= \frac{1}{2} \int \left| \frac{d\mathbb{P}_\pi}{d\mathbb{P}_0}(y) - 1 \right| d\mathbb{P}_0(y) \\
&\leq \frac{1}{2} (\mathbb{E}_0(|L_\pi(Y) - 1|^2))^{1/2},
\end{aligned}$$

where $L_\pi(Y)$ denotes the likelihood ratio between the two measures \mathbb{P}_π and \mathbb{P}_0 , and \mathbb{E}_0 denotes the expectation with respect to \mathbb{P}_0 . Combining the above arguments, we obtain the following lower bound

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq 1 - \alpha - \frac{1}{2} (\mathbb{E}_0[L_\pi^2(Y)] - 1)^{1/2}. \quad (3.1)$$

The construction of the lower bound for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ developed in (3.1) heavily relies on the construction of a prior π on the set $\Theta_a(r_\varepsilon)$. Given some sequence $\theta = (\theta_j)_{j \in \mathbb{N}}$, which will be made explicit below, we consider the symmetric prior π defined as

$$\pi = \prod_{j \in \mathbb{N}} \pi_j \quad \text{with} \quad \pi_j = \frac{1}{2} (\delta_{-\theta_j} + \delta_{\theta_j}) \quad \forall j \in \mathbb{N}. \quad (3.2)$$

(Note that $\pi(\Theta_a(r_\varepsilon)) = 1$.) Since the ξ_j are standard Gaussian random variables, we get, after some technical algebra (see [1] p. 596 or [17], supplementary material, Section 11.1), that

$$\mathbb{E}_0[L_\pi^2(Y)] = \prod_{j \in \mathbb{N}} \cosh(b_j^2 \theta_j^2 / \varepsilon^2) \leq \exp\left(\frac{1}{2\varepsilon^4} \sum_{j \in \mathbb{N}} b_j^4 \theta_j^4\right) := \exp(u_{\varepsilon,\theta}^2). \quad (3.3)$$

It is worth pointing out that the construction of the lower bound for the minimax separation radius \tilde{r}_ε (non-asymptotic approach) and the minimax separation rate \bar{r}_ε (asymptotic approach) are then both related to the study of either $\mathbb{E}_0[L_\pi^2(Y)]$ or its corresponding upper bound (3.3).

Two different interesting regimes at this point can be immediately deduced:

- First, $\mathbb{E}_0[L_\pi^2(Y)]$ tends to 1 as $\varepsilon \rightarrow 0$. In such a case, the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ is asymptotically lower bounded by $1 - \alpha$, i.e., $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq 1 - \alpha + o_\varepsilon(1)$.
- Second, $\mathbb{E}_0[L_\pi^2(Y)]$ can be upper bounded by a constant. In this case, the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ is also lower bounded by a constant. An interesting situation corresponds to the case where $\mathbb{E}_0[L_\pi^2(Y)] \leq 1 + 4(1 - \alpha - \beta)^2$ for some $\beta \in]0, 1 - \alpha[$. Then, $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq \beta$.

Moreover, a more delicate study of the term $u_{\varepsilon,\theta}^2 := \frac{1}{2\varepsilon^4} \sum_{j \in \mathbb{N}} b_j^4 \theta_j^4$ in (3.3) allows, under certain conditions to be made precise later on, to study sharp asymptotics of Gaussian type for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$.

We discuss below the two different strategies that have been investigated in the literature.

3.1.1 Non-asymptotic control

The following control has been proposed by [1] in the direct setting and it has been generalized to the inverse setting by [22]. The main idea consists of finding an explicit sequence $\theta^0 = (\theta_j^0)_{j \in \mathbb{N}}$ and a radius $r_\varepsilon > 0$ which satisfy the following three requirements:

- $\|\theta^0\| \geq r_\varepsilon$,
- $\exp\left(\frac{1}{2\varepsilon^4} \sum_{j \in \mathbb{N}} b_j^4 (\theta_j^0)^4\right) = 1 + 4(1 - \alpha - \beta)^2$,
- $\theta^0 \in \mathcal{E}_a$.

To this end, one can consider, for instance, the sequence θ^0 defined as

$$\theta_j^0 := \frac{r_\varepsilon \varepsilon^2 b_j^{-2}}{\left(\varepsilon^4 \sum_{k=1}^D b_k^{-4}\right)^{1/2}} \quad \forall j \in \{1, \dots, D\} \quad \text{and} \quad \theta_j^0 = 0 \quad \forall j > D, \quad (3.4)$$

for some (finite) parameter $D \in \mathbb{N}$ (called the bandwidth), that possibly depends on $\varepsilon > 0$.

It is evident that $\|\theta^0\| = r_\varepsilon$. Furthermore, taking into account (3.3), we get

$$\mathbb{E}_0[L_\pi^2(Y)] \leq \exp\left(\frac{1}{2\varepsilon^4} \sum_{j \in \mathbb{N}} b_j^4 (\theta_j^0)^4\right) = \exp\left[\frac{r_\varepsilon^4}{2\varepsilon^4 \sum_{j=1}^D b_j^{-4}}\right] = 1 + 4(1 - \alpha - \beta)^2, \quad (3.5)$$

as soon as

$$r_\varepsilon^2 = r_{\varepsilon, D}^2 := c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4}},$$

where

$$c(\alpha, \beta) = (2 \ln(1 + 4(1 - \alpha - \beta)^2))^{1/4} > 0. \quad (3.6)$$

In order to conclude, it remains to choose an appropriate $D \in \mathbb{N}$ such that $\theta^0 \in \mathcal{E}_a$. To this end, note that

$$\sum_{j \in \mathbb{N}} a_j^2 (\theta_j^0)^2 \leq a_D^2 \sum_{j=1}^D (\theta_j^0)^2 = a_D^2 r_{\varepsilon, D}^2 \leq 1 \quad \text{as soon as} \quad r_{\varepsilon, D}^2 \leq a_D^{-2}.$$

Hence, if we define

$$r_{\varepsilon, \star}^2 := \sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right], \quad (3.7)$$

we get, using (3.1)-(3.5),

$$\beta_{\varepsilon, \alpha}(\Theta_a(r_{\varepsilon, \star})) \geq \beta,$$

which means that the minimax separation radius \tilde{r}_ε satisfies

$$\tilde{r}_\varepsilon \geq r_{\varepsilon, \star}.$$

This corresponds to a non-asymptotic lower bound on the minimax separation radius \tilde{r}_ε . The main advantage of such a bound is that it provides a precise description of the area where minimax signal detection is impossible with prescribed values $\alpha, \beta \in]0, 1[$.

3.1.2 Asymptotic control

In the previous (non-asymptotic) approach, the main idea was to construct an explicit sequence θ and to control $\mathbb{E}_0[L_\pi^2(Y)]$. In the asymptotic approach, one instead starts from (3.3) and find the smallest possible value of $u_{\varepsilon, \theta}^2$ for which $\theta \in \Theta_a(r_\varepsilon)$. In other words, the idea is to choose a sequence $\bar{\theta} := \bar{\theta}(r_\varepsilon)$ as the solution of the following extremal problem

$$\bar{\theta}(r_\varepsilon) := \operatorname{arginf}_{\theta \in \Theta_a(r_\varepsilon)} \left\{ u_{\varepsilon, \theta}^2 := \frac{1}{2\varepsilon^4} \sum_{k \in \mathbb{N}} b_k^4 \theta_k^4 \right\}. \quad (3.8)$$

In the following, we will denote the solution of the extremal problem (3.8) as

$$u_\varepsilon(r_\varepsilon) := u_{\varepsilon, \bar{\theta}(r_\varepsilon)} := \inf_{\theta \in \Theta_a(r_\varepsilon)} \left\{ \frac{1}{2\varepsilon^4} \sum_{k \in \mathbb{N}} b_k^4 \theta_k^4 \right\}. \quad (3.9)$$

This idea has been in particular developed in the series of papers [13], [14], [15], or, more recently, in [17], in an inverse problem framework. The cases of interest correspond to the setting where $u_\varepsilon^2(r_\varepsilon)$ either tends to zero or is bounded by a constant. In that case, one can find the solution of $u_\varepsilon(r_\varepsilon)$ in (3.9) using, for instance, the standard methodology of Lagrange multipliers.

Let $\alpha \in]0, 1[$ be fixed. We immediately see from (3.1) and (3.3) that if $u_\varepsilon(r_\varepsilon) = o_\varepsilon(1)$, then

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq 1 - \alpha + o_\varepsilon(1). \quad (3.10)$$

The interesting situation, however, arises when $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$. It allows a more accurate study to asymptotically precise the shape of the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$. In particular, if $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$, it can be established that

$$\ln(L_\pi(Y)) = -\frac{u_\varepsilon^2(r_\varepsilon)}{2} + u_\varepsilon(r_\varepsilon)\xi_\varepsilon + \zeta_\varepsilon, \quad (3.11)$$

where $\xi_\varepsilon \rightarrow \xi \sim \mathcal{N}(0, 1)$ and $\zeta_\varepsilon \rightarrow 0$ (in \mathbb{P}_0 -probability) as $\varepsilon \rightarrow 0$ distribution (see Section 4.3.1 of [18] or the proof of Theorem 4.1 of [17], supplementary material, Section 11.1). By a standard change of probability measure, it follows that

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq \mathbb{E}_\pi(1 - \psi^*) = \mathbb{E}_0(\exp(\ln(L_\pi(Y)))(1 - \psi^*)),$$

where ψ^* is the likelihood ratio test defined as $\psi^* = \mathbf{1}_{\{\ln(L_\pi(Y)) > t_{1-\alpha}^*\}}$ with $t_{1-\alpha}^*$ being the $(1-\alpha)$ -quantile of the distribution of $\ln(L_\pi(Y))$ under H_0 . Hence, in view of (3.11), it is easily seen that

$$t_{1-\alpha}^* = -\frac{u_\varepsilon^2(r_\varepsilon)}{2} + u_\varepsilon(r_\varepsilon)t_{1-\alpha} + o_\varepsilon(1),$$

where $t_{1-\alpha}$ refers to the $(1-\alpha)$ -quantile of the standard Gaussian distribution. Moreover, using the mean value theorem, it follows that

$$\begin{aligned} \beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq \mathbb{E}_\pi(1 - \psi^*) &= \mathbb{E}_0(\exp(\ln(L_\pi(Y)))(1 - \psi^*)) \\ &= \Phi(t_{1-\alpha} - u_\varepsilon(r_\varepsilon)) + o_\varepsilon(1). \end{aligned} \quad (3.12)$$

(Note that, in the particular case where $r_\varepsilon > 0$ satisfies $u_\varepsilon(r_\varepsilon) = t_{1-\alpha} - t_\beta$, then, it is immediately seen that $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq \beta + o_\varepsilon(1)$.)

Remark 3.1 It is worth mentioning that one cannot determine at this point the radius $r_{\varepsilon,1} > 0$ (considered in the general strategy of Section 2.2 for constructing lower bounds), unless the sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b = (b_j)_{j \in \mathbb{N}}$ are explicitly given. We refer to, e.g., [17] for more details or to Section 4.3 where a mildly ill-posed inverse problem is treated for illustrative purposes.

In the following section, we investigate upper bounds on the minimax separation radius \tilde{r}_ε (non-asymptotic approach) and upper bounds on the minimax separation rate \bar{r}_ε (asymptotic approach). In the latter setting, we also provide, under mild conditions, sharp asymptotics of Gaussian type for the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$.

3.2 Control of the upper bounds

3.2.1 A general testing methodology

In this section, we construct appropriate tests and investigate the associated separation radius (non-asymptotic approach) and the maximal second kind error probability (asymptotic approach). Starting from signal detection problem (2.4), the underlying question is to decide whether we observe a signal or not. To this end, a possible approach is to construct an estimator \hat{d} of $d := \|\theta\|^2$ or $d := \|b\theta\|^2$. Indeed, the assertions “ $\theta = 0$ ” and “ $b\theta = 0$ ” are equivalent since the sequence $b = \{b_j\}_{j \in \mathbb{N}}$ is assumed strictly positive. We refer to [21] for an extended discussion on that subject. Then, one can use the following decision rule:

- if \hat{d} is large enough (larger than a prescribed threshold which should be precisely quantified), we reject H_0 ,

- If \hat{d} is smaller than this threshold, we do not reject H_0 .

In order to estimate $\|\theta\|^2$ (resp. $\|b\theta\|^2$), one can first construct a preliminary estimator of θ (resp. $b\theta$) and then take its squared norm. This idea has been widely investigated. We point out that, in general, the preliminary estimators cannot be directly plugged in order to estimate $\|\theta\|^2$ (resp. $\|b\theta\|^2$). Indeed, minimax estimation and minimax testing are essentially two different problems, see, e.g., [18], Sections 1.4.4 and 2.10. Nevertheless, ideas and methodologies in minimax estimation can inspire the construction of appropriate minimax testing procedures.

