Maxiset Point of View for Signal Detection in Inverse Problems

F. Autin1*, M. Clausel2**, J.-M. Freyermuth1***, and C. Marteau3****

1Aix-Marseille Univ., Inst. de Math. de Marseille, France
2Univ. de Lorraine, CNRS, IECL, Nancy, France
3Univ. Lyon, Univ. Claude Bernard Lyon 1 CNRS UMR 5208, Inst. Camille Jordan, Villeurbanne, France

Received December 20, 2018; in final form, August 17, 2019

Abstract—This paper extends the successful maxiset paradigm from function estimation to signal detection in inverse problems. In this context, the maxisets do not have the same shape compared to the classical estimation framework. Nevertheless, we introduce a robust version of these maxisets allowing to exhibit tail conditions on the signals of interest. Under this novel paradigm we are able to compare direct and indirect testing procedures.

Keywords: maxiset, inverse problems, signal detection.

AMS 2010 Subject Classification: 62G20.

DOI: 10.3103/S1066530719030037

1. INTRODUCTION

Over the last 20 years, the assessment of the performance of nonparametric function estimation methods mainly relied on the asymptotic minimax and oracle approaches. More marginally used, the maxiset paradigm has been proved to be very useful to accurately describe the behavior of some estimation procedures. In some cases, it allows one to distinguish methods having comparable minimax performance. The question of adapting the maxiset concept to the signal detection framework was often raised. This is the aim of this paper to rigorously extend this point of view to the signal detection framework and to discuss new related outcomes.

We will deal all along the paper with the Gaussian sequence space model

\[ y_k = b_k \theta_k + \varepsilon \xi_k, \quad k \in \mathbb{N}^* := \mathbb{N} \setminus \{0\}, \]

where \((y_k)_{k \in \mathbb{N}^*}\) denotes the observations, \(\theta = (\theta_k)_{k \in \mathbb{N}^*}\) is an unknown sequence of interest, \((b_k)_{k \in \mathbb{N}^*}\) is a given decreasing sequence of real numbers in \((0, 1]\) that tends to zero as \(k\) goes to infinity, \(\varepsilon\) is a noise level in \((0, 1)\) and \((\xi_k)_{k \in \mathbb{N}^*}\) is a sequence of i.i.d. standard Gaussian random variables. The model (1) allows us to describe several situations, as, e.g., nonparametric regression or estimation of a blurred function in white noise. For more details on these situations and their connection with the Gaussian sequence space model we refer the interested reader to [17].

We also stress that the model (1) allows one to deal with so-called inverse problem models as described in [6]. In such a setting, one is interested in doing inference on a function \(f\) in some Hilbert space \(H\) from indirect and blurred observation of the form

\[ Y = Af + \varepsilon \xi, \]

where \(A : H \rightarrow K\) denotes a compact operator acting from \(H\) to another Hilbert space \(K\), \(\varepsilon\) is a noise level and \(\xi\) is a Gaussian white noise. Let us denote by \(A^*\) the adjoint operator of \(A\) and remark that the two sequences \((b_k^2)_{k \in \mathbb{N}^*}\) and \((\theta_k)_{k \in \mathbb{N}^*}\) can be respectively identified as the sequence of eigenvalues of the

---

*E-mail: florent.autin@univ-amu.fr
**E-mail: marianne.clausel@univ-lorraine.fr
***E-mail: jean-marc.FREYERMUTH@univ-amu.fr
****E-mail: marteau@math.univ-lyon1.fr
operator \( A^*A \) and as the sequence of coefficients of \( f \) in the singular values decomposition (SVD) basis associated with the operator \( A \).

For inverse problem models, the minimax paradigm has been widely used in order to assess the performance of estimation procedures. Roughly speaking, given a structural constraint on the vector \( \theta \) of interest, typically of the form \( \theta \in \Theta \) for some \( \Theta \subset l_2(\mathbb{N}^*) \), one measures the performance of a given estimator \( \hat{\theta} \), that is, a measurable function of the observations, through its maximal risk

\[
R(\hat{\theta}) = \sup_{\theta \in \Theta} E_\theta \ell(\theta, \hat{\theta}),
\]

where \( \ell(\cdot, \cdot) \) denotes a given loss function. This paradigm has been widely used and discussed over the years. In several situations, a precise bound on \( R(\hat{\theta}) \) can be obtained, which allows one to characterize how the maximal risk decreases with respect to the noise level \( \varepsilon \). More precisely, one can often exhibit a continuous sequence of positive numbers \( (r_\varepsilon)_{\varepsilon > 0} \) that tends to zero when \( \varepsilon \to 0 \), the so-called rate of convergence associated with the estimation procedure \( \hat{\theta} \) with noise level \( \varepsilon \) such that \( R(\hat{\theta}) \sim r_\varepsilon \) for some positive constant \( C \). If this rate appears to be the smallest possible one, namely if there exists a positive constant \( c \) such that

\[
\inf_{\hat{\theta}} \sup_{\theta \in \Theta} E_\theta \ell(\theta, \hat{\theta}) \geq cr_\varepsilon,
\]

the sequence \( (r_\varepsilon)_{\varepsilon > 0} \) is called minimax rate of convergence over \( \Theta \) and \( \hat{\theta} \) is said to be a ‘minimax-optimal’ estimator over \( \Theta \). In the above inequality, the infimum is taken over all possible estimators \( \hat{\theta} \) of \( \theta \). We refer, e.g., to [17] and [11] for a non-exhaustive reference list.

Under the minimax estimation paradigm, the performance of two given procedures can be compared through their respective rates of convergence according to a chosen functional set \( \Theta \). However it does not always allow for comparison of performance if the two procedures are both minimax-optimal. In addition the used criterion is quite pessimistic: the minimax point of view focuses on the slowest possible estimation precision over the set \( \Theta \). To tackle these issues, an alternative point of view has been proposed in the seminal paper [12]. In the maxiset setting, given an estimation procedure \( \hat{\theta} \) and a sequence of rates \( (r_\varepsilon)_{\varepsilon > 0} \), providing the maxiset of \( \hat{\theta} \) for the rate \( (r_\varepsilon)_{\varepsilon > 0} \) means exhibiting the set \( \Theta_{MS}(\hat{\theta}) \) of sequences \( \theta \) that are estimated by \( \hat{\theta} \) at the rate \( (r_\varepsilon)_{\varepsilon > 0} \). Under this paradigm, the best performing procedure, i.e., the ‘maxiset-optimal’ procedure, would be the one whose associated maxiset includes the maxisets of the others. Note that a common criticism concerns the situation where estimation methods have nonnested maxisets. Autin et al. [3] discuss this important aspect of the maxiset approach. They explain that, firstly, examining the ‘form’ of the maxiset of an estimator brings interesting information on the signals well estimated by itself. Secondly, it may be possible to build a novel maxiset-optimal procedure by combining those with nonnested maxisets. The maxiset point of view has been generalized to various settings, see, e.g., [2], [16] or [8].

In the framework of signal detection, the minimax point of view has been widely investigated and was very fruitfully applied. We refer to [9], [4], [10] or [15] among others. Nevertheless, as in the estimation case, the minimax paradigm does not allow for a fully satisfying comparison between different testing procedures. The extension of the maxiset theory to this setting is a doorway to novel informative and rigorous mathematical study of the statistical performance of these procedures. A flavor of the maxiset approach in the signal detection framework has been discussed in [1]. Recently Ermakov [7] proposed a definition of maxisets for tests which is intrinsically different from the one presented hereafter. Moreover, we consider the more general setting of inverse problems and we exhibit the maxisets of testing procedures for a wide range of rates which allow us to properly compare the performance of estimators.

We discuss hereafter some new issues related to the maxiset setting. In particular, we aim at

- comparing two approaches for signal detection in model (2). Testing could be performed using either an indirect or a direct approach. The former requires the inversion of the operator \( A \), while the latter does not. Laurent et al. [14] compared these approaches from a minimax point of view and notably showed that, in the mildly ill-posed problem, direct approaches outperform indirect ones. In this paper we will complete their results by comparing these approaches from our newly developed maxiset point of view;
• describing two extreme regimes according to the separation rate and \((b_k)_{k \in \mathbb{N}^*}\). In the first regime, the maxiset of the testing procedures is the empty set, while in the second regime, it is not and it surprisingly does not depend on the inverse problem. Of particular interest is the transition regime also called hereafter horizon of detection, which is subtly depending on the decay of the \((b_k)_{k \in \mathbb{N}^*}\).

Novel insights on the performance of both direct and indirect approaches related to this horizon of detection will be provided.

