NON ASYMPTOTIC MINIMAX RATES OF TESTING IN SIGNAL DETECTION

YANNICK BARAUD

ABSTRACT. Let $Y = (Y_i)_{i \in I}$ be a finite or countable sequence of independent Gaussian random variables of mean $f = (f_i)_{i \in I}$ and common variance σ^2. For various sets $\mathcal{F} \subset \ell_2(I)$, the aim of this paper is to describe the minimal ℓ_2-distance between f and 0 for the problem of testing "$f = 0$" against "$f \neq 0$, $f \in \mathcal{F}$" to be possible with prescribed probabilities of error. To do so, we start with the set $\mathcal{F}$ which collects the sequences f such that $f_j = 0$ for $j > n$ and $|\{j, \ f_j \neq 0\}| \leq k$ where the numbers k and n are integers satisfying $1 \leq k \leq n$. Then we show how such a result allows to handle the cases where $\mathcal{F}$ is an ellipsoid and more generally an ℓ_p-body with $p \in]0, 2]$. Our results are not asymptotic in the sense that we do not assume that σ tends to 0. Finally, we consider the problem of adaptive testing.

1. INTRODUCTION

We consider the following statistical model

$$Y_i = f_i + \sigma \varepsilon_i, \quad i \in I \tag{1}$$

where $f = (f_i)_{i \in I}$ is an unknown sequence of real numbers (called the signal), σ a positive number and the ε_i's a sequence of i.i.d. standard Gaussian random variables. Throughout this paper, I either denotes the set $\{1, ..., N\}$ (for some integer $N \geq 1$) or $\mathbb{N}^* = \mathbb{N} \backslash \{0\}$, the notation I being useful to handle both the Gaussian regression model and the Gaussian sequence model simultaneously. The observations are given by the sequence of Gaussian random variables $Y = (Y_i)_{i \in I}$, their joint law is denoted by P_f.

Let $\mathcal{F}$ be some subset of the Hilbert space

$$\ell_2(I) = \left\{ f \in \mathbb{R}^I, \ \|f\|^2 = \sum_{i \in I} f_i^2 < +\infty \right\}.$$

The aim of the paper is to describe the minimal radius ρ for which the problem of testing "$f = 0$" against the alternative "$f \in \mathcal{F}$ and $\|f\| \geq \rho$" with prescribed probabilities of errors is possible .

More precisely, let us fix some level $\alpha \in]0, 1[$ and consider some level-α test ϕ_α with values in $\{0, 1\}$ to test "$f = 0$" against "$f \in \mathcal{F} \setminus \{0\}$" (we decide to reject the null hypothesis when $\phi_\alpha(Y) = 1$). The test ϕ_α is powerful if it rejects the null hypothesis for all $f \in \mathcal{F}$ lying outside a small ball (the smaller the better) around 0 with probability close to 1. Then, given some $\delta \in]0, 1[$ (typically small) it is natural to measure the performance of the test via the quantity $\rho(\phi_\alpha, \mathcal{F}, \delta, \sigma)$ defined by

$$
\begin{aligned}
\rho(\phi_\alpha, \mathcal{F}, \delta, \sigma) &= \inf \left\{ \rho > 0, \ \inf_{f \in \mathcal{F}, \|f\| \geq \rho} P_f \left[\phi_\alpha = 1 \right] \geq 1 - \delta \right\} \\
&= \inf \left\{ \rho > 0, \ \sup_{f \in \mathcal{F}, \|f\| \geq \rho} P_f \left[\phi_\alpha = 0 \right] \leq \delta \right\}.
\end{aligned}
$$

The aim of this paper is to describe the quantity

$$\inf_{\phi_\alpha} \rho(\phi_\alpha, \mathcal{F}, \delta, \sigma) = \rho(\mathcal{F}, \alpha, \delta, \sigma), \tag{2}$$

the infimum being taken over all the level-α tests. In the sequel we shall call this quantity the (α, δ)-minimax rate of testing over $\mathcal{F}$ (or the minimax separation rate), the word "rate" referring to the scale parameter σ which is meant to decrease towards 0 when one considers the asymptotic point of view.

It is beyond the scope of this paper to give an exhaustive review of the literature on the problem of hypothesis testing. We refer for further details to the series of papers due to Ingster (1993a,b,c) which represent a landmark in the problem of finding minimax rates of testing over non parametric alternatives. In the Gaussian white noise model, the case of ellipsoids was first considered in Ermakov (1991) where exact minimax rates of testing is stated under assumptions on the semi-axes of the ellipsoids. Other kinds of alternatives are considered in Ingster (1993a,b,c) including Hölderian functional spaces, ellipsoids in $\mathbb{L}_2$ and other function spaces... . Lepski & Spokoiny (1999) obtain minimax rates of testing over Besov bodies $\mathcal{B}_{s,p,q}(R)$ with $p \in]0, 2[$ (see also Ingster & Suslina (1998)) and show an unexpected dependence (with regard to the case $p = 2$) of the minimax rate of testing with respect

to s. Spokoiny (1996) considers the problem of finding adaptive tests and shows that adaptation is impossible without some loss of efficiency (see also Ingster (1998)). In other words, it is not possible to find a test which achieves the minimax rate of testing (up to a universal constant) simultaneously over non trivial collections of Besov bodies.