In the following, we focus on the construction of linear estimators based on observations from the GSM (1.1). Let $\omega = (\omega_j)_{j \in \mathbb{N}}$ be a filter, i.e., a sequence taking values in the interval $[0, 1]$. Then, one can estimate $\|\theta\|^2$ by the following estimator

$$\widehat{\|\theta\|^2} = \sum_{j \in \mathbb{N}} \omega_j b_j^{-2} (y_j^2 - \varepsilon^2). \quad (3.13)$$

or, in the same spirit, estimate $\|b\theta\|^2$ by the following estimator

$$\widehat{\|b\theta\|^2} = \sum_{j \in \mathbb{N}} \omega_j (y_j^2 - \varepsilon^2). \quad (3.14)$$

Various possible filters $\omega = (\omega_j)_{j \in \mathbb{N}}$ are available in the literature. Among them, one can mention, e.g., spectral cut-off filters (see Section 3.2.2), Tikhonov filters, Ingster filters (see Section 3.2.3) or filters based on other regularization approaches. For more details regarding available regularization methods, we refer, e.g., to [3], [7] and [18].

Having an estimator $\|\theta\|^2$ (resp. $\|b\theta\|^2$) of the form (3.13) (resp. (3.14)), denoted by \hat{d} , we can construct an associated test $\Psi_{\alpha, \omega}$ as

$$\Psi_{\alpha, \omega} = \mathbf{1}_{\{\hat{d} > t_{\alpha, \omega}\}},$$

where $t_{\alpha, \omega}$ is a threshold that (asymptotically) controls the first kind error probability $\alpha_\varepsilon(\Psi_{\alpha, \omega})$.

It is important to point out at this point that, having an (asymptotic) α -level test Ψ_α ,

- (non-asymptotic approach) one can try to determine the smallest possible separation radius $r_{\varepsilon, 0} := r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta) > 0$ such that the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_{\varepsilon, 0}), \Psi_\alpha)$ is at most β , for any prescribed $\alpha, \beta \in]0, 1[$,
- (asymptotic approach) one can investigate the asymptotic behavior of the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_\alpha)$, for any given $r_\varepsilon > 0$ and any prescribed $\alpha \in]0, 1[$.

Our aim below is,

1. to construct appropriate tests that reach (at least up to a constant) the lower bounds established in Sections 3.1.1 and 3.1.2 (Sections 3.2.2 and 3.2.3),
2. to bring into light hitherto unknown links between non-asymptotic and asymptotic approaches to minimax signal detection (see Section 4).

As mentioned previously, there exist several possible available filters. We focus below on two different kind of filters investigated in, e.g., [22] and [17], namely, spectral cut-off and Ingster filters, respectively.

3.2.2 Non-asymptotic control: Spectral cut-off filters

Our aim is to propose an α -level test Ψ_α such that

$$r_\varepsilon(\mathcal{E}_a, \Psi_\alpha, \beta) \leq Cr_{\varepsilon, \star},$$

for some $C \geq 1$, where $r_{\varepsilon, \star} > 0$ has been introduced in (3.7). In such a case, this will mean that lower and upper bounds for the minimax separation radius $\tilde{r}_\varepsilon > 0$ match together, up to a constant.

According to the previous discussion, the suggested test will be based on an estimation of $\|\theta\|^2$ (using (3.13)). More formally, given a bandwidth $D \in \mathbb{N}$, we define

$$\begin{aligned} \Psi_{D,P} &:= \mathbf{1}_{\{\sum_{j=1}^D b_j^{-2}(y_j^2 - \varepsilon^2) > t_{1-\alpha, D}\}} \\ &:= \mathbf{1}_{\{T_{D,P} > t_{1-\alpha, D}\}}, \end{aligned} \quad (3.15)$$

where

$$T_{D,P} := \sum_{j=1}^D b_j^{-2}(y_j^2 - \varepsilon^2) \quad (3.16)$$

and $t_{1-\alpha, D}$ denotes the $(1-\alpha)$ -quantile of $T_{D,P}$ under H_0 , i.e., the $(1-\alpha)$ -quantile of the random variable $\varepsilon^2 \sum_{j=1}^D b_j^{-2}(\xi_j^2 - 1)$.

Due to the definition of $t_{1-\alpha, D}$, the spectral cut-off test $\Psi_{D,P}$ is an α -level test. Indeed,

$$\begin{aligned} \alpha_\varepsilon(\Psi_{D,P}) := \mathbb{P}_0(\Psi_{D,P} = 1) &= \mathbb{P}_0(T_{D,P} > t_{1-\alpha, D}) \\ &= \mathbb{P}\left(\varepsilon^2 \sum_{j=1}^D b_j^{-2}(\xi_j^2 - 1) > t_{1-\alpha, D}\right) = \alpha. \end{aligned}$$

Now, we turn to the control of the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{D,P})$. To this end, denote by $t_{\beta, D}(\theta)$ the β -quantile of $T_{D,P}$ under H_1 , i.e., the term satisfying

$$\mathbb{P}_\theta(T_{D,P} \leq t_{\beta, D}(\theta)) = \beta.$$

Then, for a given $\theta \in \mathcal{E}_a$, in order to prove that

$$\mathbb{P}_\theta(\Psi_{D,P} = 0) := \mathbb{P}_\theta(T_{D,P} \leq t_{1-\alpha, D}) = \mathbb{P}_\theta\left(\sum_{j=1}^D b_j^{-2}(y_j^2 - \varepsilon^2) \leq t_{1-\alpha, D}\right) \leq \beta,$$

it suffices to show that

$$t_{1-\alpha, D} \leq t_{\beta, D}(\theta). \quad (3.17)$$

Figure 3 provides, for a fixed bandwidth $D \in \mathbb{N}$, a heuristic illustration for the comparison (3.17) between the $(1-\alpha)$ -quantile $t_{1-\alpha, D}$ and the β -quantile $t_{\beta, D}(\theta)$ of the test statistic $T_{D,P}$, defined in (3.16). In order to compare these two terms formally, we use the following proposition.

Proposition 3.1 *Let $T_{D,P}$ be the test statistic defined in (3.16), and let $t_{1-\alpha, D}$ and $t_{\beta, D}(\theta)$ be its $(1-\alpha)$ -quantile under H_0 and β -quantile under H_1 . Then, there exists a constant (explicitly computable) $C(\alpha) > 0$ such that*

$$t_{1-\alpha, D} \leq C(\alpha)\varepsilon^2 \left(\sum_{j=1}^D b_j^{-4}\right)^{1/2},$$

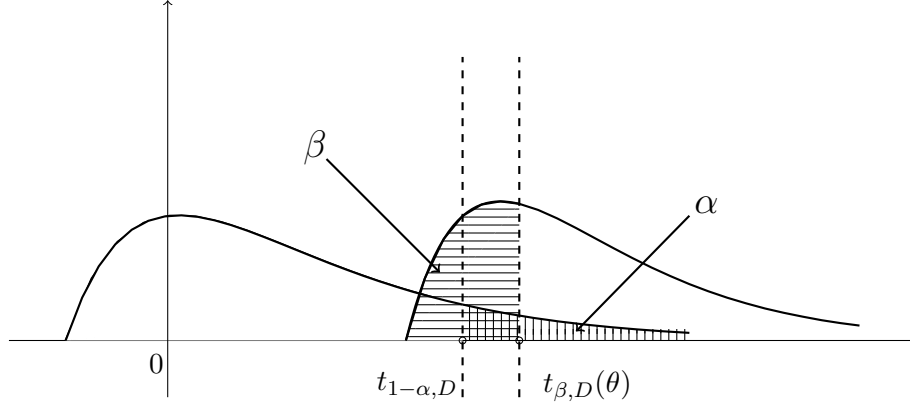


Figure 3: *Illustrative comparison of the $(1 - \alpha)$ -quantile $t_{1-\alpha,D}$ and the β -quantile $t_{\beta,D}(\theta)$ for a fixed bandwidth $D \in \mathbb{N}$. The left hand side curve displays the density of the test statistic $T_{D,P} := \sum_{j=1}^D b_j^{-2}(y_j^2 - \epsilon^2)$, defined in (3.16), under H_0 , while the one on the right hand side displays the density of the same test statistics $T_{D,P}$ under H_1 . The shaded areas are determined by the corresponding $(1 - \alpha)$ -quantile $t_{1-\alpha,D}$ (vertical lines) and β -quantile $t_{\beta,D}(\theta)$ (horizontal lines).*

and

$$t_{\beta,D} \geq \sum_{j=1}^D \theta_j^2 - 2\sqrt{\ln(1/\beta)} \sqrt{\epsilon^4 \sum_{j=1}^D b_j^{-4} + 2\epsilon^2 \sum_{j=1}^D b_j^{-2} \theta_j^2}.$$

The proof of Proposition 3.1 can be found in [22] (see the construction of the upper and lower bounds in the proof of their Proposition 2). In particular, the control of $t_{1-\alpha,D}$ and $t_{\beta,D}(\theta)$ is based on deviation inequalities of appropriate independent weighted- χ^2 random variables.

Using (3.17) and Proposition 3.1, one can easily see that

$$t_{1-\alpha,D} \leq t_{\beta,D}(\theta)$$

if and only if

$$\sum_{j=1}^D \theta_j^2 \geq C(\alpha)\epsilon^2 \left(\sum_{j=1}^D b_j^{-4} \right)^{1/2} - 2\sqrt{\ln(1/\beta)} \sqrt{\epsilon^4 \sum_{j=1}^D b_j^{-4} + 2\epsilon^2 \sum_{j=1}^D b_j^{-2} \theta_j^2},$$

which, in turn, holds true as soon as

$$\sum_{j=1}^D \theta_j^2 \geq C(\alpha, \beta)\epsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4}}, \quad (3.18)$$

where, setting $x_\gamma = \ln(1/\gamma)$, for all $\gamma \in]0, 1[$,

$$C(\alpha, \beta) = \sqrt{2x_\beta} + \sqrt{2(x_\alpha + x_\beta)} + \sqrt{2} (\sqrt{x_\alpha} + \sqrt{x_\beta})^{1/2} > 0 \quad (3.19)$$

(see [22] for more details). The condition (3.18) ensures that

$$\mathbb{P}_\theta(\Psi_{D,P} = 0) \leq \beta.$$

The main drawback of (3.18) is that it is expressed in terms of a lower bound on $\sum_{j=1}^D \theta_j^2$ instead of $\|\theta\|^2$. However, since $\theta \in \mathcal{E}_a$, it follows that $\sum_{j>D} \theta_j^2 \leq a_D^{-2}$. Hence,

$$\forall \theta \in \mathcal{E}_a, \quad \|\theta\|^2 \geq C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \quad \Rightarrow \quad \mathbb{P}_\theta(\Psi_{D,P} = 0) \leq \beta. \quad (3.20)$$

Moreover, we point out that the term in the left hand side of (3.20) corresponds to the sum of two antagonist quantities. Since our aim is to obtain the weaker possible bound on the energy condition, we choose a bandwidth $D := D^* \in \mathbb{N}$ such that

$$\beta_\varepsilon(\Theta_a(r_\varepsilon^*), \Psi_{D^*,P}) := \sup_{\theta \in \Theta_a(r_\varepsilon^*)} \mathbb{P}_\theta(\Psi_{D^*,P} = 0) \leq \beta,$$

where

$$(r_\varepsilon^*)^2 := \inf_{D \in \mathbb{N}} \left[C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \right]. \quad (3.21)$$

Finally, thanks to (3.7) and (3.21), the values of $r_{\varepsilon, \star}$ and r_ε^* are coherent. In Section 4, we show that, under some weak conditions on the sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b = (b_j)_{j \in \mathbb{N}}$, non-asymptotic lower and upper bounds for the minimax separation radius \tilde{r}_ε match together, up to a constant.

3.2.3 Asymptotic control: Ingster filters

We consider a different approach, since the testing procedure will be based on an estimation of $\|b\theta\|^2$ (using (3.14)). We will deal with a specific kind of filters which have been, to the best of our knowledge, introduced by Yuri I. Ingster in a series of papers (see, e.g., [13], [14], [15], [18]).

Let $\bar{\theta} = \bar{\theta}(r_\varepsilon) \in \Theta_a(r_\varepsilon)$ be the solution of the extremal problem (3.8). Then, we define the Ingster filters $\omega_{j,r_\varepsilon} = (\omega_{j,r_\varepsilon})_{j \in \mathbb{N}}$ as

$$\omega_{j,r_\varepsilon} = \frac{b_j^2 \bar{\theta}_j^2}{\sqrt{2 \sum_{k \in \mathbb{N}} b_k^4 \bar{\theta}_k^4}} \quad \forall j \in \mathbb{N}. \quad (3.22)$$

As discussed in Section 3.2.1, one can use a test of the form

$$\begin{aligned} \Psi_{r_\varepsilon, I} &:= \mathbf{1}_{\{\sum_{j \in \mathbb{N}} \omega_{j,r_\varepsilon} (y_j^2 - \varepsilon^2) > \varepsilon^2 t_{1-\alpha}\}} \\ &= \mathbf{1}_{\{\sum_{j \in \mathbb{N}} \omega_{j,r_\varepsilon} ((y_j/\varepsilon)^2 - 1) > t_{1-\alpha}\}} \\ &:= \mathbf{1}_{\{T_{r_\varepsilon, I} > t_{1-\alpha}\}}, \end{aligned} \quad (3.23)$$

where

$$T_{r_\varepsilon, I} = \sum_{j \in \mathbb{N}} \omega_{j,r_\varepsilon} \left(\left(\frac{y_j}{\varepsilon} \right)^2 - 1 \right), \quad (3.24)$$

and $t_{1-\alpha}$ denotes the $(1 - \alpha)$ -quantile of a standard Gaussian random variable.