In Section 2, we recall the minimax paradigm and present the maxiset point of view for signal detection problems. Thereafter, in Section 3, we state our maxiset results for both the indirect and direct approaches (see Theorems 3.1 and 3.2). In Section 4 we succeed in comparing these two approaches in the light of the maxiset theory in many cases (see Proposition 4.1) as, for instance, in the mildly ill-posed inverse problem (see Proposition 4.2). Following the conclusion on the novelty of our results in Section 5, we postpone all the related proofs to Section 6.

2. SIGNAL DETECTION IN INVERSE PROBLEMS

2.1. The Minimax Paradigm for Inverse Problems in Signal Detection

We consider the sequence space model (1)

\[ y_k = b_k \theta_k + \varepsilon_k, \quad k \in \mathbb{N}^*. \]

The signal detection problem aims at determining whether or not the observations \(y = (y_k)_{k \in \mathbb{N}^*}\) contain some signal. This question can be formalized by the following hypotheses testing problem:

\[ H_0: \theta = 0_{l_2(\mathbb{N}^*}) \quad \text{against} \quad H_1: \theta \in \Theta, \quad \|\theta\|^2 \geq r_\varepsilon^2. \]  

(3)

for some continuous sequence of positives numbers \(r = (r_\varepsilon)_{\varepsilon > 0}\) that tends to 0 as \(\varepsilon \to 0\). Here \(\| \cdot \|\) is the Euclidean norm.

In the alternative hypothesis \(H_1\), the set \(\Theta\) denotes a subset of \(l_2(\mathbb{N}^*)\). The requirement \(\theta \in \Theta\) can be thought of either as a structural constraint on the signal or as a regularity condition on the underlying function \(f\) in model (2). At the same time, the constraint \(\|\theta\|^2 \geq r_\varepsilon^2\) corresponds to an energy condition that allows us to quantify the amount of signal available in the observations. Another problem closely related to signal detection is pattern recognition. In this case, one aims at testing the goodness of fit between the observations and a given reference signal \(\theta^o\). Having in mind that these two problems are equivalent up to the change of variable \(\theta \leftarrow \theta - \theta^o\), we shall only focus on the sequel in the signal detection problem.

In the sequel we denote by \(\Delta = \Delta(y)\) a testing procedure. It is a measurable function of \(y = (y_k)_{k \in \mathbb{N}^*}\) such that \(\Delta \in \{0, 1\}\), with the convention that we reject \(H_0\) if \(\Delta = 1\), which corresponds to acceptance of \(H_1\), and do not reject \(H_0\) otherwise. Given a level \(\alpha \in (0, 1/2)\), \(\Delta = \Delta_\alpha(y)\) is called a level-\(\alpha\) test if and only if the Type-I error is no larger than \(\alpha\), i.e.,

\[ P_{0_{l_2(\mathbb{N}^*})}(\Delta_\alpha = 1) \leq \alpha. \]

The risk of the testing procedure \(\Delta_\alpha\) under the alternative hypothesis is often measured through the maximal Type-II error over the set \(\Theta\), as

\[ \beta(\Delta_\alpha, \Theta, r_\varepsilon) := \sup_{\theta \in \Theta, \|\theta\|^2 \geq r_\varepsilon^2} P_\theta(\Delta_\alpha = 0). \]

In particular, given a level \(\beta \in (0, 1/2)\), the level-\(\alpha\) test \(\Delta_\alpha\) is said to have power \(1 - \beta\) if its maximal Type-II error can be bounded by \(\beta\), namely \(\beta(\Delta_\alpha, \Theta, r_\varepsilon) \leq \beta\). In this context, the minimax paradigm has been at the core of several investigations over the last decades. Given both \(\alpha\) and \(\beta \in (0, 1/2)\) some fixed levels and a given set \(\Theta\), the separation rate \(R_\varepsilon(\Delta_\alpha, \Theta)\) associated with a level-\(\alpha\) testing procedure \(\Delta_\alpha\), which may depend on \(\varepsilon\), is defined as

\[ R_\varepsilon(\Delta_\alpha, \Theta) = \inf \{ r_\varepsilon > 0, \beta(\Delta_\alpha, \Theta, r_\varepsilon) \leq \beta \}. \]
Then, a natural way to provide a decision rule is to construct an estimator of this quantity. According to the choice of the set \( \Theta \) is then defined as \( r^*_\varepsilon \), where
\[
 r^*_\varepsilon := \inf_{\Delta} R_\varepsilon(\Delta, \Theta),
\]
the infimum being taken over all possible level-\( \alpha \) testing procedures \( \Delta \). We refer to [9] or [4] for exhaustive discussions on these definitions. Determining the minimax separation rate for a given problem is quite informative. For substantial account on the subject, see, e.g., [9], [4], [5], [15] or [13] among others.

### 2.2. Indirect and Direct Approaches

Several testing procedures have been proposed in order to deal with the testing problem (3). In this section we will focus on two \( \chi^2 \)-based test statistics that have been proved to perform well in some minimax sense.

The amount of signal contained in the observations can be measured through the quantity \( \|\theta\|^2 \). Then, a natural way to provide a decision rule is to construct an estimator of this quantity.

**Indirect approach** (IA): According to the sequence model (1), each coefficient \( \theta_k \) can be estimated by \( b_k^{-1} y_k \) provided that the sequence of \( b = (b_k)_{k \in \mathbb{N}^*} \) is known. For a given sequence of positive integers \( (D_\varepsilon)_{\varepsilon > 0} \), which is decreasing with respect to \( \varepsilon \), this leads to the testing procedure
\[
\Delta_{\alpha,\varepsilon}^{IA} = 1_{\{T_{D_\varepsilon} > t_{\alpha,\varepsilon}\}}, \quad \text{where} \quad T_{D_\varepsilon} = \sum_{k=1}^{D_\varepsilon} b_k^{-2}(y_k^2 - \varepsilon^2)
\]
and \( t_{\alpha,\varepsilon} \) is a threshold value that allows one to control the Type-I error of the test. The integer \( D_\varepsilon \) plays a similar role to a regularization parameter in estimation. According to the choice of the set \( \Theta \), specific choices for \( D_\varepsilon \) are available. We refer, e.g., to [15] for more details. We stress that a weighted version of this procedure has been proposed and investigated in several papers, as, e.g., [10], allowing to obtain sharp asymptotic results.

**Direct approach** (DA): Since the sequence model (1) is derived from the model (2), we remark that the test defined in (4) is essentially based on an inversion of the operator \( A \) at hand. Such an inversion appears to be quite natural in an estimation context in which we provide a reconstruction of the unknown function \( f \). In the signal detection framework, this is no longer required. Indeed, setting \( \vartheta = b \cdot \theta \) (i.e., \( \vartheta_k = b_k \theta_k \) for all \( k \in \mathbb{N}^* \)), we can remark that both assertions \( \theta = 0 \) \( \epsilon_2(\mathbb{N}^*) \) and \( \vartheta = 0 \) \( \epsilon_2(\mathbb{N}^*) \) are equivalent. In other words, the direct testing problem associated with the hypotheses
\[
\tilde{H}_0: \vartheta = 0 \epsilon_2(\mathbb{N}^*) \quad \text{against} \quad \tilde{H}_1: \|\vartheta\|^2 \geq \varepsilon^2,
\]
where \( \mu = (\mu_\varepsilon)_{\varepsilon > 0} \) is some continuous sequence of positive numbers that tends to zero as \( \varepsilon \to 0 \) and \( \tilde{\Theta} \) some set in \( \epsilon_2(\mathbb{N}^*) \), only differs from (3) by the alternative hypothesis. In some sense, (5) does not take into account the fact that the data are distorted by a compact operator: we treat the data as a ‘direct’ problem and deal with a model of the form
\[
Y = g + \varepsilon \xi \quad \text{or} \quad y_k = \vartheta_k + \varepsilon \xi_k \quad \forall k \in \mathbb{N}^*,
\]
where \( g = Af \in K \). Consequently, for a given sequence of positive integers \( (D_\varepsilon)_{\varepsilon > 0} \) that is decreasing with respect to \( \varepsilon \), we can introduce for any \( \varepsilon \in (0, 1) \) the testing procedure \( \Delta_{\alpha,\varepsilon}^{DA} \) as
\[
\Delta_{\alpha,\varepsilon}^{DA} = 1_{\{S_{D_\varepsilon} > s_{\alpha,\varepsilon}\}}, \quad \text{where} \quad S_{D_\varepsilon} = \sum_{k=1}^{D_\varepsilon} (y_k^2 - \varepsilon^2),
\]
the term \( S_{D_\varepsilon} \) corresponds to an estimator of \( \|\vartheta\|^2 \), and \( s_{\alpha,\varepsilon} \) denotes an appropriate threshold, allowing a control of the Type-I error. This test provides interesting performance when dealing with (5). But this is also the case for the testing problem (3) in some specific situations. We refer to [14] for an extensive discussion on the subject with a minimax point of view. One of the aims of this paper is to complete this discussion using a maxiset point of view. This notion is extended to the signal detection context in the next section.
2.3. The Maxiset Point of View in Signal Detection Problem