A common feature of those results is their asymptotic character. In this paper we give non asymptotic results, mainly focusing on the problem of finding sharp lower bounds for the minimax rate of testing. However, asymptotic (upper) and lower bounds for the quantity $\rho(\mathcal{F}, \alpha, \delta, \sigma)$ can be deduced from our result by making σ tend to 0. In the regression framework, it is convenient to set $\sigma = 1/\sqrt{N}$ in order to obtain separation rate with respect to $\| \ \|_N = \| \ \|/\sqrt{N}$. The asymptotics are then obtained by letting N grow towards infinity as usual.

This paper was originally motivated by the following question: in the regression framework, what is the minimax rate for testing 0 against the class of signals which have their components equal to 0, except at most D of them? This situation corresponds to the reception of a sparse signal (at least $N - D$ components of the signal are 0 with D/N small), the problem being to determine some lower bound on the signal energy, $\|f\|^2$, for the detection to be possible with probability close to 1 and the probability of false alarm close to 0.

In Section 2, we give a partial answer to this question (a lower bound and an upper bound on the minimax rate of testing which are equal up to a possible $\ln(N)$ factor). An interesting feature of the result is that, for suitable values of D, the minimax rate of testing and the minimax rate of estimation are of the same order which is, as far as we know, seldom the case.

Another particular feature of this result is that it allows to derive non asymptotic lower bounds for the minimax rates of testing over ellipsoids and more generally over ℓ_p-bodies (also called ellipsoids in ℓ_p). A similar approach was adopted by Birgé & Massart (1999) for the related problem of estimation. To our knowledge the statement of lower bounds for the minimax rate of testing over general ℓ_p-bodies (that is under no assumption on the decay of the semi-axes) is new.

These results allow to recover those first established by Ermakov for ellipsoids (relaxing thus the assumptions on the semi-axes) and by Lepski & Spokoiny for some Besov bodies $\mathcal{B}_{s,p,q}(R)$ with $s > 0$, $R > 0$, $p \in]0, 2]$, $q \geq p$, this set being related to ℓ_p-bodies with semi-axes of the form k^{-s}.

The paper is organized as follows. As already mentioned, Section 2 is devoted to the problem of the detection of a sparse signal. Non asymptotic upper and lower bounds for the minimax rates of testing over ellipsoids are given in Section 3, the more general case of ℓ_p-bodies (with $p \in]0, 2]$) being treated in Section 4. The case of Besov bodies is considered in Section 5. The problem of adaptive testing is considered in Section 6 and the proofs are postponed to the last sections.

To end this section we introduce some notations that will be repeatedly used along the paper.

For any $\mathcal{F} \subset \ell_2(I)$ and $\alpha \in]0, 1[$, we denote by $\beta(\mathcal{F})$ the quantity

$$\beta(\mathcal{F}) = \inf_{\phi_\alpha} \sup_{f \in \mathcal{F}} P_f \left[\phi_\alpha = 0\right],$$

the infimum being taken over all tests ϕ_α with values in $\{0, 1\}$ satisfying $P_0[\phi_\alpha = 1] \leq \alpha$. By convention $\beta(\mathcal{F}) = 0$ if $\mathcal{F} = \emptyset$. For $x, y \in \mathbb{R}$, we set

$$x \wedge y = \inf\{x, y\}, \quad x \vee y = \sup\{x, y\}, \quad \lceil x \rceil = \inf\{n \in \mathbb{N}, \ n \geq x\},$$

and for all integers n, k such that $0 \leq k \leq n$,

$$C_n^k = \frac{n!}{k!(n-k)!}.$$

Throughout this paper the numbers α and $\delta \in]0, 1 - \alpha]$ are fixed and in order to keep our formulas as short as possible, we set

$$\eta = 2(1 - \alpha - \delta) \quad \text{and} \quad \mathcal{L}(\eta) = \ln(1 + \eta^2) < \ln 5.$$

Lastly, $C, C', C''...$ denote constants that may vary from line to line.

2. Detecting non zero coordinates

2.1. **The problem at hand.** Let I be either $\{1, ..., N\}$ or $\mathbb{N}^*$ and let $(e_j)_{j \geq 1}$ be the orthonormal family of vectors of $\ell_2(I)$ defined by

$$(e_j)_i = 1 \text{ if } i = j \text{ and } (e_j)_i = 0 \text{ otherwise.} \tag{3}$$

When I is finite the space $\ell_2(I)$ is merely $\mathbb{R}^N$ and the e_j's the canonical basis. For each pair of integers (n, k) with $k \in \{1, ..., n\}$ ($n \leq N$ when $I = \{1, ..., N\}$), let $\mathcal{M}(k, n)$ be the class of all the subsets of $\{1, ..., n\}$ of cardinality k. Now for all $m \in \mathcal{M}(k, n)$ and $D \geq 1$, let us set

$$S_m = \text{span}\{e_j, \ j \in m\} \quad \text{and} \quad S_D = \text{span}\{e_j, \ j \in \{1, ..., D\}\},$$

where span(A) denotes the linear space generated by $A \subset \ell_2(I)$.