Since $T_{r_\varepsilon, I} = \xi_\varepsilon$, where ξ_ε is the quantity appeared in (3.11), from the proof of the corresponding lower bounds, it follows that $T_{r_\varepsilon, I} \rightarrow \xi \sim \mathcal{N}(0, 1)$ (in \mathbb{P}_0 probability) as $\varepsilon \rightarrow 0$. Hence, we immediately see that,

$$\alpha_\varepsilon(\Psi_{r_\varepsilon, I}) = \Phi(t_{1-\alpha}) + o_\varepsilon(1) = \alpha + o_\varepsilon(1). \quad (3.25)$$

(This means that $\Psi_{r_\varepsilon, I}$ is, asymptotically, an α -level test.)

We now consider the corresponding maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon, I})$. The following cases are of particular interest:

- $u_\varepsilon(r_\varepsilon) = o_\varepsilon(1)$. According to (3.10), $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \geq 1 - \alpha + o_\varepsilon(1)$. It is then impossible to distinguish H_0 from H_1 , meaning that, according to Definition 2.2, we have an asymptotical trivial test. Thus, it is not needed to further study this case.
- $u_\varepsilon(r_\varepsilon) = O_\varepsilon(1)$. In this case, under the mild condition $\sup_{j \in \mathbb{N}} \omega_{j,r_\varepsilon} = o_\varepsilon(1)$, we establish a sharp asymptotic of Gaussian type for $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon,I})$.
- $u_\varepsilon(r_\varepsilon) \rightarrow \infty$ as $\varepsilon \rightarrow 0$. In this case, we establish that $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon,I}) = o_\varepsilon(1)$.

To this end, simple algebra leads to the following expressions of the expectation and the variance of the test statistics $T_{r_\varepsilon,I}$:

$$\mathbb{E}_\theta[T_{r_\varepsilon,I}] = \varepsilon^{-2} \sum_{j \in \mathbb{N}} \omega_{j,r_\varepsilon} b_j^2 \theta_j^2, \quad \text{Var}_\theta[T_{r_\varepsilon,I}] = 1 + 4\varepsilon^{-2} \sum_{j \in \mathbb{N}} \omega_{j,r_\varepsilon}^2 b_j^2 \theta_j^2. \quad (3.26)$$

Introduce the standardized random variable $\tilde{T}_{r_\varepsilon,I}$ defined as

$$\tilde{T}_{r_\varepsilon,I} = \frac{T_{r_\varepsilon,I} - \mathbb{E}_\theta[T_{r_\varepsilon,I}]}{\sqrt{\text{Var}_\theta[T_{r_\varepsilon,I}]}}$$

where the $\mathbb{E}_\theta[T_{r_\varepsilon,I}]$ and $\text{Var}_\theta[T_{r_\varepsilon,I}]$ have been computed in (3.26). Define

$$\omega_{0,r_\varepsilon} := \sup_{j \in \mathbb{N}} \frac{b_j^2 \bar{\theta}_j^2}{\sqrt{2 \sum_{k \in \mathbb{N}} b_k^4 \bar{\theta}_k^4}} := \sup_{j \in \mathbb{N}} \omega_{j,r_\varepsilon}, \quad (3.27)$$

where $\bar{\theta} = \bar{\theta}(r_\varepsilon) \in \Theta_a(r_\varepsilon)$ is the extremal sequence, i.e., the solution of the extremal problem (3.8). (Note that, using (3.27), $1 \leq \text{Var}_\theta[T_{r_\varepsilon,I}] \leq 1 + 4\omega_{0,r_\varepsilon} \mathbb{E}_\theta[T_{r_\varepsilon,I}]$.)

In order to proceed, we need the following proposition.

Proposition 3.2 *Let $T_{r_\varepsilon,I}$ be the test statistic defined in (3.24) Let $h(r_\varepsilon, \theta) := \mathbb{E}_\theta[T_{r_\varepsilon,I}]$, where $\mathbb{E}_\theta[T_{r_\varepsilon,I}]$ is computed in (3.26). Then*

$$\inf_{\theta \in \Theta_a(r_\varepsilon)} h(r_\varepsilon, \theta) = u_\varepsilon(r_\varepsilon).$$

The proof of Proposition 3.2 can be found in [17], supplementary material, Lemma 11.1.

Case 1 ($u_\varepsilon(r_\varepsilon) = O_\varepsilon(1)$) Following the proof of Theorem 4.1 of [17], supplementary material, Section 11.1, using Lyapunov's conditions and (3.26), it follows that, as soon as $\omega_{0,r_\varepsilon} = o_\varepsilon(1)$,

- $\tilde{T}_{r_\varepsilon,I}$ is asymptotically standard Gaussian under \mathbb{P}_θ .
- $\text{Var}_\theta[T_{r_\varepsilon,I}] = 1 + o_\varepsilon(1)$.

Therefore, we get that

$$\mathbb{P}_\theta(\Psi_{r_\varepsilon,I} = 0) = \mathbb{P}_\theta \left(\tilde{T}_{r_\varepsilon,I} \leq \frac{t_{1-\alpha} - \mathbb{E}_\theta[T_{r_\varepsilon,I}]}{\sqrt{\text{Var}_\theta[T_{r_\varepsilon,I}]}} \right) = \Phi(t_{1-\alpha} - \mathbb{E}_\theta[T_{r_\varepsilon,I}]) + o_\varepsilon(1).$$

Using Proposition 3.2, we arrive at

$$\begin{aligned} \beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon,I}) &= \sup_{\theta \in \Theta_a(r_\varepsilon)} \mathbb{P}_\theta(\Psi_{r_\varepsilon,I} = 0) \\ &= \sup_{\theta \in \Theta_a(r_\varepsilon)} \Phi(t_{1-\alpha} - \mathbb{E}_\theta[T_{r_\varepsilon,I}]) + o_\varepsilon(1) \\ &:= \sup_{\theta \in \Theta_a(r_\varepsilon)} \Phi(t_{1-\alpha} - h(r_\varepsilon, \theta)) + o_\varepsilon(1) \\ &= \Phi \left(t_{1-\alpha} - \inf_{\theta \in \Theta_a(r_\varepsilon)} h(r_\varepsilon, \theta) \right) + o_\varepsilon(1) \\ &= \Phi(t_{1-\alpha} - u_\varepsilon(r_\varepsilon)) + o_\varepsilon(1). \end{aligned} \quad (3.28)$$

Therefore, using (3.28), we get that

$$\begin{aligned}\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) &\leq \beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon, I}) \\ &= \Phi(t_{1-\alpha} - u_\varepsilon(r_\varepsilon)) + o_\varepsilon(1).\end{aligned}\tag{3.29}$$

(Note that, in the particular case that $r_\varepsilon > 0$ satisfies $u_\varepsilon(r_\varepsilon) = t_{1-\alpha} - t_\beta$, it is immediately seen that $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \leq \beta + o_\varepsilon(1)$.)

Case 2 ($u_\varepsilon(r_\varepsilon) \rightarrow +\infty$ as $\varepsilon \rightarrow 0$) Using Proposition 3.2 and (3.26), it follows that, for all $\theta \in \Theta_a(r_\varepsilon)$,

$$\mathbb{E}_\theta[T_{r_\varepsilon, I}] := h(r_\varepsilon, \theta) \geq \inf_{\theta \in \Theta_a(r_\varepsilon)} h(r_\varepsilon, \theta) = u_\varepsilon(r_\varepsilon) \rightarrow +\infty \quad \text{as } \varepsilon \rightarrow 0.$$

Therefore, using Markov's inequality,

$$\begin{aligned}\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon, I}) &:= \sup_{\theta \in \Theta_a(r_\varepsilon)} \mathbb{P}_\theta(\Psi = 0) \\ &= \sup_{\theta \in \Theta_a(r_\varepsilon)} \mathbb{P}_\theta \left(\tilde{T}_{r_\varepsilon, I} \leq \frac{t_{1-\alpha} - \mathbb{E}_\theta[T_{r_\varepsilon, I}]}{\sqrt{\text{Var}_\theta[T_{r_\varepsilon, I}]}} \right) \\ &\leq \sup_{\theta \in \Theta_a(r_\varepsilon)} \mathbb{P}_\theta \left(|\tilde{T}_{r_\varepsilon, I}| \geq \frac{\mathbb{E}_\theta[T_{r_\varepsilon, I}] - t_{1-\alpha}}{\sqrt{\text{Var}_\theta[T_{r_\varepsilon, I}]}} \right) \\ &\leq \sup_{\theta \in \Theta_a(r_\varepsilon)} \frac{\text{Var}_\theta[T_{r_\varepsilon, I}]}{(t_{1-\alpha} - \mathbb{E}_\theta[T_{r_\varepsilon, I}])^2} \\ &\leq \sup_{\theta \in \Theta_a(r_\varepsilon)} \frac{1 + 4\omega_{0, r_\varepsilon} h(r_\varepsilon, \theta)}{(h(r_\varepsilon, \theta) - t_{1-\alpha})^2} \\ &\sim \frac{1}{\inf_{\theta \in \Theta_a(r_\varepsilon)} h(r_\varepsilon, \theta)} := \frac{1}{u_\varepsilon(r_\varepsilon)} = o_\varepsilon(1).\end{aligned}\tag{3.30}$$

Remark 3.2 It is worth mentioning that one cannot determine at this point the radius $r_{\varepsilon, 2} > 0$ (considered in the general strategy of Section 2.2 for constructing upper bounds). This more or less amounts to solve the equation $u_\varepsilon(r_{\varepsilon, 2}) = \mathcal{O}_\varepsilon(1)$ as $\varepsilon \rightarrow 0$. This cannot be accomplished unless the sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b = (b_j)_{j \in \mathbb{N}}$ are explicitly given. We refer again to, e.g., [17] for more details or to Section 4.3 where an example of a mildly ill-posed inverse problem is treated for illustrative purposes.

Below, we first formalize the results for the lower and upper bounds presented above and explain their meaning for practical purposes (Section 4.1). We then bring into light hitherto unknown links between non-asymptotic and asymptotic approaches to minimax signal detection (Section 4.2 and Section 4.3).

4 Connections between non-asymptotic and asymptotic frameworks

4.1 General Results

In Section 3, lower and upper bounds on the minimax separation radius \tilde{r}_ε and the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ were independently treated. In the following theorems, these results are gathered in unified manners.

We first focus our attention to the non-asymptotic paradigm.

Theorem 4.1 (*Non-asymptotic framework*) Assume that $Y = (Y_j)_{j \in \mathbb{N}}$ are observations from the GSM (1.1), and consider the signal detection problem (2.4) with \mathcal{F} defined in (2.3). Let $\alpha, \beta \in]0, 1[$ be given. Then, for every $\varepsilon > 0$, the minimax separation radius \tilde{r}_ε is controlled by

$$\sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \leq \tilde{r}_\varepsilon^2 \leq \inf_{D \in \mathbb{N}} \left[C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \right], \quad (4.1)$$

where the constants $c(\alpha, \beta)$ and $C(\alpha, \beta)$ are respectively given in (3.6) and (3.19).

In order to shed some light on the meaning of (4.1), the following comments are in order:

- One cannot ensure that both lower and upper bounds on the minimax separation radius \tilde{r}_ε in (4.1) match together, unless (weak) conditions on the sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b = (b_j)_{j \in \mathbb{N}}$ are at hand. A discussion on that point is provided below (see Theorem 4.1).
- Nevertheless, we point out that these bounds are coherent since they involve the same quantities, namely, a_D^{-2} and $\varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4}}$, for any given bandwidth $D \in \mathbb{N}$, as well as positive constants $c(\alpha, \beta)$ and $C(\alpha, \beta)$, depending on α and β only.
- A careful look into the discussion presented in the previous section indicates that the term a_D^{-2} can be related to a (in fact an upper bound on the) ‘bias’ term in the sense that it measures the amount of signal that is missed using the spectral cut-off test $\Psi_{D,P}$ (see, e.g., (3.18)). Recall that the sequence $a = (a_j)_{j \in \mathbb{N}}$ characterizes the smoothness of the underlying signal θ . Obviously, the smoother the signal of interest, the easier the testing problem in the sense that the minimax separation radius \tilde{r}_ε becomes smaller.
- In the same spirit, $\varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4}}$ can be related to a ‘standard deviation’ term that corresponds to the estimation of the term $\|\theta\|^2$ using the spectral cut-off test $\Psi_{D,P}$. When $b_j = 1$, for all $j \in \mathbb{N}$, (i.e., the direct problem) this term is of order $\varepsilon^2 \sqrt{D}$. This particular case has been discussed in detail in [1], Section 3. On the other hand, the case when $b_j \rightarrow 0$ as $j \rightarrow +\infty$, corresponds to ill-posed inverse problems. In this case, the signal detection problem becomes harder in the sense that the minimax separation radius \tilde{r}_ε strongly depends on the decay of the sequence $b = (b_j)_{j \in \mathbb{N}}$ towards 0 and becomes larger than the corresponding one in the direct problem.