First we recall that in the minimax setting, given a functional set Θ, we usually determine a sequence of rates \( r = (r_\varepsilon)_{\varepsilon > 0} \) such that for any \( \varepsilon \in (0, 1) \),
\[
\inf_{\theta} \sup_{\varepsilon \in \Theta} \mathbb{E}_\theta \| \hat{\theta} - \theta \|^2 \geq cr_\varepsilon^2, \quad \sup_{\theta \in \Theta} \mathbb{E}_\theta \| \hat{\theta}_{D_\varepsilon} - \theta \|^2 \leq Cr_\varepsilon^2, \tag{8}
\]
for some constants \( c, C > 0 \). Under the maxiset paradigm, we are giving a sequence of rates \( r = (r_\varepsilon)_{\varepsilon > 0} \) and we are exhibiting the largest set of sequences \( \Theta \subset l_2(\mathbb{N}^+) \) for which (8) holds.

We will now adapt the maxiset point of view to the signal detection framework. Given a sequence of rates \( r = (r_\varepsilon)_{\varepsilon > 0} \), we determine the largest set for which the maximal Type-II error can be controlled with a prescribed error. This is formalized in the following definition.

**Definition 2.1.** Consider the testing problem defined in (3). For a fixed \((\alpha, \beta) \in (0, 1/2)^2\), let \( \Delta_\alpha = (\Delta_{\alpha, \varepsilon})_{\varepsilon > 0} \) be a sequence of testing procedures, and let \((r_\varepsilon)_{\varepsilon > 0}\) be a continuous sequence of positive numbers that tends to zero when \( \varepsilon \to 0 \). For the testing problem (3), the maxiset of \( \Delta_\alpha \) associated with the separation rate \( \varepsilon \) is the set of all \( \theta \in l_2(\mathbb{N}^+) \) such that, for all \( \varepsilon \in (0, 1) \),
\[
\mathbb{P}_{0_2(0^*)}(\Delta_{\alpha, \varepsilon} = 1) \leq \alpha \quad \text{and} \quad \sup_{\theta, \| \theta \|^2 \geq r_\varepsilon} \mathbb{P}_\theta(\Delta_{\alpha, \varepsilon} = 0) \leq \beta. \tag{9}
\]

In the following, we denote the maxiset of the sequence of level-\( \alpha \) testing procedures \( \Delta_\alpha \) by \( MS(\Delta_\alpha, r) := MS((\Delta_{\alpha, \varepsilon})_{\varepsilon > 0}, (r_\varepsilon)_{\varepsilon > 0}) \). Note that it corresponds to the set
\[
MS(\Delta_\alpha, r) = \left\{ (\theta_k)_{k \in \mathbb{N}^*} \in l_2(\mathbb{N}^*) : \forall \varepsilon \in (0, 1), \left[ \| \theta \|^2 \geq r_\varepsilon^2 \Rightarrow \mathbb{P}_\theta(\Delta_{\alpha, \varepsilon} = 0) \leq \beta \right] \right\}. \tag{10}
\]

This definition can be generalized in a straightforward way to the testing problem defined in (5), replacing \( \| \theta \|^2 \) by \( \| b \cdot \theta \|^2 \) in (9) to get:
\[
MS(\Delta_\alpha, r) = \left\{ (\theta_k)_{k \in \mathbb{N}^*} \in l_2(\mathbb{N}^*) : \forall \varepsilon \in (0, 1), \left[ \| b \cdot \theta \|^2 \geq r_\varepsilon^2 \Rightarrow \mathbb{P}_\theta(\Delta_{\alpha, \varepsilon} = 0) \leq \beta \right] \right\}. \tag{11}
\]

In Section 3, we derive an explicit expression of the maxisets for some tests based on \( \chi^2 \) statistics (see Theorems 3.1 and 3.2). Following Definition 2.1 we let the reader be convinced that, for a given sequence of testing procedures \( \Delta_\alpha = (\Delta_{\alpha, \varepsilon})_{\varepsilon > 0} \), there is an embedding result between the maxisets associated with different choices of detection rates.

**Proposition 2.1.** Let \((r'_\varepsilon)\) and \((r_\varepsilon)\) be two sequences of detection rates such that \( r'_\varepsilon \propto \varepsilon \) as \( \varepsilon \to 0 \). Let us consider a sequence of testing procedures \( \Delta_\alpha = (\Delta_{\alpha, \varepsilon})_{\varepsilon > 0} \). Then, for some \( C > 0 \)
\[
MS(\Delta_\alpha, r) \subset MS(\Delta_\alpha, Cr').
\]

This embedding entails that the set of detectable sequences enlarges as soon as we relax the constraint requirement on the minimal energy.

3. TESTING PROCEDURES AND THEIR MAXISET PERFORMANCE

In this section, we first provide a description of the maxisets with respect to (10) and associated with the procedures \((\Delta_{\alpha, \varepsilon}^{lA})_{\varepsilon > 0}\) and \((\Delta_{\alpha, \varepsilon}^{DA})_{\varepsilon > 0}\) defined in (4) and (7) respectively. From now on, we fix the two thresholds \( t_{\alpha, \varepsilon} \) and \( s_{\alpha, \varepsilon} \) involved in the definitions of these testing procedures as
\[
t_{\alpha, \varepsilon} = C_{\alpha, 1} \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \quad \text{and} \quad s_{\alpha, \varepsilon} = C_{\alpha, 2} \varepsilon^2 \sqrt{D_\varepsilon} \quad \forall \varepsilon \in (0, 1), \tag{12}
\]
where \( C_{\alpha, 1}, C_{\alpha, 2} > 0 \) denote two explicit constants that guarantee that the considered procedures have a Type-I error controlled by \( \alpha \) for all \( \varepsilon \in (0, 1) \). Undoubtedly, the smaller is \( \alpha \), the bigger \( C_{\alpha, 1} \) and \( C_{\alpha, 2} \) are necessary to choose. For the sake of simplicity, we do not use the \( 1 - \alpha \) quantile of the respective test statistics \((S_{D_\varepsilon})_{\varepsilon > 0}\) and \((T_{D_\varepsilon})_{\varepsilon > 0}\). Such a change will not modify the spirit of the results displayed in this paper but will induce more technical details.

Below we characterize the maxisets associated with the considered testing procedures for general separation rates. In such a setting, these sets are poorly informative but they highlight valuable information on the problem provided that we impose some structural constraints.
3.1. A General Characterization of the Maxisets

In this section we start with introducing the two sets $F_{r,D}(C)$ and $G_{r,D}(C)$, which will be of the utmost importance in the sequel.

**Definition 3.1.** Let $(D_s)_{s>0}$ be a sequence of positive integers that is decreasing with respect to $s$. Let $r = (r_s)_{s>0}$ be a chosen sequence of rates. For any $C > 0$ we set

$$F_{r,D}(C) = \left\{ \theta \in l_2(\mathbb{N}^s), \forall \varepsilon \in (0,1); \left[ \| \theta \|^2 \geq r^2_\varepsilon \Rightarrow \sum_{k=1}^{D_s} \theta_k^2 > C \varepsilon^2 \sqrt{\sum_{k=1}^{D_s} b_k^{-4}} \right] \right\},$$

$$G_{r,D}(C) = \left\{ \theta \in l_2(\mathbb{N}^s), \forall \varepsilon \in (0,1); \left[ \| b \cdot \theta \|^2 \geq r^2_\varepsilon \Rightarrow \sum_{k=1}^{D_s} b_k^2 \theta_k^2 > C \varepsilon^2 \sqrt{D_\varepsilon} \right] \right\}.$$

The following results show that these sets provide a characterization of the maxisets associated with the procedures $(\Delta_{A,\varepsilon}^{IA})_{\varepsilon>0}$ and $(\Delta_{A,\varepsilon}^{DA})_{\varepsilon>0}$.