In summary, in order to precisely compute the minimax separation radius \tilde{r}_ε , explicit sequences of $a = (a_j)_{j \in \mathbb{N}}$ and $b = (b_j)_{j \in \mathbb{N}}$ are needed to control the trade-off between the two antagonist terms, i.e., the ‘bias’ and the ‘standard deviation’ terms, a_D^{-2} and $\varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4}}$, respectively. This will be elaborated on Section 4.3 below, where an example of a mildly ill-posed inverse problem is used for illustrative purposes.

We now turn our attention to the asymptotic paradigm.

Theorem 4.2 (*Asymptotic framework*) Assume that $Y = (Y_j)_{j \in \mathbb{N}}$ are observations from the GSM (1.1), and consider the signal detection problem (2.4) with \mathcal{F} defined in (2.3). Let a radius $r_\varepsilon > 0$ be fixed, and let $\alpha \in]0, 1[$ be given. Let $u_\varepsilon(r_\varepsilon)$ and ω_{0,r_ε} denote the solution of the extremal problem defined in (3.9) and the term introduced in (3.27), respectively.

(a) If

$$u_\varepsilon(r_\varepsilon) = o_\varepsilon(1),$$

then

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) = 1 - \alpha + o_\varepsilon(1).$$

(b) If

$$u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1) \quad \text{and} \quad \omega_{0,r_\varepsilon} = o_\varepsilon(1),$$

then

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) = \Phi(t_{1-\alpha} - u_\varepsilon(r_\varepsilon)) + o_\varepsilon(1).$$

(c) If

$$u_\varepsilon(r_\varepsilon) \rightarrow +\infty \quad \text{as} \quad \varepsilon \rightarrow 0,$$

then

$$\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) = o_\varepsilon(1).$$

It is evident from Theorem 4.2 that the minimax signal detection problem in the asymptotic framework essentially reduces to the study of the extremal problem (3.8). Indeed, the corresponding solution given in (3.9) governs both the lower and the upper bounds on the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$. The three different regimes mentioned in Theorem 4.2 are of particular interest and require at this step some additional explanations:

- If $u_\varepsilon(r_\varepsilon) = o_\varepsilon(1)$, then, according to Definition 2.2, an asymptotical non-trivial minimax hypothesis testing problem is not possible. In other words, it is impossible to distinguish between H_0 and H_1 .
- If $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$ and $\omega_{0,r_\varepsilon} = o_\varepsilon(1)$, then one can precisely describe the shape of the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ since it possesses a sharp asymptotic of Gaussian type. It is also evident that $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) \in]0, 1 - \alpha[$. This means that the minimax signal detection problem is asymptotically non-trivial (i.e., $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) > 0$) but that H_0 and H_1 can be asymptotically always separated (i.e., $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) < 1 - \alpha$). Note that in this particular case that $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$ and $\omega_{0,r_\varepsilon} = o_\varepsilon(1)$, the Ingster test $\Psi_{r_\varepsilon,I}$ defined in (3.23) is asymptotical minimax according to Definition 2.8.
- If $u_\varepsilon(r_\varepsilon) \rightarrow +\infty$ as $\varepsilon \rightarrow 0$, then the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) = o_\varepsilon(1)$. In particular, the test $\Psi_{r_\varepsilon,I}$ constructed in (3.23)-(3.24) asymptotically always separates H_0 from H_1 .

Remark 4.1 Theorem 4.2 does not treat the case where

$$u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1) \quad \text{and} \quad \omega_{0,r_\varepsilon} \not\rightarrow 0 \quad \text{as} \quad \varepsilon \rightarrow 0. \quad (4.2)$$

In such a case, the lower bound (3.12) is still valid but can be, in fact, improved by showing that

$$\liminf_{\varepsilon \rightarrow 0} \beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon)) > 1 - \alpha \quad \text{for any} \quad \alpha \in]0, 1[,$$

i.e., the minimax signal detection problem is asymptotically trivial (see the proof of Theorem 4.1 of [17], supplementary material, Section 11.1.). It is worth pointing out at this point that if (4.2) holds, then the minimax second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$ asymptotically belongs to the set $\{0, 1 - \alpha\}$, for any $\alpha \in]0, 1[$, depending on the behavior of any given $r_\varepsilon > 0$.

The case where $\omega_{0,r_\varepsilon} \not\rightarrow 0$ as $\varepsilon \rightarrow 0$ exists in, e.g., the case of severely ill-posed inverse problems with the class of analytic functions (super-smooth functions), i.e., $b_j \asymp e^{-jt}$, $j \in \mathbb{N}$, for some $t > 0$, and $a_j \asymp e^{js}$, $j \in \mathbb{N}$, for some $s > 0$, respectively. Indeed,

$$\omega_{0,r_\varepsilon} := \sup_{j \in \mathbb{N}} \frac{b_j^2 \bar{\theta}_j^2}{\sqrt{2 \sum_{k \in \mathbb{N}} b_k^4 \bar{\theta}_k^4}} \sim \frac{z_0^2 e^{-2tm}}{z_0^2 e^{-2tm}} \asymp 1 \not\rightarrow 0, \quad \text{as} \quad \varepsilon \rightarrow 0,$$

for some quantities $z_0 \in \mathbb{R}$ and $m \in [1, \infty)$ (see Theorem 4.3 and Remark 4.4 in [17]). We also refer to Section 4.3.3 below for a similar computation in a mildly ill-posed inverse problem setting.

Remark 4.2 Theorem 4.2 does not provide an immediate expression for the minimax separation rate \bar{r}_ε . In practice, however, both terms $r_{\varepsilon,1}$ and $r_{\varepsilon,2}$ required in the construction of the lower and upper bounds, respectively, sketched in Section 2.2, are derived from the same equation: $u_\varepsilon(r_{\varepsilon,1}) = u_\varepsilon(r_{\varepsilon,2}) = \mathcal{O}_\varepsilon(1)$. Then one can, ‘in general’, check the implications

$$\frac{r_\varepsilon}{r_{\varepsilon,1}} \rightarrow 0 \Rightarrow u_\varepsilon(r_\varepsilon) = o_\varepsilon(1) \quad \text{and} \quad \frac{r_\varepsilon}{r_{\varepsilon,2}} \rightarrow +\infty \Rightarrow u_\varepsilon(r_\varepsilon) \rightarrow +\infty, \quad \text{as } \varepsilon \rightarrow 0,$$

which, thanks to Theorem 4.2 and Definition 2.7, allows one to conclude. As mentioned previously, this task cannot be accomplished unless explicit expressions on the sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b = (b_j)_{j \in \mathbb{N}}$ are given. Explicit calculation of the minimax separation rate \bar{r}_ε in a mildly ill-posed inverse problem is provided in Section 4.3.2.

The proofs of the assertions in Theorem 4.1 and Theorem 4.2 are direct consequences of the discussion provided in Section 3, concerning the control of the upper and lower bounds, for both minimax separation radius \tilde{r}_ε and maximal second kind error probability $\beta_{\varepsilon,\alpha}(\Theta_a(r_\varepsilon))$. Detailed arguments and related discussions can be found in, e.g., [1], [17] and [22].

4.2 Deriving the minimax separation rate \bar{r}_ε from bounds on the minimax separation radius \tilde{r}_ε .

The following theorem shows that, under some mild conditions on the growth of the sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b^{-1} = (b_j^{-1})_{j \in \mathbb{N}}$, one can derive the asymptotic order of the minimax separation rate \bar{r}_ε from the bounds on the minimax separation radius \tilde{r}_ε given in (4.1).

Proposition 4.1 *Assume that $Y = (Y_j)_{j \in \mathbb{N}}$ are observations from the GSM (1.1), and consider the signal detection problem (2.4) with \mathcal{F} defined in (2.3). Assume that both sequences $a = (a_j)_{j \in \mathbb{N}}$ and $b^{-1} = (b_j^{-1})_{j \in \mathbb{N}}$ are non-decreasing and that they satisfy*

$$a_\star \leq \frac{a_{D-1}}{a_D} \leq a^\star \quad \text{and} \quad b_\star \leq \frac{b_{D-1}}{b_D} \leq b^\star \quad \text{for all } D > 1, \quad (4.3)$$

for some constants $0 < a_\star \leq a^\star < \infty$ and $0 < b_\star \leq b^\star < \infty$. Let $\alpha, \beta \in]0, 1[$ be given. Then, there exists a constant $C \geq 1$ such that

$$\inf_{D \in \mathbb{N}} \left[C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \right] \leq C \sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right],$$

where the constants $c(\alpha, \beta)$ and $C(\alpha, \beta)$ are respectively given in (3.6) and (3.19). In particular, both lower and upper bounds in (4.1) are of the same order.

The control (4.1) hence proposes a sharp description of the minimax separation radius \tilde{r}_ε as soon as (4.3) is satisfied. Since $\bar{r}_\varepsilon/\tilde{r}_\varepsilon = O_\varepsilon(1)$, the asymptotic minimax separation rate \bar{r}_ε can thus be determine from (4.1). On the other hand, a non-asymptotic bound that matches asymptotic known results can be considered as (rate) optimal. Hence, although the motivations differ, both asymptotic and non-asymptotic approaches provide a similar description on the minimax signal detection problem at hand.

Remark 4.3 We note also that the condition (4.3) is satisfied for various combinations of interest, among them: (i) mildly ill-posed inverse problems ($b_j \asymp j^{-t}$, $j \in \mathbb{N}$, for some $t > 0$) with ordinary smooth functions ($a_j \asymp j^s$, $j \in \mathbb{N}$, for some $s > 0$), (ii) severely ill-posed inverse problems ($b_j \asymp e^{-jt}$, $j \in \mathbb{N}$, for some $t > 0$) with ordinary smooth functions ($a_j \asymp j^s$, $j \in \mathbb{N}$, for some $s > 0$), and (iii) mildly ill-posed inverse problems ($b_j \asymp j^{-t}$, $j \in \mathbb{N}$, for some $t > 0$) with super-smooth functions ($a_j \asymp e^{js}$, $j \in \mathbb{N}$, for some $s > 0$). Among the possible situations where the condition (4.3) is not satisfied, one can mention, for instance, power-exponential behaviors ($a_j \asymp e^{j^l}$, $j \in \mathbb{N}$, for some $s > 0$ and $l > 1$, or $b_j \asymp e^{-j^r}$, $j \in \mathbb{N}$, for some $t > 0$ and $r > 1$).

PROOF. Let the bandwidth $D_0 \in \mathbb{N}$ satisfy

$$D_0 = \arg \sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right]. \quad (4.4)$$

We restrict ourselves to the following case

$$a_{D_0}^{-2} \leq c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D_0} b_j^{-4}}.$$

(The other case follows similarly along the same lines of proof.) Then, thanks to (4.3), we get

$$\begin{aligned} \inf_{D \in \mathbb{N}} \left[C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \right] &\leq C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D_0} b_j^{-4} + a_{D_0}^{-2}} \\ &\leq (C(\alpha, \beta) + c(\alpha, \beta)) \varepsilon^2 \sqrt{\sum_{j=1}^{D_0} b_j^{-4}} \\ &\leq C \varepsilon^2 \sqrt{\sum_{j=1}^{D_0-1} b_j^{-4}}, \end{aligned}$$

for some constant $C > 0$ that can be explicitly computed. Note that

$$a_{D_0}^{-2} \leq c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D_0} b_j^{-4}} \quad \text{implies} \quad a_{D_0-1}^2 > c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^{D_0-1} b_j^{-4}},$$

since, otherwise, we arrive at a contradiction, due to the definition of $D_0 \in \mathbb{N}$ in (4.4). Hence,

$$\begin{aligned} \inf_{D \in \mathbb{N}} \left[C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \right] &\leq C \varepsilon^2 \sqrt{\sum_{j=1}^{D_0-1} b_j^{-4}}, \\ &\leq C \left[\varepsilon^2 \sqrt{\sum_{j=1}^{D_0-1} b_j^{-4} \wedge a_{D_0-1}^{-2}} \right], \\ &\leq C \sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right]. \end{aligned}$$

In other words, there exists some constant $C \geq 1$ such that

$$\begin{aligned} &\sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \\ &\leq \inf_{D \in \mathbb{N}} \left[C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \right] \leq C \sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right]. \quad (4.5) \end{aligned}$$

Hence, the lower and the upper bounds in (4.1) are of the same order. This concludes the proof of the proposition. \square