**Theorem 3.1.** Consider $(\alpha, \beta) \in (0, 1/2)^2$, Let $(t_{\alpha,\varepsilon})_{\varepsilon>0}$ and $(s_{\alpha,\varepsilon})_{\varepsilon>0}$ satisfy (12). Consider the two sequences of testing procedures $(\Delta_{A,\varepsilon}^{IA})_{\varepsilon>0}$ and $(\Delta_{A,\varepsilon}^{DA})_{\varepsilon>0}$ defined respectively in (4) and (7). We have the two following maxiset results for any choice of detection rates $r = (r_s)_{s>0}$ and $\mu = (\mu_s)_{s>0}$:

1. There exist two positive constants $C_{\min}(\alpha, \beta)$ and $C_{\max}(\alpha, \beta)$ depending on $C_{\alpha,1}$ and $\beta$ such that

$$F_{r,D}(C_{\max}(\alpha, \beta)) \subset MS(\Delta_{A}^{IA}, r) \subset F_{r,D}(C_{\min}(\alpha, \beta)),$$

rewritten as $MS(\Delta_{A}^{IA}, r) = F_{r,D}$.

2. There exist two positive constants $C'_{\min}(\alpha, \beta)$ and $C'_{\max}(\alpha, \beta)$ depending on $C_{\alpha,2}$ and $\beta$ such that

$$G_{\mu,D}(C'_{\max}(\alpha, \beta)) \subset MS(\Delta_{A}^{DA}, \mu) \subset G_{\mu,D}(C'_{\min}(\alpha, \beta)),$$

rewritten as $MS(\Delta_{A}^{DA}, \mu) = G_{\mu,D}$.

**Remark 3.1.** In Section 6, we provide explicit values of the constants $C_{\max}(\alpha, \beta)$ and $C_{\min}(\alpha, \beta)$ (see (19) and (21)). The constants $C'_{\max}(\alpha, \beta)$ and $C'_{\min}(\alpha, \beta)$ are obtained from the values of $C_{\max}(\alpha, \beta)$ and $C_{\min}(\alpha, \beta)$ by replacing $C_{\alpha,1}$ by $C'_{\alpha,2}$.

Surprisingly, the maxisets in the testing case have a completely different form compared to results obtained in the estimation case. Indeed, according to [12], the constraint that a given procedure attains the rate $(r_s)_{s>0}$ induces a tail constraint on the signal of interest in the estimation problem. This is no longer the case in the signal detection problem. Theorem 3.1 indicates that the testing procedures $(\Delta_{A,\varepsilon}^{IA})_{\varepsilon>0}$ are only able to detect signals satisfying

$$\sum_{k=1}^{D_s} \theta_k^2 > C \varepsilon^2 \sqrt{\sum_{k=1}^{D_s} b_k^{-4}}, \quad \forall \varepsilon \in (0,1), \quad \ldots (13)$$

for $C$ large enough. Analogously, Theorem 3.1 indicates that the testing procedures $(\Delta_{A,\varepsilon}^{DA})_{\varepsilon>0}$ are only able to detect signals satisfying

$$\sum_{k=1}^{D_s} b_k^2 \theta_k^2 > C \varepsilon^2 \sqrt{D_\varepsilon}, \quad \forall \varepsilon \in (0,1), \quad \ldots (14)$$
In particular, inequalities (13) and (14) indicate that there should be enough signal on the low frequencies used by the test statistics $T_{D_\varepsilon}$ and $S_{D_\varepsilon}$. Nothing is said regarding the high frequency content, i.e., the coefficients after the rank $D_\varepsilon$. The maxiset results of Theorem 3.1 contrast with usual ones since they are not stated in terms of smoothness spaces. Moreover this constraint has already been highlighted in, e.g., [15] hence, at the first glance, this maxiset result is poorly informative. Nevertheless in the following section, we prove that an additional structural assumption on the maxiset provides valuable information on the signal that can be detected by the procedures we are interested in.

### 3.2. A Robust Version of Maxisets for Tests

The main message of the previous section is that the functions that can be detected by the sequences of testing procedures $(\Delta_{\alpha,\varepsilon})_{\varepsilon>0}$ and $(\Delta_{\alpha,\varepsilon}^A)_{\varepsilon>0}$ are the ones that have enough energy on low frequencies. It means in particular that our testing procedures are very sensitive to any distortion of the signal involving low frequencies. In what follows, we shall require some robustness of our testing procedures with respect to this low frequency content of the signal, provided that we have enough information. This structural constraint on the maxisets of interest can be reformulated in a more formal way as follows.

**Definition 3.2.** A set $\mathcal{H} \subset l_2(\mathbb{N}^*)$ is said to be robust with respect to filtering if for any filter $h = (h_k) \in \ell_\infty(\mathbb{N}^*)$ such that $\|h\|_{\ell_\infty(\mathbb{N}^*)} \leq 1$

$$\theta \in \mathcal{H} \Rightarrow h \cdot \theta \in \mathcal{H},$$

where $\cdot$ denotes the componentwise multiplication.

The constraint displayed in Definition 3.2 is motivated by practical situations in which signals are often preprocessed or filtered. Hence it seems natural to ensure that, given a function set $\mathcal{H}$ of interest, the sequence $h \cdot \theta = (h_k\theta_k)_{k \in \mathbb{N}}$ belongs to $\mathcal{H}$ provided $\theta = (\theta_k)_{k \in \mathbb{N}} \in \mathcal{H}$.

We stress that such a condition is for instance satisfied by all the sets $\Theta \subset l_2(\mathbb{N}^*)$ of the form

$$\Theta(\gamma, L) = \{ v = (v_k)_{k \in \mathbb{N}^*} \in l_2(\mathbb{N}^*) : \sum_{k \in \mathbb{N}^*} w_k \gamma_k^2 < L \},$$

for some positive sequence $(w_k)_{k \in \mathbb{N}^*}$ and positive constants $\gamma$ and $L$. Such a set describes some smoothness conditions through the decay of the coefficients of the function of interest.

We now define the so-called robust maxisets. They satisfy both the requirement of Definition 2.1 and the robustness with respect to filtering property.

**Definition 3.3.** Consider the testing problem defined in (3). For a fixed $(\alpha, \beta) \in (0, 1/2)^2$, let $\Delta_\alpha = (\Delta_{\alpha,\varepsilon})_{\varepsilon>0}$ be a sequence of testing procedures and $(r_\varepsilon)_{\varepsilon>0}$ be a sequence of rates. For the testing problem (3), the robust maxiset of $\Delta_\alpha$ associated with the sequence of separation rates $r = (r_\varepsilon)_{\varepsilon>0}$ is the set of all $\theta \in l_2(\mathbb{N}^*)$ that is robust with respect to filtering and such that, for all $\varepsilon \in (0, 1)$

$$\mathbb{P}_{0_{l_2(\mathbb{N}^*)}}(\Delta_{\alpha,\varepsilon} = 1) \leq \alpha \quad \text{and} \quad \sup_{\theta, \|\theta\|^2 \geq r_\varepsilon^2} \mathbb{P}_\theta(\Delta_{\alpha,\varepsilon} = 0) \leq \beta.$$

This definition can be generalized in a straightforward way to the testing problem defined in (5), replacing $\|\theta\|^2$ by $\|b \cdot \theta\|^2$.

We now give a characterization of robust maxisets associated with the two sequences of testing procedures $(\Delta_{\alpha,\varepsilon}^A)_{\varepsilon>0}$ and $(\Delta_{\alpha,\varepsilon}^A)_{\varepsilon>0}$ defined respectively in (4) and (7). We first need to define robust versions of $\mathcal{F}_{r,D}$ and $\mathcal{G}_{\mu,D}$:

**Definition 3.4.** Let $r = (r_\varepsilon)_{\varepsilon>0}$ be a sequence of rates and $D = (D_\varepsilon)_{\varepsilon>0}$ be a sequence of positive integers that is decreasing with respect to $\varepsilon$. For any $C > 0$ we define the sets

$$\mathcal{F}_{r,D}^{\text{filt}}(C) = \left\{ \theta \in l_2(\mathbb{N}^*), \forall \varepsilon \in (0, 1); \sum_{k > D_\varepsilon} \theta_k^2 < r_\varepsilon^2 - C\varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \right\},$$

$$\mathcal{G}_{r,D}^{\text{filt}}(C) = \left\{ \theta \in l_2(\mathbb{N}^*), \forall \varepsilon \in (0, 1); \sum_{k > D_\varepsilon} b_k^2 \theta_k^2 < r_\varepsilon^2 - C\varepsilon^2 \sqrt{D_\varepsilon} \right\}.$$
Undoubtedly, following the choices of \( D = (D_\varepsilon)_{\varepsilon > 0}, r = (r_\varepsilon)_{\varepsilon > 0}, \) and \( C > 0, \) the sequence spaces above can be identical to the empty space. In the sequel, in order to get a nonvoid theory we mainly focus on the cases where these sequence spaces are not empty.

**Definition 3.5.** For any chosen sequences \( D = (D_\varepsilon)_{\varepsilon > 0}, r = (r_\varepsilon)_{\varepsilon > 0}, \) and \( C > 0, \) we say that \((r, D, C)\) is \( \mathcal{F} \)-admissible (respectively \( \mathcal{G} \)-admissible) if and only if \( \mathcal{F}_{r, D}^{\text{fill}}(C) \) (respectively \( \mathcal{G}_{r, D}^{\text{fill}}(C) \)) is not the empty space.

**Remark 3.2.** Following Definition 3.5, note that \((r, D, C)\) is \( \mathcal{F} \)-admissible and \( \mathcal{G} \)-admissible for rates of detection that do not converge to zero too fast as \( \varepsilon \to 0.\)

**Theorem 3.2.** Consider \((\alpha, \beta) \in (0, 1/2)^2.\) Let \((t_{\alpha, \varepsilon})_{\varepsilon > 0} \) and \((s_{\alpha, \varepsilon})_{\varepsilon > 0} \) satisfy (12). Consider the two sequences of testing procedures \((\Delta^I_{\alpha, \varepsilon})_{\varepsilon > 0} \) and \((\Delta^D_{\alpha, \varepsilon})_{\varepsilon > 0} \) defined in (4) and (7) respectively.

Denote by \( MS^{\text{fill}}(\Delta^I_{\alpha}, r) \) and \( MS^{\text{fill}}(\Delta^D_{\alpha}, \mu) \) respectively their robust maxisets associated with the chosen rates \( r = (r_\varepsilon)_{\varepsilon > 0} \) and \( \mu = (\mu_\varepsilon)_{\varepsilon > 0} \). We have the following maxiset results:

1. If \((r, D, C_{\max}(\alpha, \beta))\) is \( \mathcal{F} \)-admissible, then
   \[
   \mathcal{F}_{r, D}^{\text{fill}}(C_{\max}(\alpha, \beta)) \subset MS^{\text{fill}}(\Delta^I_{\alpha}, r) \subset \mathcal{F}_{\sqrt{2r}, D}^{\text{fill}}(C_{\min}(\alpha, \beta)),
   \]
   rewritten as \( MS^{\text{fill}}(\Delta^I_{\alpha}, r) = \mathcal{F}_{r, D}^{\text{fill}}.\)

2. If \((\mu, D, C'_{\max}(\alpha, \beta))\) is \( \mathcal{G} \)-admissible, then
   \[
   \mathcal{G}_{\mu, D}^{\text{fill}}(C'_{\max}(\alpha, \beta)) \subset MS^{\text{fill}}(\Delta^D_{\alpha}, \mu) \subset \mathcal{G}_{\sqrt{2\mu}, D}^{\text{fill}}(C'_{\min}(\alpha, \beta)),
   \]
   rewritten as \( MS^{\text{fill}}(\Delta^D_{\alpha}, \mu) = \mathcal{G}_{\mu, D}^{\text{fill}}.\)

**Remark 3.3.** The constants stated in Theorem 3.2 are similar to those given in Theorem 3.1.

We observe that, as in the estimation case, the robust maxiset depends on the tail of the sequence \( \theta = (\theta_k)_{k \in \mathbb{N}^*} \) of interest. According to the relative growth of the levels of possible energies of the signal and the sums of the power of the eigenvalues of the operator involved in the signal detection problem, the nature of the robust maxiset related to the sequence of testing procedures might be different. Consider the case where the sequence of testing procedures is \( \Delta^I_{\alpha, \varepsilon}. \) There are two extreme regimes:

1. First regime: \( r_\varepsilon^2 = o(\varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}}) \) as \( \varepsilon \to 0. \) In this case, Theorem 3.2 implies that the robust maxiset is the empty set. It means that the sequence of testing procedures \( \Delta^D_{\alpha} = (\Delta^D_{\alpha, \varepsilon})_{\varepsilon > 0} \) is not able to detect signals when the chosen sequence \( r = (r_\varepsilon)_{\varepsilon > 0} \) converges toward zero too fast.

2. Second regime: \( \varepsilon^2 \sum_{k=1}^{D_\varepsilon} b_k^{-4} = o(r_\varepsilon^2) \) as \( \varepsilon \to 0. \) In this case, the robust maxiset is non empty and does not depend on the operator. In particular, provided that the signal has enough energy, the performance of the detection procedure does not depend on the underlying inverse problem that is considered.

The transition regime where the two sequences are equally balanced \( (r_\varepsilon^2 = O(\varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}})) \) is called hereafter the horizon of detection. Here, the robust maxiset can be explicitly embedded as follows:

\[
MS^{\text{fill}}(\Delta^I_{\alpha}, r) \subset \left\{ \theta \in l_2(\mathbb{N}^*), \forall \varepsilon \in (0, 1); \sum_{k>D_\varepsilon} \theta_k^2 < Kr_\varepsilon^2 \right\},
\]

for some positive constant \( K. \) The set on the right-hand side of the previous embedding provides a control on the tail of the sequence of interest by the considered rate \( (r_\varepsilon)_{\varepsilon > 0}. \) In particular, the faster the sequence \( (b_k)_{k \in \mathbb{N}^*} \) converges toward zero, the farther the horizon of detection.
Proposition 4.2. Let \( r = (r_\varepsilon)_{\varepsilon > 0} \), \( \mu = (\mu_\varepsilon)_{\varepsilon > 0} \), and \( D = (D_\varepsilon)_{\varepsilon > 0} \) such that \( \mu_\varepsilon = b_{D_\varepsilon} r_\varepsilon \) for any \( \varepsilon \in (0, 1) \). Then, provided that
\[
C'_{\text{max}}(\alpha, \beta) \sqrt{k} \leq C_{\text{min}}(\alpha, \beta) b_k^2 \sum_{j=1}^{k} b_j^{-4} \quad \text{for all} \quad k \in \mathbb{N}^*,
\]
we get
\[
MS^{\text{fill}}(\Delta_{\alpha}^{DA}, r) \subset MS^{\text{fill}}(\Delta_{\alpha}^{DA}, \mu).
\]

This proposition indicates that all the sequences that can be detected by \( \Delta_{\alpha}^{DA} \) can be also be detected by \( \Delta_{\alpha}^{DA} \). In other words, the direct test appears to be more efficient in the sense that its associated maxiset is larger. A question which one may ask is whether the inclusion is strict or not. In order to provide an answer, we will consider a specific setting and prove that the indirect testing method may lose some functions that can be detected by the direct one.

We are now considering the classical setting of the mildly ill-posed inverse problem, namely we assume that \( b_k \sim k^{-t} \) for some \( t > 0 \) and any \( k \). We also assume that we are in the case where the calibration is the minimax one. In this case the two terms \( r_\varepsilon^2 \) and \( \sqrt{\sum_{k=1}^{D_k} b_k^{-4}} \) are equally balanced, so that we are in the limit case, described in Section 3.1.

Proposition 4.2. Let \( s, t > 0 \). Consider the case, where \( D_\varepsilon \sim \varepsilon^{-\frac{4s+4}{4s+4t+1}} \) for any \( \varepsilon \in (0, 1) \). Assume that \( r_\varepsilon \sim \varepsilon^{-\frac{4s}{4s+4t+1}} \) and \( \mu_\varepsilon \sim b_{D_\varepsilon} r_\varepsilon \sim \varepsilon^{-\frac{4s+4}{4s+4t+1}} \) with \( b \sim (b_k)_{k \in \mathbb{N}^*} \sim (k^{-t})_{k \in \mathbb{N}^*} \). Then there exist sequences \( \theta = (\theta_k)_{k \in \mathbb{N}^*} \) such that
\[
\theta \in MS^{\text{fill}}(\Delta_{\alpha}^{DA}, \mu) \quad \text{but} \quad \theta \notin MS^{\text{fill}}(\Delta_{\alpha}^{IA}, r).
\]

Remark 4.1. With the choice of operator given in Proposition 4.2 and appropriate choices of calibration thresholds, (17) is clearly satisfied and therefore (18) holds and the embedding is strict.
In this paper, we adapt the maxiset approach for signal detection in inverse problem. This paradigm provides a novel way for researchers to assess the performance of their testing procedures. We illustrate this ability by comparing the so-called direct and indirect approaches. The maxiset approach provides three novel arguments describing the better performance of direct methods. Firstly, we have established that, in many cases, direct methods are associated with strictly larger maxisets. Secondly, the horizon of detection is always the same for the direct method while being dependent on the decay of the eigenvalues of the operator for the indirect method. Thirdly, when comparing to the indirect method, the direct method starts to detect at faster rates since, for any $\varepsilon \in (0, 1)$, $\varepsilon^2 \sqrt{D_\varepsilon} \leq \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}}$.