Remark 4.4 According to Definition 2.8, and as soon as (4.3) is satisfied, the spectral cut-off test $\Psi_{D,P}$ defined in (3.15), with bandwidth $D := D_0 \in \mathbb{N}$ selected as in (4.4), is asymptotical minimax consistent. Indeed, given $\theta \in \mathcal{E}_a$ and a radius r_ε such that $\|\theta\| \geq r_\varepsilon$,

$$\begin{aligned} \mathbb{P}_\theta(\Psi_{D_0,P} = 0) &= \mathbb{P}_\theta(T_{D_0,P} \leq t_{1-\alpha,D_0}), \\ &= \mathbb{P}_\theta\left(T_{D_0,P} - \mathbb{E}[T_{D_0,P}] \leq t_{1-\alpha,D_0} - \sum_{j=1}^{D_0} \theta_j^2\right), \\ &\leq \mathbb{P}_\theta\left(|\mathbb{E}[T_{D_0,P}] - T_{D_0,P}| \geq \sum_{j=1}^{D_0} \theta_j^2 - t_{1-\alpha,D_0}\right), \\ &\leq \frac{\varepsilon^4 \sum_{j=1}^{D_0} b_j^{-4}}{\left(\sum_{j=1}^{D_0} \theta_j^2 - t_{1-\alpha,D_0}\right)^2}. \end{aligned}$$

Then, since $\bar{r}_\varepsilon/\tilde{r}_\varepsilon = O_\varepsilon(1)$ as $\varepsilon \rightarrow 0$, using Proposition 3.1, we get

$$\begin{aligned} \sum_{j=1}^{D_0} \theta_j^2 - t_{1-\alpha,D_0} &\geq \|\theta\|^2 - C \left(\varepsilon^2 \sqrt{\sum_{j=1}^{D_0} b_j^{-4}} + a_{D_0}^{-2} \right), \\ &\geq r_\varepsilon^2 - C\tilde{r}_\varepsilon^2, \\ &\geq r_\varepsilon^2(1 - o_\varepsilon(1)), \end{aligned}$$

as soon as $r_\varepsilon/\tilde{r}_\varepsilon \rightarrow +\infty$ as $\varepsilon \rightarrow 0$. Finally, we obtain

$$\mathbb{P}_\theta(\Psi_{D_0,P} = 0) \leq \frac{C\tilde{r}_\varepsilon^4}{r_\varepsilon^4(1 - o_\varepsilon(1))} = o_\varepsilon(1),$$

which entails

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{D_0,P}) = o_\varepsilon(1) \quad \text{if} \quad \frac{r_\varepsilon}{\tilde{r}_\varepsilon} \rightarrow +\infty \quad \text{as} \quad \varepsilon \rightarrow 0.$$

4.3 An illustrative example: a mildly ill-posed inverse problem

Our aim below is to illustrate the results presented in Section 4.1 and Section 4.2. To this end, we address the minimax signal detection problem of a mildly ill-posed inverse problem. Namely, we will assume that

$$a_j \asymp j^s, \quad \text{for some } s > 0, \quad \text{and} \quad b_j \asymp j^{-t}, \quad \text{for some } t > 0, \quad \text{for all } j \in \mathbb{N}. \quad (4.6)$$

Our aim in this context is multifold:

- First, we consider an asymptotic analysis of the minimax separation radius \tilde{r}_ε based on the inequality (4.1).
- Second, we explicitly compute the minimax separation rate \bar{r}_ε through a careful analysis of the extremal problem (3.9).
- Third, we provide a non-asymptotic analysis of the Ingster test $\Psi_{r_\varepsilon,I}$, defined in (3.23). In particular, we compute its associated separation radius $r_\varepsilon(\mathcal{E}_a, \Psi_{r_\varepsilon,I}, \beta)$ and show that, up to constant, it coincides with the minimax separation radius \tilde{r}_ε .
- Fourth, we present an asymptotic analysis of the spectral cut-off test $\Psi_{D,P}$, defined in (3.15). In particular, for an appropriate bandwidth $D := \tilde{D} \in \mathbb{N}$, we prove that the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D},P})$, is asymptotically bounded from above by a quantity that possesses a Gaussian shape.

These results demonstrate that tools used to derive ‘asymptotic’ results can be exploited to draw ‘non-asymptotic’ conclusions, and vice-versa.

4.3.1 Asymptotic analysis of the minimax separation radius \tilde{r}_ε

We are interested in the asymptotic behavior of the minimax separation radius \tilde{r}_ε . Recall from Theorem 4.1 that, for any $\varepsilon > 0$,

$$\sup_{D \in \mathbb{N}} \left[c(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} \wedge a_D^{-2}} \right] \leq \tilde{r}_\varepsilon^2 \leq \inf_{D \in \mathbb{N}} \left[C(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4} + a_D^{-2}} \right].$$

Moreover, according to (4.5), both the upper and the lower bounds in the above inequality are of the same order. Indeed, the constraint (4.3) of Theorem 4.1 is satisfied in the setting (4.6). Hence, we are now able to characterize the asymptotic value of the minimax separation radius \tilde{r}_ε .

Simple algebra shows that

$$\sum_{j=1}^D b_j^{-4} = \sum_{j=1}^D j^{4t} = CD^{4t+1}(1 + o(1)) \quad \text{as } D \rightarrow +\infty,$$

for some constant $C > 0$. Hence, the bandwidth $D_0 \in \mathbb{N}$, introduced in (4.4), satisfies

$$\varepsilon^2 \sqrt{\sum_{j=1}^{D_0} b_j^{-4}} = \mathcal{O}(a_{D_0}^{-2}) \quad \text{iff} \quad \varepsilon^2 D_0^{2t+1/2} = \mathcal{O}(D_0^{-2s}) \quad \text{iff} \quad D_0 = \mathcal{O}_\varepsilon \left(\varepsilon^{\frac{-2}{2s+2t+1/2}} \right).$$

We then deduce from the previous computation that the minimax separation radius \tilde{r}_ε satisfies

$$\tilde{r}_\varepsilon^2 = \mathcal{O}(D_0^{-2s}) = \mathcal{O}_\varepsilon \left(\varepsilon^{\frac{4s}{2s+2t+1/2}} \right). \quad (4.7)$$

4.3.2 Computation of the minimax separation rate \bar{r}_ε

Following Remark 4.2, an explicit computation of the function $r_\varepsilon \mapsto u_\varepsilon(r_\varepsilon)$ is required in order to retrieve the minimax separation rate \bar{r}_ε from the solution of the equation $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$.

We first need to solve the extremal problem (3.9) defined as

$$u_\varepsilon^2(r_\varepsilon) = \frac{1}{2\varepsilon^4} \inf_{\theta \in \Theta_a(r_\varepsilon)} \sum_{j \in \mathbb{N}} b_j^4 \theta_j^4. \quad (4.8)$$

This problem is solved via Lagrange multipliers. In particular, the extremal sequence, i.e., the solution of the above mentioned extremal problem, appears to be of the form

$$\bar{\theta}_j^2(r_\varepsilon) = z_0^2 b_j^{-4} (1 - \mathcal{A} a_j^2)_+, \quad j \in \mathbb{N},$$

where the quantities $z_0 := z_{0,\varepsilon}$ and $\mathcal{A} := \mathcal{A}_\varepsilon$ are determined by the equations

$$\sum_{j \in \mathbb{N}} b_j^{-2} \bar{\theta}_j^2(r_\varepsilon) \quad \text{and} \quad \sum_{j \in \mathbb{N}} a_j^2 b_j^{-2} \bar{\theta}_j^2(r_\varepsilon) = 1. \quad (4.9)$$

Remark 4.5 The quantity \mathcal{A} determines the so-called *efficient dimension* m in specific ill-posed inverse problems: since a_j is an increasing sequence, the efficient dimension is the quantity $m = m_\varepsilon \in [1, \infty)$ such that $\mathcal{A} a_{[m]}^2 \leq 1 < \mathcal{A} a_{[m]+1}^2$, see, e.g., [17], supplementary material, Section 11. Moreover, a unique solution to the system of equations (4.9) exists for $r_\varepsilon > 0$ small enough, due to the fact that $\sum_{j \in \mathbb{N}} b_j^{-1} = +\infty$ (see, Proposition 11.2 of [17], supplementary material, Section 11).

The equations (4.8)-(4.9) are immediately rewritten in the form

$$\begin{cases} r_\varepsilon^2 = z_0^2 J_1, \\ 1 = z_0^2 \mathcal{A}^{-1} J_2, \end{cases} \quad (4.10)$$

with

$$\begin{aligned} J_1 &= \sum_{j \in \mathbb{N}} b_j^{-4} (1 - \mathcal{A} a_j^2)_+, \\ J_2 &= \mathcal{A} \sum_{j \in \mathbb{N}} a_j^2 b_j^{-4} (1 - \mathcal{A} a_j^2)_+. \end{aligned}$$

In particular, the extremal problem (3.9) takes the form

$$u_\varepsilon^2(r_\varepsilon) = \varepsilon^{-4} z_0^4 J_0 / 2, \quad \text{where } J_0 = J_1 - J_2 = \sum_{j \in \mathbb{N}} b_j^{-4} (1 - \mathcal{A} a_j^2)_+^2. \quad (4.11)$$

Setting $\mathcal{R} = \mathcal{A}^{-1/2s}$, simple computations lead to

$$\begin{aligned} J_1 &= \sum_{j \in \mathbb{N}} b_j^{-4} (1 - \mathcal{A} a_j^2)_+, \\ &= \sum_{j: a_j^2 \leq \mathcal{A}^{-1}} b_j^{-4} - \mathcal{A} \sum_{j: a_j^2 \leq \mathcal{A}^{-1}} b_j^{-4} a_j^2, \\ &= \sum_{j: j \leq \mathcal{R}} j^{4t} - \mathcal{A} \sum_{j: j \leq \mathcal{R}} j^{4t+2s}, \\ &= \mathcal{C}_1 \mathcal{R}^{4t+1} (1 + o(1)) \quad \text{as } \mathcal{R} \rightarrow +\infty. \end{aligned}$$

Using similar algebra, one can prove that

$$J_2 = \mathcal{C}_2 \mathcal{R}^{4t+1} (1 + o(1)) \quad \text{and} \quad J_0 = \mathcal{C}_0 \mathcal{R}^{4t+1} (1 + o(1)) \quad \text{as } \mathcal{R} \rightarrow +\infty.$$

In particular, we get from (4.10) that

$$r_\varepsilon^2 = \mathcal{A} \frac{J_1}{J_2} = \mathcal{R}^{-2s} \frac{\mathcal{C}_1}{\mathcal{C}_2} (1 + o(1)) \quad \text{as } \mathcal{R} \rightarrow +\infty. \quad (4.12)$$

Therefore, combining the above results,

$$\begin{aligned} u_\varepsilon(r_\varepsilon) &= \left(\frac{r_\varepsilon}{\varepsilon} \right)^4 \frac{J_0}{2J_1^2} \\ &= \left(\frac{r_\varepsilon}{\varepsilon} \right)^4 \frac{\mathcal{C}_0 \mathcal{R}^{4t+1}}{2\mathcal{C}_2^2 \mathcal{R}^{2(4t+1)}} (1 + o(1)) \quad \text{as } \mathcal{R} \rightarrow +\infty, \\ &= \left(\frac{r_\varepsilon}{\varepsilon} \right)^4 \frac{r_\varepsilon^{-(4t+1)/s}}{r_\varepsilon^{-(8t+2)/s}} \mathcal{O}(1) \quad \text{as } r_\varepsilon \rightarrow 0, \\ &= \mathcal{O} \left(\varepsilon^{-4} r_\varepsilon^{(4s+4t+1)/s} \right) \quad \text{as } r_\varepsilon \rightarrow 0. \end{aligned} \quad (4.13)$$

The expression in (4.13) provides an explicit form for the function $r_\varepsilon \mapsto u_\varepsilon(r_\varepsilon)$ that is required in order to retrieve the minimax separation rate \bar{r}_ε from the solution of the equation $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$. Our next task is to solve this equation. Using (4.13), we immediately get

$$u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1) \quad \text{iff} \quad \varepsilon^{-4} r_\varepsilon^{(4s+4t+1)/s} = \mathcal{O}_\varepsilon(1) \quad \text{iff} \quad r_\varepsilon = \mathcal{O}_\varepsilon \left(\varepsilon^{\frac{2s}{2s+2t+1/2}} \right). \quad (4.14)$$

In order to conclude our discussion, we need to prove that the minimax separation rate \bar{r}_ε is of the following order

$$r_{\varepsilon,0} = \mathcal{O}_\varepsilon \left(\varepsilon^{\frac{2s}{2s+2t+1/2}} \right). \quad (4.15)$$

To this end, we remark that, for any $r_\varepsilon > 0$,

- If $r_\varepsilon/r_{\varepsilon,0} \rightarrow 0$ then, using (4.13), it is easily seen that $u_\varepsilon(r_\varepsilon) = o_\varepsilon(1)$. Hence, according to Theorem 4.2, $\beta_\varepsilon(\Theta_a(r_\varepsilon)) = 1 - \alpha + o_\varepsilon(1)$.
- If $r_\varepsilon/r_{\varepsilon,0} \rightarrow +\infty$ then, using (4.13), it is easily seen that $u_\varepsilon(r_\varepsilon) \rightarrow +\infty$, as $\varepsilon \rightarrow 0$. Hence, according to Theorem 4.2, $\beta_\varepsilon(\Theta_a(r_\varepsilon)) = o_\varepsilon(1)$.