At the core of our future investigations is the adaptation of the maxiset paradigm to characterize the largest set of operators such that the performance of an estimation or a testing method reaches a prescribed rate of convergence for a given functional space, leading to the new concept of max-class of operators.

6. PROOFS

6.1. Technical Results

In this section, we recall and slightly extend some results that will be useful in the following. More details regarding these results, e.g., context and extended discussions, can be found in [4] and [15].

**Proposition 6.1.** Fix $(\alpha, \beta) \in (0, 1/2)^2$. There exist $C_{\text{max}}(\alpha, \beta)$ and $C_{\text{min}}(\alpha, \beta)$ such that, for all $\varepsilon \in (0, 1)$ and $D_\varepsilon \in \mathbb{N}^*$,

\[
\begin{align*}
(i) \quad & \sum_{k=1}^{D_\varepsilon} \theta_k^2 > C_{\text{max}}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \Rightarrow P_\theta(\Delta_{\alpha,\varepsilon}^I = 0) \leq \beta, \\
(ii) \quad & \sum_{k=1}^{D_\varepsilon} \theta_k^2 \leq C_{\text{min}}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \Rightarrow P_\theta(\Delta_{\alpha,\varepsilon}^I = 0) > \beta.
\end{align*}
\]

**Proof.** We start with the proof of item (i). A more precise proof is provided in [15]. In particular, the authors take advantage on available results on $\chi^2$ weighted statistics to characterize more precisely the dependence of the constant $C_{\text{max}}(\alpha, \beta)$ on the values of $\alpha$ and $\beta$. For the sake of completeness, we reproduce a simpler version of the proof, based on the Markov Inequality. Recall that we defined, for any $\varepsilon \in (0, 1)$, $T_{D_\varepsilon}$ as

\[ T_{D_\varepsilon} = \sum_{k=1}^{D_\varepsilon} b_k^{-2}(y_k^2 - \varepsilon^2) \quad \text{and} \quad t_{\alpha,\varepsilon} = C_{\alpha,1} \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \text{ for some constant } C_{\alpha,1} > 0. \]

Clearly

- $E_\theta(T_{D_\varepsilon}) = \sum_{k=1}^{D_\varepsilon} \theta_k^2$,
- $\text{Var}_\theta(T_{D_\varepsilon}) = 2\varepsilon^4 \sum_{k=1}^{D_\varepsilon} b_k^{-4} + 4\varepsilon^2 \sum_{k=1}^{D_\varepsilon} b_k^{-2} \theta_k^2$.

Let $\theta = (\theta_k)_{k \in \mathbb{N}^*}$ be such that

\[ \sum_{k=1}^{D_\varepsilon} \theta_k^2 > C_{\text{max}}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}}. \]

Provided that $C_{\text{max}}(\alpha, \beta) > C_{\alpha,1}$, by using the Bienayme–Chebyshev inequality, one gets

\[ P_\theta(\Delta_{\alpha,\varepsilon}^I = 0) = P_\theta(T_{D_\varepsilon} \leq t_{\alpha,\varepsilon}) = P_\theta(E_\theta(T_{D_\varepsilon}) - T_{D_\varepsilon} \geq E_\theta(T_{D_\varepsilon}) - t_{\alpha,\varepsilon}) \]
We let the reader be convinced that the smaller \( \beta \) the larger the chosen \( C_{\text{max}}(\alpha, \beta) \) given in (19).

**Remark 6.1.** We let the reader be convinced that the smaller \( \beta \) the larger the chosen \( C_{\text{max}}(\alpha, \beta) \) given in (19).

We now prove item (ii) of Proposition 6.1. To prove that \( \mathbb{P}_\theta(\Delta_{\alpha, \varepsilon}^I = 0) > \beta \) is equivalent to proving that

\[
\mathbb{P}_\theta(T_{D\varepsilon} - \mathbb{E}(T_{D\varepsilon}) > t_{\alpha, \varepsilon} - \mathbb{E}(T_{D\varepsilon})) < 1 - \beta.
\]

(20)

To show this inequality, we apply Lemma 2 of Laurent et al. (2012) with \( x := -\log(1 - \beta) \) and \( \sigma_k \equiv \varepsilon b_k^{-1} \). Setting \( \Sigma = \varepsilon^4 \sum_{k=1}^{D\varepsilon} b_k^{-4} + 2\varepsilon^2 \sum_{k=1}^{D\varepsilon} b_k^{-2} \theta_k^2 \), we then get that (20) holds provided that

\[
t_{\alpha, \varepsilon} - \mathbb{E}(T_{D\varepsilon}) \geq 2\sqrt{\Sigma x} + 2 \max_{l=1, \ldots, D\varepsilon} b_l^{-2} \varepsilon^2 x.
\]

Observe that

\[
\sqrt{\Sigma} \leq \varepsilon^2 \sqrt{\sum_{k=1}^{D\varepsilon} b_k^{-4}} + 2\varepsilon^2 \sum_{k=1}^{D\varepsilon} b_k^{-2} \theta_k^2
\]

\[
\leq \varepsilon^2 \sqrt{\sum_{k=1}^{D\varepsilon} b_k^{-4}} + \sqrt{2\varepsilon} \max_{l=1, \ldots, D\varepsilon} b_l^{-1} \sqrt{\sum_{k=1}^{D\varepsilon} \theta_k^2}
\]

\[
\leq \left(1 + \sqrt{2C_{\text{min}}(\alpha, \beta)}\right) \varepsilon^2 \sqrt{\sum_{k=1}^{D\varepsilon} b_k^{-4}},
\]

the last inequality follows from the assumption

\[
\sum_{k=1}^{D\varepsilon} \theta_k^2 \leq C_{\text{min}}(\alpha, \beta) \varepsilon^2 \sqrt{\sum_{k=1}^{D\varepsilon} b_k^{-4}}.
\]

Since \( \beta < 1/2, \ v < 1 \) (and therefore \( x < \sqrt{x} \)), a sufficient condition to get the inequality \( t_{\alpha, \varepsilon} - \mathbb{E}(T_{D\varepsilon}) \geq 2\sqrt{\Sigma x} + 2 \max_{l=1, \ldots, D\varepsilon} b_l^{-2} \varepsilon^2 x \) is

\[
C_{\alpha, \varepsilon} \varepsilon^2 \sum_{k=1}^{D\varepsilon} b_k^{-4} - \sum_{k=1}^{D\varepsilon} \theta_k^2 \geq 2\left(1 + \sqrt{2C_{\text{min}}(\alpha, \beta)}\right) \varepsilon^2 \sqrt{-\log(1 - \beta)}
\]

\[
+ 2b_{D\varepsilon}^{-2} \varepsilon^2 \sqrt{-\log(1 - \beta)}.
\]
This relationship is satisfied if \( C_{\text{min}}(\alpha, \beta) \) is small enough. The choice

\[
C_{\text{min}}(\alpha, \beta) = \left( \sqrt{-2 \log(1 - \beta) + (C_{\alpha,1} - 4\sqrt{-\log(1 - \beta)})} - \sqrt{-2 \log(1 - \beta)} \right)^{\frac{1}{2}}
\] (21)

can be done in that context. This finishes the proof since this condition is compatible with the fact that \( \Delta_{\alpha,\varepsilon}^{IA} \) is an \( \alpha \)-level test for a chosen \( C_{\alpha,1} \) sufficiently large.

\[\square\]

**Remark 6.2.** The smaller the \( \beta \) the more difficult it is to detect signals. The increased difficulty of the signal detection problem is captured by the value of \( C_{\text{min}}(\alpha, \beta) \) given in (21). Indeed, we let the reader be convinced that the smaller the \( \beta \) the larger the chosen \( C_{\text{min}}(\alpha, \beta) \) and therefore the larger the \( l_2 \) norm of the first terms of the signal \( \theta \) is required to ensure that \( P_{\theta}(\Delta_{\alpha,\varepsilon}^{IA} = 0) \leq \beta \) where \( \varepsilon \in (0,1) \).

In the same spirit as Proposition 6.1, the following result holds too.

**Proposition 6.2.** There exist \( C'_{\text{min}}(\alpha, \beta) \) and \( C'_{\text{max}}(\alpha, \beta) \) such that, for all \( \varepsilon \in (0,1) \),

\[
(i) \quad \sum_{k=1}^{D_x} b_k^2 \theta_k^2 > C'_{\text{max}}(\alpha, \beta) \varepsilon^2 \sqrt{D_x} \quad \Rightarrow \quad P_{\theta}(\Delta_{\alpha,\varepsilon}^{DA} = 0) \leq \beta,
\]

\[
(ii) \quad \sum_{k=1}^{D_x} b_k^2 \theta_k^2 < C'_{\text{min}}(\alpha, \beta) \varepsilon^2 \sqrt{D_x} \quad \Rightarrow \quad P_{\theta}(\Delta_{\alpha,\varepsilon}^{DA} = 0) > \beta.
\]

Since the proof of Proposition 6.2 is analogous to that of Proposition 6.1, we omit it. We point out that the values of \( C'_{\text{min}}(\alpha, \beta) \) and \( C'_{\text{max}}(\alpha, \beta) \) are obtained from the ones of \( C_{\text{min}}(\alpha, \beta) \) and \( C_{\text{max}}(\alpha, \beta) \) by replacing \( C_{\alpha,1} \) by \( C_{\alpha,2} \).

\[\square\]

6.2. Proof of Theorem 3.1

We first prove the following embedding property: \( \mathcal{F}_{r,D}(C_{\text{max}}(\alpha, \beta)) \subset MS(\Delta_{\alpha,\varepsilon}^{IA}, r) \) and thereafter we prove that the embedding property \( MS(\Delta_{\alpha,\varepsilon}^{IA}, r) \subset \mathcal{F}_{r,D}(C_{\text{min}}(\alpha, \beta)) \) holds whatever the choice of the detection rate \( r = (r_\varepsilon)_{\varepsilon>0} \).

First we prove that \( \mathcal{F}_{r,D}(C_{\text{max}}(\alpha, \beta)) \subset MS(\Delta_{\alpha,\varepsilon}^{IA}, r) \). Fix \( \varepsilon \in (0,1) \) and let \( \theta \in \mathcal{F}_{r,D}(C_{\text{max}}(\alpha, \beta)) \) satisfying \( \|\theta\|^2 \geq r_{\varepsilon}^2 \). Then

\[
\sum_{k=1}^{D_x} \theta_k^2 > C_{\text{max}}(\alpha, \beta) \sqrt{\sum_{k=1}^{D_x} b_k^4} \quad \text{and} \quad P_{\theta}(\Delta_{\alpha,\varepsilon}^{IA} = 0) \leq \beta
\]

because of item (i) of Proposition 6.1.

Conversely, we shall prove that \( MS(\Delta_{\alpha,\varepsilon}^{IA}, r) \subset \mathcal{F}_{D}(C_{\text{min}}(\alpha, \beta)) \), or equivalently, that

\[
\theta \notin \mathcal{F}_{r,D}(C_{\text{min}}(\alpha, \beta)) \quad \Rightarrow \quad \theta \notin MS(\Delta_{\alpha,\varepsilon}^{IA}, r).
\]

Assume that \( \theta \notin \mathcal{F}_{D}(C_{\text{min}}(\alpha, \beta)) \). Then there exists \( \varepsilon_0 \in (0,1) \) such that

\[
\|\theta\|^2 \geq r_{\varepsilon_0}^2 \quad \text{and} \quad \sum_{k=1}^{D_{0\theta}} \theta_k^2 \leq C_{\text{min}}(\alpha, \beta) \varepsilon_0^2 \sqrt{\sum_{k=1}^{D_{0\theta}} b_k^4}.
\]

Following item (ii) of Proposition 6.1, one deduces that

\[
\|\theta\|^2 \geq r_{\varepsilon_0}^2 \quad \text{and} \quad P_{\theta}(\Delta_{\alpha,\varepsilon_0} = 0) > \beta.
\] (22)

Therefore we immediately deduce that \( \theta \notin MS(\Delta_{\alpha,\varepsilon}^{IA}, r) \).

The proof of the second part of Theorem 3.1 follows exactly the same lines and is left to the interested reader.
6.3. Proof of Theorem 3.2

As in the previous proof, we concentrate our attention on the maxiset associated with the sequence of testing procedures $(\Delta_{\alpha,\varepsilon})_{\varepsilon > 0}$. We recall that our aim is to prove that the robust maxiset $MS^{\text{filt}}(\Delta_{\alpha}^{IA}, r)$ can be identified, up to a constant, with the set $\mathcal{F}_{r,D}^{\text{filt}}(C)$ defined as

$$\mathcal{F}_{r,D}^{\text{filt}}(C) = \left\{ \theta \in l_2(\mathbb{N}^*) \setminus \{0\} : \sum_{k > D_\varepsilon} \theta^2_k \leq r^2_{\varepsilon} - C\varepsilon^2 \sum_{k=1}^{D_\varepsilon} b_k^{-4} \right\}.$$ 

We begin the proof with the first inclusion that is $\mathcal{F}_{r,D}^{\text{filt}}(C_{\max}(\alpha, \beta)) \subset MS^{\text{filt}}(\Delta_{\alpha}^{IA}, r)$. It can be easily seen that for all $\varepsilon \in (0, 1)$

$$\|\theta\|^2 \geq r^2_{\varepsilon} \quad \text{and} \quad \theta \in \mathcal{F}_{r,D}^{\text{filt}}(C_{\max}(\alpha, \beta)) \Rightarrow \sum_{k > D_\varepsilon} \theta^2_k < \|\theta\|^2 - C_{\max}(\alpha, \beta)\varepsilon^2 \sum_{k=1}^{D_\varepsilon} b_k^{-4} 
\Rightarrow \sum_{k=1}^{D_\varepsilon} \theta^2_k > C_{\max}(\alpha, \beta)\varepsilon^2 \sum_{k=1}^{D_\varepsilon} b_k^{-4} 
\Rightarrow \mathbb{P}_\theta(\Delta_{\varepsilon, \alpha, r}^{IA} = 0) \leq \beta \quad \text{because of item (i) of Proposition 6.1.}$$

Since one can easily check that $\mathcal{F}_{r,D}^{\text{filt}}(C_{\max}(\alpha, \beta))$ is robust by filtering, this entails the following embedding: $\mathcal{F}_{r,D}^{\text{filt}}(C_{\max}(\alpha, \beta)) \subset MS^{\text{filt}}(\Delta_{\alpha}^{IA}, r)$.

Now we turn our attention to the proof of the second inclusion, that is

$$MS^{\text{filt}}(\Delta_{\alpha}^{IA}, r) \subset \mathcal{F}_{\sqrt{2r}, D}^{\text{filt}}(C_{\min}(\alpha, \beta)).$$

We consider the situation, where for any $\varepsilon \in (0, 1)$

$$C_{\max}(\alpha, \beta)\varepsilon^2 \sum_{k=1}^{D_\varepsilon} b_k^{-4} < r^2_{\varepsilon}$$

(otherwise $\mathcal{F}_{r,D}^{\text{filt}}(C_{\max}(\alpha, \beta))$ is empty).

Let $\theta \notin \mathcal{F}_{\sqrt{2r}, D}^{\text{filt}}(C_{\min}(\alpha, \beta))$. In particular, since clearly $C_{\max}(\alpha, \beta) \geq C_{\min}(\alpha, \beta)$, there exists $\varepsilon_1 \in (0, 1)$ such that

$$\sum_{k > D_{\varepsilon_1}} \theta^2_k \geq 2r^2_{\varepsilon_1} - C_{\min}(\alpha, \beta)\varepsilon_1^2 \sum_{k=1}^{D_{\varepsilon_1}} b_k^{-4} \geq 2r^2_{\varepsilon_1} - C_{\max}(\alpha, \beta)\varepsilon^2 \sum_{k=1}^{D_\varepsilon} b_k^{-4} > r^2_{\varepsilon_1}.$$ 

We now consider a specific filter and set $h = (1_{k > D_{\varepsilon_1}})_{k \in \mathbb{N}^*}$, which clearly belongs to $l_\infty(\mathbb{N}^*)$. The sequence $h \cdot \theta$ satisfies both

$$\|h \cdot \theta\|^2 = \sum_{k > D_{\varepsilon_1}} \theta^2_k > r^2_{\varepsilon_1} \quad \text{and} \quad \mathbb{P}_{h \cdot \theta}(\Delta_{\varepsilon, \alpha, 1}^{IA} = 0) > \beta,$$

where the last inequality is obtained thanks to item (ii) of Proposition 6.1 and the fact that $h_k \theta_k = 0$ for all $k \leq D_{\varepsilon_1}$. This entails that $h \cdot \theta \notin MS(\Delta_{\alpha}^{IA}, r)$ and therefore $\theta \notin MS^{\text{filt}}(\Delta_{\alpha}^{IA}, r)$ because of the robustness with respect to filtering constraint. Hence we get

$$\theta \notin \mathcal{F}_{\sqrt{2r}, D}^{\text{filt}}(C_{\min}(\alpha, \beta)) \quad \Rightarrow \quad \theta \notin MS(\Delta_{\alpha}^{IA}, r),$$

which implies that

$$MS^{\text{filt}}(\Delta_{\alpha}^{IA}, r) \subset \mathcal{F}_{\sqrt{2r}, D}^{\text{filt}}(C_{\min}(\alpha, \beta)).$$
The proof of the second part of Theorem 3.2 follows exactly the same lines and is left to the interested reader.