Therefore, Definition 2.7 allows to conclude that $r_{\varepsilon,0}$ in (4.15) is indeed the minimax separation rate \bar{r}_ε . (Note that, in view of (4.7) and (4.15), the minimax separation radius \tilde{r}_ε and the minimax separation rate \bar{r}_ε are of the same asymptotic order, as expected according to previous discussion.)

4.3.3 Non-asymptotic analysis of the Ingster test $\Psi_{r_\varepsilon, I}$

We present a non-asymptotic study of the Ingster test $\Psi_{r_\varepsilon, I}$, defined in (3.23). We show that the statistical performances of the Ingster test $\Psi_{r_\varepsilon, I}$ and the spectral cut-off test $\Psi_{D, P}$ defined in (3.15) are comparable. In particular, the Ingster test $\Psi_{r_\varepsilon, I}$ appears to be powerful in the sense of Definition 2.5, namely,

$$r_\varepsilon(\mathcal{E}_a, \Psi_{r_\varepsilon, I}, \beta) \leq \mathcal{C}\tilde{r}_\varepsilon,$$

for a fixed $\beta \in]0, 1[$ and some constant $\mathcal{C} \geq 1$, for an appropriately selected radius $r_\varepsilon > 0$.

Proposition 4.2 *Let $\alpha, \beta \in]0, 1[$ be given. Define*

$$\rho_\varepsilon^2 := \inf_{\mathcal{R} \geq 2} [\mathcal{C}_0 \mathcal{R}^{-2s} \vee c'(\alpha, \beta) \varepsilon^2 \mathcal{R}^{2t+1/2}], \quad (4.16)$$

for some positive constants \mathcal{C}_0 and $c'(\alpha, \beta)$ that can be explicitly computed. Let $\Psi_I^* := \Psi_{\rho_\varepsilon, I}$, where $\Psi_{\cdot, I}$ is the Ingster test defined in (3.23), and let $\rho_\varepsilon > 0$ be the radius defined in (4.16). Then, there exists constants $\mathcal{C} \geq 1$ and $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, the separation radius of Ψ_I^* satisfies

$$r_\varepsilon(\mathcal{E}_a, \Psi_I^*, \beta) \leq \mathcal{C}\tilde{r}_\varepsilon. \quad (4.17)$$

PROOF OF PROPOSITION 4.2. Let $r_\varepsilon > 0$ be a given radius. Using the same arguments as in (3.30), we get

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon, I}) \leq \sup_{\theta \in \Theta_a(r_\varepsilon)} \frac{1 + 4\omega_{0, r_\varepsilon} h(r_\varepsilon, \theta)}{(h(r_\varepsilon, \theta) - t_{1-\alpha})^2}.$$

Then, there exists an explicit constant $C_{\alpha, \beta} > 0$ such that

$$\begin{aligned} u_\varepsilon(r_\varepsilon) \geq C_{\alpha, \beta} &\Rightarrow h(r_\varepsilon, \theta) \geq C_{\alpha, \beta}, \quad \forall \theta \in \Theta_a(r_\varepsilon) \\ &\Rightarrow \frac{1 + 4\omega_{0, r_\varepsilon} h(r_\varepsilon, \theta)}{(h(r_\varepsilon, \theta) - t_{1-\alpha})^2} \leq \beta \quad \forall \theta \in \Theta_a(r_\varepsilon) \\ &\Rightarrow \sup_{\theta \in \Theta_a(r_\varepsilon)} \frac{1 + 4\omega_{0, r_\varepsilon} h(r_\varepsilon, \theta)}{(h(r_\varepsilon, \theta) - t_{1-\alpha})^2} \leq \beta \\ &\Rightarrow \beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon, I}) \leq \beta. \end{aligned} \quad (4.18)$$

Our task now is to find a condition on $r_\varepsilon > 0$ that will guarantee the validity of the above inequality $u_\varepsilon(r_\varepsilon) \geq C_{\alpha, \beta}$. Working along the lines of Section 4.3.2, we then arrive at

$$r_\varepsilon^2 = z_0^2 J_1, \quad \mathcal{A} = z_0^2 J_2 \quad \text{and} \quad u_\varepsilon^2(r_\varepsilon) = \varepsilon^{-4} r_\varepsilon^4 \frac{J_0}{2J_1^2}.$$

Hence, we see that

$$u_\varepsilon(r_\varepsilon) \geq C_{\alpha, \beta} \quad \text{iff} \quad r_\varepsilon^2 \geq \sqrt{2} C_{\alpha, \beta} \varepsilon^2 \frac{J_1}{J_0^{1/2}}. \quad (4.19)$$

Moreover,

$$J_1 = \sum_{j \in \mathbb{N}} b_j^{-4} (1 - Aa_j^2)_+ \leq \sum_{j: j \leq \mathcal{R}} j^{4t} \leq C'_1 \mathcal{R}^{4t+1},$$

and

$$J_1 \geq \sum_{j: j \leq \mathcal{R}} j^{4t} - \mathcal{R}^{-2s} \sum_{j: j \leq \mathcal{R}} j^{4t+2s} \geq C_1 \mathcal{R}^{4t+1},$$

for all $\mathcal{R} \geq 2$, where $\mathcal{R} = \mathcal{A}^{-1/2s}$, and for some positive constants C_1, C'_1 (depending on s and t only). In the same spirit, we can also prove that, for all $\mathcal{R} \geq 2$,

$$C_2 \mathcal{R}^{4t+1} \leq J_2 \leq C'_2 \mathcal{R}^{4t+1} \quad \text{and} \quad C_0 \mathcal{R}^{4t+1} \leq J_0 \leq C'_0 \mathcal{R}^{4t+1},$$

for some positive constants C_0, C'_0, C_1, C'_1 (depending on s and t only). Hence, we get that

$$\sqrt{2} C_{\alpha, \beta} \varepsilon^2 \frac{J_1}{J_0^{1/2}} \leq c'(\alpha, \beta) \varepsilon^2 \mathcal{R}^{2t+1/2}, \quad (4.20)$$

for some constant $c'(\alpha, \beta) > 0$. Therefore, we deduce from (4.19)-(4.20), that

$$r_\varepsilon^2 \geq c'(\alpha, \beta) \varepsilon^2 \mathcal{R}^{2t+1/2} \quad \Rightarrow \quad u_\varepsilon(r_\varepsilon) \geq C_{\alpha, \beta}.$$

Using the same kind of algebra, we get from (4.12) that $r_\varepsilon^2 = \mathcal{R}^{-2s} J_1 / J_0$. Hence, for all $\mathcal{R} \geq 2$,

$$C_0 \mathcal{R}^{-2s} \leq r_\varepsilon^2 \leq C_1 \mathcal{R}^{-2s},$$

for some positive constants C_0, C_1 . Finally, for all $\mathcal{R} \geq 2$,

$$r_\varepsilon^2 \geq C_0 \mathcal{R}^{-2s} \vee c'(\alpha, \beta) \varepsilon^2 \mathcal{R}^{2t+1/2} \quad \Rightarrow \quad u_\varepsilon(r_\varepsilon) \geq C_{\alpha, \beta} \quad \Rightarrow \quad \beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{r_\varepsilon, I}) \leq \beta.$$

Hence, taking $r_\varepsilon := \rho_\varepsilon$, where ρ_ε is defined in (4.16), we immediately get that

$$\beta_\varepsilon(\Theta_a(\rho_\varepsilon), \Psi_I^*) \leq \beta,$$

which implies

$$r_\varepsilon(\mathcal{E}_a, \Psi_I^*, \beta) \leq \rho_\varepsilon.$$

To conclude, it suffices to show that there exists a constant $\mathcal{C} \geq 1$ and $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, $\rho_\varepsilon \leq \mathcal{C} \tilde{r}_\varepsilon$. This, however, holds true working along the lines of the proof of (4.5). This concludes the proof of the proposition. \square

Concerning Proposition 4.2, the following comments are in order:

- The considered Ingster test Ψ_I^* , designed for asymptotic purposes, can be, somehow, employed in the non-asymptotic framework. It appears, that we recover existing non-asymptotic upper bounds, namely, for all $0 < \varepsilon < \varepsilon_0$, the Ingster test Ψ_I^* is powerful according to Definition 2.5. The value $\varepsilon_0 > 0$ guarantees that the optimal bandwidth $D_0 \in \mathbb{N}$ in (4.4) satisfies the requirement $2 \leq D_0 < +\infty$ which, in turn, ensures that ρ_ε and \tilde{r}_ε are, indeed, of the same order.
- The term ρ_ε involved in the construction of $\Psi_I^* := \Psi_{I, \rho_\varepsilon}$ plays the role of a tuning (regularization) parameter. In a sense, the parameter ρ_ε plays a similar role to the bandwidth D^* in (3.21). Hence, it provides a trade-off between the two competing terms ‘bias’ and ‘standard deviation’ involved in (4.16).
- As we have seen in (3.25), the Ingster test $\Psi_{r_\varepsilon, I}$ is an asymptotic α -level test for all $r_\varepsilon > 0$. Hence, a non-asymptotic control of the first kind error probability $\alpha_\varepsilon(\Psi_I^*)$ would be necessary in order to provide a fully non-asymptotic treatment for the Ingster test Ψ_I^* . This can be easily accomplished by replacing the $(1 - \alpha)$ -quantile $t_{1-\alpha}$ of a standard Gaussian random variable in (3.23) by an appropriate $(1 - \alpha)$ -quantile. Then, the upper bound (4.17) presented in Proposition 4.2 still holds true, up to some constants.

4.3.4 Asymptotics of Gaussian type for the spectral cut-off test $\Psi_{D,P}$

To conclude, we present an asymptotic analysis of the spectral cut-off test $\Psi_{D,P}$, defined in (3.15). In particular, as $\epsilon \rightarrow 0$, we prove that the maximal second kind error probability $\beta_\epsilon(\Theta_a(r_\epsilon), \Psi_{\tilde{D},P})$, for an appropriate bandwidth $\tilde{D} \in \mathbb{N}$, is asymptotically bounded from above by a quantity that possesses a Gaussian shape.

Proposition 4.3 *Let $\alpha \in]0, 1[$ be given. Let $r_\epsilon > 0$ be a radius satisfying $u_\epsilon(r_\epsilon) = \mathcal{O}_\epsilon(1)$. Let also $\Psi_{\tilde{D},P}$ be the spectral cut-off test defined in (3.15) with bandwidth $\tilde{D} \in \mathbb{N}$ satisfying*

$$\tilde{D} := \arg \max_{D \in \mathbb{N}} \left\{ C(\alpha) \epsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4}} + a_D^{-2} \leq \frac{r_\epsilon^2}{2} \right\}, \quad (4.21)$$

where the positive constant $C(\alpha)$ is defined in Proposition 3.1. Then, for any sequence $h_\epsilon \in]0, 1[$ satisfying $h_\epsilon = o_\epsilon(1)$ and $h_\epsilon \tilde{D}^{1/4} \rightarrow +\infty$ as $\epsilon \rightarrow 0$,

$$\beta_\epsilon(\Theta_a(r_\epsilon), \Psi_{\tilde{D},P}) \leq \Phi \left(t_{1-\alpha} - (1 - h_\epsilon) \frac{a_{\tilde{D}}^{-2}}{\epsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \right) + o_\epsilon(1). \quad (4.22)$$

PROOF OF PROPOSITION 4.3 Consider the spectral cut-off test $\Psi_{\tilde{D},P}$ defined in (3.15), with bandwidth $\tilde{D} \in \mathbb{N}$ selected as in (4.21). For any $\theta \in \Theta_a(r_\epsilon)$ and any sequence $h_\epsilon \in]0, 1[$ (that will be made precise later on)

$$\begin{aligned} \mathbb{P}_\theta(\Psi_{\tilde{D},P} = 0) &= \mathbb{P}_\theta \left(T_{\tilde{D},P} \leq t_{1-\alpha, \tilde{D}} \right) \\ &= \mathbb{P}_\theta \left(\sum_{j=1}^{\tilde{D}} b_j^{-2} (y_j^2 - \epsilon^2) \leq t_{1-\alpha, \tilde{D}} \right) \\ &= \mathbb{P}_\theta \left(\epsilon^2 \sum_{j=1}^{\tilde{D}} b_j^{-2} (\xi_j^2 - 1) + 2\epsilon \sum_{j=1}^{\tilde{D}} b_j^{-1} \theta_j \xi_j \leq t_{1-\alpha, \tilde{D}} - \sum_{j=1}^{\tilde{D}} \theta_j^2 \right) \\ &\leq \mathbb{P}_\theta \left(\epsilon^2 \sum_{j=1}^{\tilde{D}} b_j^{-2} (\xi_j^2 - 1) \leq (1 - h_\epsilon) \left(t_{1-\alpha, \tilde{D}} - \sum_{j=1}^{\tilde{D}} \theta_j^2 \right) \right) \\ &\quad + \mathbb{P}_\theta \left(2\epsilon \sum_{j=1}^{\tilde{D}} b_j^{-1} \theta_j \xi_j \leq h_\epsilon \left(t_{1-\alpha, \tilde{D}} - \sum_{j=1}^{\tilde{D}} \theta_j^2 \right) \right) \\ &:= T_1 + T_2, \end{aligned} \quad (4.23)$$

where, for the last inequality, we used the fact that, for any $t \in \mathbb{R}$ and any random variables X and Y ,

$$\{X + Y \leq t\} \subseteq \{X \leq (1 - h_\epsilon)t\} \cup \{Y \leq h_\epsilon t\}.$$

Below, our aim is

- to show that, asymptotically, T_1 has a Gaussian shape of the form (4.33),
- to prove that $T_2 = o_\epsilon(1)$,
- to study the asymptotic behavior of the threshold $t_{1-\alpha, \tilde{D}}$.

Control of T_1 : For any $\delta > 0$, simple algebra shows that, for any bandwidth $D \in \mathbb{N}$,

$$\frac{\varepsilon^{2(2+\delta)} \sum_{j=1}^D b_j^{-2(2+\delta)} \mathbb{E}|\xi_j^2 - 1|^{2+\delta}}{\left(\varepsilon^2 \sqrt{\sum_{j=1}^D b_j^{-4}}\right)^{2+\delta}} \asymp \frac{D^{2(2+\delta)t+1}}{D^{(4t+1)(2+\delta)/2}} \asymp D^{-\delta/2} = o(1) \quad \text{as } D \rightarrow +\infty.$$

Hence, by Lyapunov's condition,

$$\frac{\varepsilon^2 \sum_{j=1}^{\tilde{D}} b_j^{-2} (\xi_j^2 - 1)}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \quad \text{as } \varepsilon \rightarrow 0 \quad (\text{since } \tilde{D} \rightarrow +\infty).$$

Then, it follows that

$$T_1 = \Phi \left(\frac{(1 - h_\varepsilon) \left(t_{1-\alpha, \tilde{D}} - \sum_{j=1}^{\tilde{D}} \theta_j^2 \right)}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \right) + o_\varepsilon(1). \quad (4.24)$$

Control of T_2 : Since $(\xi_j)_{j \in \mathbb{N}}$ are independent standard Gaussian random variables,

$$\begin{aligned} T_2 &:= \mathbb{P}_\theta \left(2\varepsilon \sum_{j=1}^{\tilde{D}} b_j^{-1} \theta_j \xi_j \leq h_\varepsilon \left(t_{1-\alpha, \tilde{D}} - \sum_{j=1}^{\tilde{D}} \theta_j^2 \right) \right) \\ &= \mathbb{P}_\theta \left(Z \leq \frac{h_\varepsilon \left(t_{1-\alpha, \tilde{D}} - \sum_{j=1}^{\tilde{D}} \theta_j^2 \right)}{\varepsilon \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-2} \theta_j^2}} \right) \quad (\text{where } Z \sim \mathcal{N}(0, 1)). \end{aligned} \quad (4.25)$$

Then, according to (4.21), for any $\theta \in \Theta_a(r_\varepsilon)$,

$$\begin{aligned} \frac{\sum_{j=1}^{\tilde{D}} \theta_j^2 - t_{1-\alpha, \tilde{D}}}{\varepsilon \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-2} \theta_j^2}} &\geq \frac{\|\theta\|^2 - t_{1-\alpha, \tilde{D}} - a_{\tilde{D}}^{-2}}{\varepsilon (\max_{1 \leq j \leq \tilde{D}} b_j^{-1}) \|\theta\|} \geq \frac{1}{C} \left(\frac{\|\theta\|^2 - r_\varepsilon^2/2}{r_\varepsilon \|\theta\|} \right) \tilde{D}^{1/4} \\ &\geq \frac{\tilde{D}^{1/4} \|\theta\|}{2C r_\varepsilon} \\ &\geq \frac{1}{2C} \tilde{D}^{1/4} \rightarrow +\infty \quad \text{as } \varepsilon \rightarrow 0 \quad (\text{since } \tilde{D} \rightarrow +\infty), \end{aligned} \quad (4.26)$$

where for the second inequality we used the fact that

$$\varepsilon \left(\max_{1 \leq j \leq \tilde{D}} b_j^{-1} \right) \asymp \varepsilon \tilde{D}^t \asymp \frac{\left(\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}} \right)^{1/2}}{\tilde{D}^{1/4}} \leq C \frac{r_\varepsilon}{\tilde{D}^{1/4}},$$

for some constant $C > 0$. Hence, using (4.25) and (4.26), it follows that, as soon as $h_\varepsilon \tilde{D}^{1/4} \rightarrow +\infty$ as $\varepsilon \rightarrow 0$,

$$T_2 = o_\varepsilon(1). \quad (4.27)$$

Behavior of $t_{1-\alpha, \tilde{D}}$: First we show that

$$(1 - h_\varepsilon) \frac{t_{1-\alpha, \tilde{D}}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} = t_{1-\alpha} + o_\varepsilon(1). \quad (4.28)$$

Indeed, according to the definition of $t_{1-\alpha, \tilde{D}}$,

$$\begin{aligned} & \mathbb{P}_0 \left(\sum_{j=1}^{\tilde{D}} b_j^{-2} (y_j^2 - \varepsilon^2) \leq t_{1-\alpha, \tilde{D}} \right) = 1 - \alpha \\ \Leftrightarrow & \mathbb{P} \left(\frac{\varepsilon^2 \sum_{j=1}^{\tilde{D}} b_j^{-2} (\xi_j^2 - 1)}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \leq \frac{t_{1-\alpha, \tilde{D}}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \right) = 1 - \alpha \\ \Leftrightarrow & \Phi_\varepsilon^{-1}(1 - \alpha) = \frac{t_{1-\alpha, \tilde{D}}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}}, \end{aligned}$$

where, for any $s \in \mathbb{R}$,

$$\Phi_\varepsilon(s) := \mathbb{P} \left(\frac{\varepsilon^2 \sum_{j=1}^{\tilde{D}} b_j^{-2} (\xi_j^2 - 1)}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \leq s \right).$$

Then, as above, using the Central Limit Theorem with Lyapunov's condition and Lemma 21.2 in [26], we get

$$\begin{aligned} & \Phi_\varepsilon(s) \rightarrow \Phi(s) \quad \text{as } \varepsilon \rightarrow 0 \quad (\forall s \in \mathbb{R}) \\ \Leftrightarrow & \Phi_\varepsilon^{-1}(u) \rightarrow \Phi^{-1}(u) \quad \text{as } \varepsilon \rightarrow 0 \quad (\forall u \in]0, 1[) \end{aligned} \quad (4.29)$$

In particular, for any $\alpha \in]0, 1[$,

$$\begin{aligned} & \Phi_\varepsilon^{-1}(1 - \alpha) \rightarrow \Phi^{-1}(1 - \alpha) \quad \text{as } \varepsilon \rightarrow 0 \\ \Leftrightarrow & \frac{t_{1-\alpha, \tilde{D}}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} = t_{1-\alpha} + o_\varepsilon(1). \end{aligned} \quad (4.30)$$

Finally, taking into account that $h_\varepsilon = o_\varepsilon(1)$, (4.28) holds true.

Completing the proof: Using (4.24) and (4.28), it follows that

$$T_1 = \Phi \left(t_{1-\alpha} - \frac{(1 - h_\varepsilon) \sum_{j=1}^{\tilde{D}} \theta_j^2}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} + o_\varepsilon(1) \right) + o_\varepsilon(1). \quad (4.31)$$

According to (4.21), for any $\theta \in \Theta_a(r_\varepsilon)$,

$$\begin{aligned} (1 - h_\varepsilon) \frac{\sum_{j=1}^{\tilde{D}} \theta_j^2}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} &= (1 - h_\varepsilon) \frac{\|\theta\|^2 - \sum_{j=1}^{\tilde{D}} \theta_j^2}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \\ &\geq (1 - h_\varepsilon) \frac{r_\varepsilon^2 - a_{\tilde{D}}^{-2}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \\ &\geq (1 - h_\varepsilon) \frac{a_{\tilde{D}}^{-2}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}}. \end{aligned} \quad (4.32)$$

Hence, using (4.23), (4.24), (4.27) and the Mean Value Theorem,

$$\begin{aligned} \beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D}, P}) &= \sup_{\theta \in \Theta_a(r_\varepsilon)} \mathbb{P}_\theta(\Psi_{\tilde{D}, P} = 0) \\ &\leq \Phi \left(t_{1-\alpha} - (1 - h_\varepsilon) \frac{a_{\tilde{D}}^{-2}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \right) + o_\varepsilon(1). \end{aligned}$$

Hence, (4.22) holds true, and this completes the proof of the proposition.

□

Concerning Proposition 4.3, the following comments are in order:

- The maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D}, P})$ associated to the spectral cut-off test $\Psi_{\tilde{D}, P}$, with bandwidth $\tilde{D} \in \mathbb{N}$ selected as in (4.21), is asymptotically bounded from above by a quantity that possesses a Gaussian shape. It is worth mentioning that this spectral cut-off test $\Psi_{\tilde{D}, P}$ is of the same type as the one introduced in Section 3.2.2. In particular, by construction, the spectral cut-off $\Psi_{\tilde{D}, P}$ is still an α -level test. Nevertheless, the bandwidth $\tilde{D} \in \mathbb{N}$ defined in (4.21), is selected in a different manner in order to accommodate the asymptotic paradigm. Indeed, this regularization parameter $\tilde{D} \in \mathbb{N}$ now depends on the radius r_ε . Notice that this is comparable to the construction of the Ingster test $\Psi_{r_\varepsilon, I}$ introduced in (3.23), where the Ingster filters $\omega_{j, r_\varepsilon}$ defined in (3.22) explicitly depend on the radius r_ε .
- The asymptotic upper bound of the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D}, P})$ obtained in (4.22) is coherent with the non-asymptotic analysis provided in Section 3.2.2 (see, in particular, (3.20) and (3.21)). Indeed, in order to guarantee that, for any $\beta \in]0, 1[$, $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D}, P})$ is (asymptotically) upper bounded by β , we have to solve the equation $a_D^{-2} \asymp C_{\alpha, \beta} \varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}$, for some constant $C_{\alpha, \beta} > 0$ (whose value depends on the tools used to control $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D}, P})$).
- In order to conclude our discussion, we provide a heuristic comparison between the asymptotic upper bound of the maximal second kind error probability $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D}, P})$ and the sharp asymptotics of Gaussian type obtained in Theorem 4.2. Working as in Section 4.3.2, we get that, as $\varepsilon \rightarrow 0$,

$$u_\varepsilon(r_\varepsilon) \sim \frac{r_\varepsilon^4}{\varepsilon^4} \mathcal{R}_\varepsilon^{-(4t+1)} \sim \left(\frac{\mathcal{R}_\varepsilon^{-2s}}{\varepsilon^2 \mathcal{R}_\varepsilon^{2t+1/2}} \right)^2, \quad \text{where } \mathcal{R}_\varepsilon \text{ satisfies } \mathcal{R}_\varepsilon^{-s} \sim r_\varepsilon.$$

Note that, thanks to (4.13), $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$ implies that $r_\varepsilon \sim \varepsilon^{2s/(2s+2t+1/2)}$, as $\varepsilon \rightarrow 0$. Moreover, in view of (4.7), $\tilde{D}^{-2s} \sim a_D^{-2} \sim r_\varepsilon^2$ as $\varepsilon \rightarrow 0$. Hence, according to the definition of the bandwidth \tilde{D} given in (4.21), as soon as $u_\varepsilon(r_\varepsilon) = \mathcal{O}_\varepsilon(1)$, in some sense, we have that

$$u_\varepsilon(r_\varepsilon) \sim \frac{a_{\tilde{D}}^{-2}}{\varepsilon^2 \sqrt{\sum_{j=1}^{\tilde{D}} b_j^{-4}}} \quad \text{as } \varepsilon \rightarrow 0.$$

In particular, it means that we can find a $c \in]0, 1[$ such that

$$\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi_{\tilde{D}, P}) \leq \Phi(t_{1-\alpha} - c u_\varepsilon(r_\varepsilon)) + o_\varepsilon(1). \quad (4.33)$$

According to Theorem 4.2, it is immediately seen that

$$\beta_{\varepsilon, \alpha}(\Theta_a(r_\varepsilon)) = \Phi(t_{1-\alpha} - u_\varepsilon(r_\varepsilon)) + o_\varepsilon(1) < \Phi(t_{1-\alpha} - c u_\varepsilon(r_\varepsilon)) + o_\varepsilon(1). \quad (4.34)$$

Hence, the spectral cut-off test $\Psi_{\tilde{D}, P}$ defined in (4.22), with bandwidth $\tilde{D} \in \mathbb{N}$ selected as in (4.21), does not provide sharp asymptotics of Gaussian type. Indeed, it is not designed for that purpose: the spectral cut-off filters associated to this test appear to be quite ‘rough’ in such setting compared to the Ingster filters defined in (3.22) (see Figure 4 for a graphical illustration).