This concludes the proof. 

6.4. Proof of Proposition 4.1

The proof of Proposition 4.1 directly follows from Theorem 3.2 and from the fact that 
\[ \mathcal{F}_{r,D}(C_{\min}(\alpha, \beta)) \subset \mathcal{G}_{\mu,D}(C'_{\max}(\alpha, \beta)). \]

Let us prove this embedding of sequence spaces. Assume first that \( \theta \in \mathcal{F}_{r,D}(C_{\min}(\alpha, \beta)) \), then, for any \( \varepsilon \in (0, 1) \),
\[
\sum_{k > D_\varepsilon} b_k^2 \theta^2_k \leq b_{D_\varepsilon}^2 \left( \sum_{k > D_\varepsilon} \theta^2_k \right) < b_{D_\varepsilon}^2 \left( r_\varepsilon^2 - C_{\min}(\alpha, \beta)\varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \right).
\]
Since \( \mu_\varepsilon = b_{D_\varepsilon} r_\varepsilon \) and \( b_{D_\varepsilon}^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \geq \frac{C'_{\max}(\alpha, \beta)}{C_{\min}(\alpha, \beta)} \sqrt{D_\varepsilon} \), we get from (17)
\[
\sum_{k > D_\varepsilon} b_k^2 \theta^2_k < \mu_\varepsilon^2 - C'_{\max}(\alpha, \beta)\varepsilon^2 \sqrt{D_\varepsilon},
\]
which exactly means that \( \theta \in \mathcal{G}_{\mu,D}(C'_{\max}(\alpha, \beta)) \).

6.5. Proof of Proposition 4.2

We consider the case, where \( b_k \sim k^{-t} \) for any \( k \in \mathbb{N}^* \). One has
\[
r_\varepsilon^2 \sim \varepsilon^2 \sqrt{\sum_{k=1}^{D_\varepsilon} b_k^{-4}} \sim \varepsilon^2 D_\varepsilon^{(1+4t)/2} \sim D_\varepsilon^{-(1+4(s+t))/2} D_\varepsilon^{(1+4t)/2} = D_\varepsilon^{-2s}.
\]
We are then in the limit case. According to the value of \( C \), \( \mathcal{F}_{r,D}(C) \) may be empty or may be as follows:
\[
\mathcal{F}_{r,D}(C) := \left\{ \theta \in l_2(\mathbb{N}^*) : \forall \varepsilon \in (0, 1), \sum_{k > D_\varepsilon} \theta^2_k \leq CD_\varepsilon^{-2s} \right\} = \left\{ \theta \in l_2(\mathbb{N}^*) : \sup_{K \in \mathbb{N}^*} K^{-2s} \sum_{k > K} \theta^2_k \leq C \right\},
\]
which is a Besov space. Note that since the maxiset is only between the two sets \( \mathcal{F}_{r,D}(C_1) \) and \( \mathcal{F}_{r,D}(C_2) \) with \( C_1 \) and \( C_2 \) unknown constants, we cannot give an equality concerning the maxiset. Using similar computations, we get that
\[
\mathcal{G}_{\mu,D}(C) = \left\{ \theta \in l_2(\mathbb{N}^*) : \sup_{K \in \mathbb{N}^*} K^{2(s+t)} \sum_{k > K} \theta^2_k \leq C \right\}.
\]
Our aim below is to exhibit a sequence \( \theta \) such that \( \theta \in \mathcal{G}_{\mu,D}(C) \) for some \( C > 0 \) but \( \theta \notin \mathcal{F}_{r,D}(C') \), whatever the value of the constant \( C' \). To this end, let us consider the sequence \( \theta = (\theta_k)_{k \in \mathbb{N}^*} \) such that
\[
\theta_k = \frac{2^{-js}}{\sqrt{k}} \quad \forall k \in \{2^j, \ldots, 2^{j+1} - 1\}, \quad j \in \mathbb{N}^*, \quad \theta_k = 0 \text{ otherwise}.
\]
Let \( K \in \mathbb{N}^* \) be fixed and \( j \) be such that \( 2^{j-1} \leq K \leq 2^j \). We can check that
\[
\sum_{k > K} b_k^2 \theta^2_k \leq \sum_{k > K} k^{-(2t+1)2^{-2js}} \sim 2^{-2js} K^{-2t} \sim K^{-2(s+t)}.
\]
Hence \( \theta \in \mathcal{G}_{\mu,D}(C) \) for some \( C > 0 \).

At the same time, since \( K \leq 2^j \),
\[
\sum_{k > K} \theta^2_k \geq \sum_{k = 2^j}^{2^{j+1} - 1} \frac{2^{-2js}}{k} \sim 2^{-2js} \log(2^j) \sim 2^{-2js} j \sim \log(K) K^{-2s}.
\]
So \( \theta \notin \mathcal{F}_{r,D}(C') \), whatever \( C' > 0 \).
REFERENCES

1. J. Aston, F. Autin, G. Claeskens, J.-M. Freyermuth, and Ch. Pouet, “Minimax Optimal Procedures for Testing the Structure of Multidimensional Functions”, Appl. Comput. Harmonic Anal. 46 (2), 288–311 (2019).
2. F. Autin, “Maxiset for Density Estimation on $\mathbb{R}$”, Math. Methods Statist. 15 (2), 123–145 (2006).
3. F. Autin, J.-M. Freyermuth, and R. von Sachs, “Combining Thresholding Rules: A New Way to Improve the Performance of Wavelet Estimators”, J. Nonparametric Statist. 24 (4), 905–922 (2012).
4. Y. Baraud, “Non-Asymptotic Minimax Rates of Testing in Signal Detection”, Bernoulli 8 (5), 577–606 (2002).
5. C. Butucea, “Goodness-of-Fit Testing and Quadratic Functional Estimation from Indirect Observations”, Ann. Statist. 35 (5), 1907–1930 (2007).
6. L. Cavalier, Inverse Problems in Statistics, in Lecture Notes in Statistics, Vol. 203: Inverse Problems and High-Dimensional Estimation (Springer, Heidelberg, 2011), pp. 3–96.
7. M. Ermakov, “On Maxispace of Nonparametric Tests”, arXiv:1708.04985 (2018).
8. T. Hohage and F. Weidling, “Characterizations of Variational Source Conditions, Converse Results, and Maxisets of Spectral Regularization Methods”, SIAM J. Numer. Anal. 55 (2), 598–620 (2017).
9. Y. Ingster and I. Suslina, Nonparametric Goodness-of-Fit Testing under Gaussian Models, in Springer Science & Business Media (Springer, 2003), Vol. 169.
10. Y. I. Ingster, T. Sapatinas, and I. A. Suslina, “Minimax Signal Detection in Ill-Posed Inverse Problems”, Ann. Statist. 40 (3), 1524–1549 (2012).
11. J. Johannes, S. Van Bellegem, and A. Vanhems, “Convergence Rates for Ill-Posed Inverse Problems with an Unknown Operator”, Econometric Theory 27 (3), 522–545 (2011).
12. G. Kerkyacharian and D. Picard, “Minimax or Maxisets?”, Bernoulli 8 (2), 219–253 (2002).
13. C. Lacour and T. M. Pham Ngoc, “Goodness-of-Fit Test for Noisy Directional Data”, Bernoulli 20 (4), 2131–2168 (2014).
14. 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).
15. B. Laurent, J.-M. Loubes, and C. Marteau, “Non-Asymptotic Minimax Rates of Testing in Signal Detection with Heterogeneous Variances”, Electron. J. Statist. 6, 91–122 (2012).
16. V. Rivoirard and K. Tribouley, “The Maxiset Point of View for Estimating Integrated Quadratic Functionals”, Statist. Sinica 18 (1), 255–279 (2008).
17. A. V. Tsybakov, Introduction to Nonparametric Estimation, in Springer Series in Statistics (Springer, New York, 2009), Revised and extended from the 2004 French original, translated by Vladimir Zaiats.