- If we define a radius $\bar{r}_\varepsilon^* > 0$ to satisfy $u(\bar{r}_\varepsilon^*) = t_{1-\alpha} - t_\beta$, for prescribed $\alpha, \beta \in]0, 1[$, then, using Theorem 4.2, we immediately get

$$\beta_{\varepsilon, \alpha}(\Theta_a(\bar{r}_\varepsilon^*)) = \Phi(t_{1-\alpha} - u_\varepsilon(\bar{r}_\varepsilon^*)) + o_\varepsilon(1) = \beta + o_\varepsilon(1).$$

Furthermore, according to the definition of the separation radius $\tilde{r}_{\varepsilon, \tilde{D}}^* := r_\varepsilon(\mathcal{E}_a, \Psi_{\tilde{D}, P}, \beta)$ for the spectral cut-off test $\Psi_{\tilde{D}, P}$ defined in (4.22), with bandwidth $\tilde{D} \in \mathbb{N}$ selected as in (4.21), we have

$$\beta_{\varepsilon, \alpha}(\Theta_a(\tilde{r}_{\varepsilon, \tilde{D}}^*)) \leq \beta.$$

However, we conjecture that it is not possible to prove that

$$\frac{\tilde{r}_{\varepsilon, \tilde{D}}^*}{\tilde{r}_\varepsilon^*} = 1 + o_\varepsilon(1).$$

In other words, the spectral cut-off tests appear to be quite ‘rough’ in order to provide the optimal constants of the associated rates for the considered minimax signal detection problem.

Remark 4.6 Proposition 4.3 holds true in a general setting. Indeed, by looking at its proof (the control of T_1 and T_2), the only condition needed to prove (4.22) is that

$$\exists \delta > 0 \quad \text{such that} \quad \frac{\max_{1 \leq j \leq D} b_j^{-2}}{\sqrt{\sum_{j=1}^D b_j^{-4}}} = o(D^{-\delta}) \quad \text{as} \quad D \rightarrow +\infty. \quad (4.35)$$

It is easily seen the condition (4.35) is satisfied in various settings, namely, direct problems (i.e., $b_j = 1$, $j \in \mathbb{N}$), well-posed inverse problems (i.e., $b_j > b_0$, for some $b_0 > 0$, $j \in \mathbb{N}$) and mildly ill-posed problems (i.e., $b_j \asymp j^{-jt}$, $j \in \mathbb{N}$, for some $t > 0$). We point out, however, that it is not satisfied, for instance, in exponential or power-exponential behaviors (i.e., $b_j \asymp e^{-j^r t}$, $j \in \mathbb{N}$, for some $t > 0$ and $r \geq 1$), discussed in Remark 4.3). It is worth mentioning that condition (4.35) is, in general, comparable to the condition $\omega_{0, r_\varepsilon} = o_\varepsilon(1)$, discussed in Theorem 4.2. For more details on the asymptotic expression of $\omega_{0, r_\varepsilon}$ in mildly ill-posed inverse problems, we refer to the proof of Theorem 4.2 of [17], supplementary material, Section 11.3.

5 Conclusions

We discussed non-asymptotic and asymptotic approaches to minimax signal detection through a unified treatment and provided, in some sense, an overview of this specialized area. In particular, we considered a Gaussian sequence model that contains classical models as special cases, such as, direct, well-posed inverse and ill-posed inverse problems. We compared the construction of lower and upper bounds for the minimax separation radius (non-asymptotic approach) and the minimax separation rate (asymptotic approach), and brought into light hitherto unknown similarities and links between these two associated minimax signal detection paradigms. An example of a mildly ill-posed inverse problem was used for illustrative purposes. In particular, tools used to derive ‘asymptotic’ results can be exploited to draw ‘non-asymptotic’ conclusions, and vice-versa. To this end, we note that in these considerations we have worked with certain ellipsoids in the space of squared-summable sequences of real numbers, with a ball of positive radius removed, and we confined our attention to the Neyman-Pearson criterion.

There are various ways that the above results could be possibly extended. For instance, for the same smoothness classes, similar investigations, could be easily obtained for the total-error probability criterion defined as the sum of the type I and maximal type II error probabilities of a given test Ψ , i.e.,

$$\zeta_\varepsilon(\Theta_a(r_\varepsilon), \Psi) = \alpha_\varepsilon(\Psi) + \beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi),$$

where $\alpha_\varepsilon(\Psi)$ and $\beta_\varepsilon(\Theta_a(r_\varepsilon), \Psi)$ are defined in (2.5) and (2.6), respectively. Note that, by defining

$$\zeta_\varepsilon(\Theta_a(r_\varepsilon)) = \inf_{\tilde{\Psi}} [\zeta_\varepsilon(\Theta_a(r_\varepsilon), \tilde{\Psi})]$$

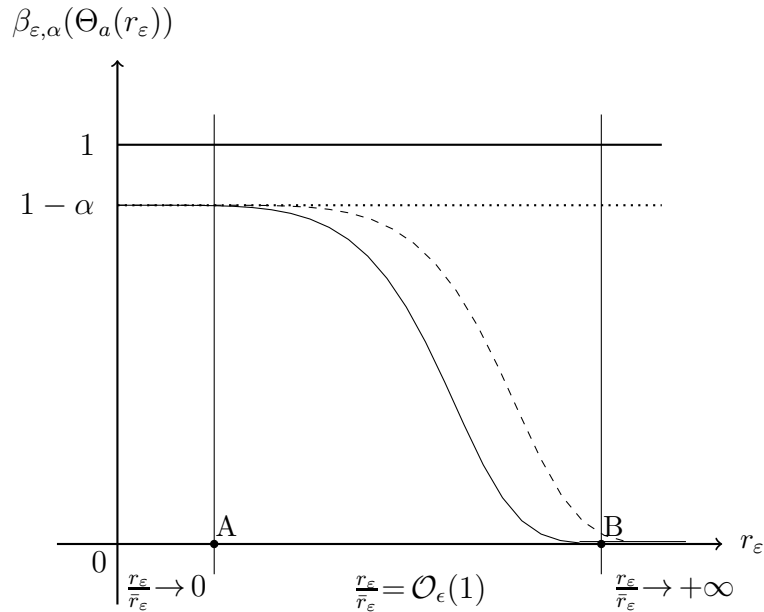


Figure 4: The solid curve represents the function $\Phi(t_{1-\alpha} - u_\epsilon(r_\epsilon))$ while the dashed curve displays the function $\Phi(t_{1-\alpha} - cu_\epsilon(r_\epsilon))$, for some $c \in]0, 1[$ and a radius $r_\epsilon > 0$ satisfying $u_\epsilon(r_\epsilon) = \mathcal{O}_\epsilon(1)$ (see (4.34)). The solid curve is associated to the sharp asymptotics of Gaussian type for the maximal second kind error probability $\beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon))$ while the dashed curve is associated to the asymptotic upper bound of the maximal second kind error probability $\beta_\epsilon(\Theta_a(r_\epsilon), \Psi_{\tilde{D}, P})$ of the spectral cut-off test $\Psi_{\tilde{D}, P}$, with bandwidth $\tilde{D} \in \mathbb{N}$ selected as in (4.21).

where the infimum is taken over all possible tests $\tilde{\Psi}$, it is known that (see, e.g., [18], Chapter 2) that

$$\zeta_\epsilon(\Theta_a(r_\epsilon)) = \inf_{\alpha \in]0, 1[} [\alpha + \beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon))],$$

where $\beta_{\epsilon, \alpha}(\Theta_a(r_\epsilon))$ is the minimax second kind error probability defined in Definition 2.6.

Similar investigations for the Neyman-Pearson criterion and/or the total-error probability criterion should also be possible for other classes \mathcal{F} of signals, such as those characterized by their non-zero coefficients (dense or sparse signals) and l_p -bodies with $p \in]0, 2]$ (see, e.g., [24], [1], [17], [22]). In the same spirit, several contributions have been proposed in various regression and density models which provide attractive frameworks for investigation in the minimax testing theory (see, e.g., [12], [9], [11], [4], [16], [5], [20]).

For the sake of brevity and clarity in our presentation, we have also not discussed adaptation issues of the involved testing procedures in the considered minimax signal detection paradigms. Indeed, the filters used to design the spectral cut-off (non-asymptotic framework) and Ingster (asymptotic framework) tests explicitly depend on the form of the sequence $(a_j)_{j \in \mathbb{N}}$ that measures the smoothness of the signal θ , which is, in general, unknown in practice. It is therefore of paramount importance in practical applications to provide minimax testing procedures that do not explicitly depend on the associated smoothness parameter. This is, usually, referred to as the ‘adaptation’ problem (see, e.g., [1], [18], [17], [23]).

However, all the above investigations need careful attention that is beyond the scope of the present work.

References

- [1] Y. Baraud. Non-asymptotic minimax rates of testing in signal detection. *Bernoulli*, 8(5):577–606, 2002.
- [2] Y. Baraud, S. Huet, and B. Laurent. A new test of linear hypothesis in regression. In *Goodness-of-Fit Tests and Model Validity (Paris, 2000)*, Stat. Ind. Technol., pages 195–207. Birkhäuser Boston, Boston, MA, 2002.
- [3] N. Bissantz, T. Hohage, A. Munk, and F. Ruymgaart. Convergence rates of general regularization methods for statistical inverse problems and applications. *SIAM J. Numer. Anal.*, 45(6):2610–2636, 2007.
- [4] C. Butucea. Goodness-of-fit testing and quadratic functional estimation from indirect observations. *Ann. Statist.*, 35(5):1907–1930, 2007.
- [5] C. Butucea, C. Matias, and C. Pouet. Adaptive goodness-of-fit testing from indirect observations. *Annales de l’I.H.P., Probabilités et Statistiques*, 45(2):352–372, 2009.
- [6] L. Cavalier. Inverse problems in statistics. In *Inverse Problems and High-Dimensional Estimation*, volume 203 of *Lect. Notes Stat. Proc.*, pages 3–96. Springer, Heidelberg, 2011.
- [7] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
- [8] M.S. Ermakov. Minimax detection of signal in gaussian white noise. *Th. Probab. Appl.*, 35(4):667–679, 1991.
- [9] M.S. Ermakov. Minimax nonparametric testing of hypotheses on a distribution density. *Th. Probab. Appl.*, 39:396–416, 1994.
- [10] J. Fan and I. Gijbels. *Local Polynomial Modelling and its Applications*, volume 66 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1996.
- [11] M. Fromont and B. Laurent. Adaptive goodness-of-fit tests in a density model. *Ann. Statist.*, 34(2):1–45, 2006.
- [12] E. Guerre and P. Lavergne. Optimal minimax rates for nonparametric specification testing in regression models. *Econometric Theory*, 18:1139–1171, 2002.
- [13] Yu.I. Ingster. Asymptotically minimax hypothesis testing for nonparametric alternatives. I. *Math. Methods Statist.*, 2(2):85–114, 1993.
- [14] Yu.I. Ingster. Asymptotically minimax hypothesis testing for nonparametric alternatives. II. *Math. Methods Statist.*, 2(3):171–189, 1993.
- [15] Yu.I. Ingster. Asymptotically minimax hypothesis testing for nonparametric alternatives. III. *Math. Methods Statist.*, 2(4):249–268, 1993.
- [16] Yu.I. Ingster and T. Sapatinas. Minimax goodness-of-fit testing in multivariate nonparametric regression. *Math. Methods Statist.*, 18(3):241–269, 2009.
- [17] Yu.I. Ingster, T. Sapatinas, and I.A. Suslina. Minimax signal detection in ill-posed inverse problems. *Ann. Statist.*, 40:1524–1549, 2012.
- [18] Yu.I. Ingster and I.A. Suslina. *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*, volume 169 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2003.

- [19] I.M. Johnstone. *Gaussian Estimation: Sequence and Wavelet Models*. 2013. Unpublished Monograph.
- [20] C. Lacour and T.M. Pham Ngoc. Goodness-of-fit test for noisy directional data. *Bernoulli*, 20:2131–2168, 2014.
- [21] B. Laurent, J.-M. Loubes, and C. Marteau. Testing inverse problems: a direct or an indirect problem? *J. Statist. Plann. Inference*, 141(5):1849–1861, 2011.
- [22] B. Laurent, J.-M. Loubes, and C. Marteau. Non asymptotic minimax rates of testing in signal detection with heterogeneous variances. *Electron. J. Stat.*, 6:91–122, 2012.
- [23] C. Marteau and P. Mathé. General regularization schemes for signal detection in inverse problems. *Math. Methods Statist.*, 23:176–200, 2014.
- [24] V.G. Spokoiny. Adaptive hypothesis testing using wavelets. *Ann. Statist.*, 24:2477–2498, 1996.
- [25] A.B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, New York, 2009. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.
- [26] A.W. van der Vaart. *Asymptotics Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, New York, 1998.
- [27] G. Wahba. *Spline Models for Observational Data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.