In this section we study the case where $\mathcal{F}$ is given by

$$\mathcal{F} = \bigcup_{m \in \mathcal{M}(k,n)} S_m, \tag{4}$$

2.2. **Lower bounds.** To start with, let us consider the elementary case where $n = k = D \geq 1$, that is when $\mathcal{F} = S_D$.

Proposition 1. *Let us set*

$$\rho_D^2 = \sqrt{2\mathcal{L}(\eta)D}\ \sigma^2, \tag{5}$$

then for all $\rho \leq \rho_D$,

$$\beta\left(\{f \in S_D,\ \|f\| = \rho\}\right) \geq \delta.$$

Comments: The result can be described in words in the following way: whatever the level-α test ϕ_α, there exists some signal $f \in S_D$ satisfying $\|f\| \geq \rho_D$ for which the error of second kind, $P_f[\phi_\alpha = 0]$, is at least δ. This implies the lower bound

$$\rho\left(S_D, \alpha, \delta, \sigma\right) \geq \rho_D,$$

the left-hand side of this inequality being defined by (2).

The Gaussian distribution being invariant under orthogonal transformations, the same result holds for $\mathcal{F}$ being any linear space of dimension D.

Let us now turn to the general case.

Theorem 1. *Let $\mathcal{F}$ be given by (4) and let us set*

$$\rho_{k,n}^2 \;=\; k\ln\left(1 + \mathcal{L}(\eta)\frac{n}{k^2} + \sqrt{2\mathcal{L}(\eta)\frac{n}{k^2} + \left(\mathcal{L}(\eta)\frac{n}{k^2}\right)^2}\right)\sigma^2. \tag{6}$$

Then for all $\rho \leq \rho_{k,n}$,

$$\beta\left(\{f \in \mathcal{F}, \|f\| = \rho\}\right) \geq \delta.$$

If $\alpha + \delta \leq 59\%$ then one has

$$\rho_{k,n}^2 \geq k\ln\left(1 + \frac{n}{k^2} \vee \sqrt{\frac{n}{k^2}}\right)\sigma^2. \tag{7}$$

2.3. **Upper bounds.** Let us now discuss the sharpness of the results stated in the previous section. For this aim we introduce some additional notations and define some special tests based on χ^2-statistics. For each finite subset m of $\mathbb{N}^*$ we set

$$\phi_{m,\alpha} = \mathbf{1} \left\{ \sum_{i \in m} Y_i^2 > t_{|m|,\alpha} \sigma^2 \right\} \tag{8}$$

where for each $d \in \mathbb{N}^*$, $t_{d,\alpha}$ satisfies

$$P[Z_d^2 > t_{d,\alpha}] = \alpha \text{ if } Z_d^2 \sim \chi^2(d). \tag{9}$$

Lastly, we denote by $\phi_{D,\alpha}$ the test defined by

$$\phi_{D,\alpha} = \phi_{\{1,\ldots,D\},\alpha} = \mathbf{1} \left\{ \sum_{i=1}^{D} Y_j^2 \geq t_{D,\alpha} \sigma^2 \right\}. \tag{10}$$

The following result holds.

Proposition 2. *Let $\mathcal{F}$ be defined by (4). The test ϕ_α^* defined by*

$$\phi_\alpha^* = \left[\sup_{m \in \mathcal{M}(k,n)} \phi_{m,\alpha/(2C_n^k)} \right] \vee \phi_{n,\alpha/2},$$

satisfies

$$P_0 \left[\phi_\alpha^* = 1 \right] \leq \alpha \quad \text{and} \quad P_f \left[\phi_\alpha^* = 0 \right] \leq \delta,$$

for all $f \in \mathcal{F}$ such that

$$\|f\|^2 \geq C_1 \left[\left(k \ln \left(e\frac{n}{k} \right) \right) \wedge \sqrt{n} \right] \sigma^2.$$

One can take $C_1 = 2(\sqrt{5} + 4) \ln(2e/(\alpha\delta))$.

Comments: The results of Theorem 1 and Proposition 2 show that (for reasonable values of α and δ) the quantity $\rho^2 = \rho^2(\mathcal{F}, \alpha, \delta, \sigma)$ satisfies

$$k \ln \left(1 + \frac{n}{k^2} \vee \sqrt{\frac{n}{k^2}} \right) \sigma^2 \leq \rho^2 \leq C_1 \left[\left(k \ln \left(e\frac{n}{k} \right) \right) \wedge \sqrt{n} \right] \sigma^2.$$

To analyze further these inequalities, we take $\sigma^2 = 1$ and distinguish between the values of k.

- When $k = n = D$,

we see that the lower and the upper bound are both of order $\sqrt{D}$, which shows that the result of Proposition 1 is sharp and that an optimal test is merely obtained by rejecting the null hypothesis when $\sum_{j=1}^{D} Y_j^2$ is large enough.

- When $k \leq n^{\gamma}$ for some $\gamma < 1/2$,

the lower and the upper bound are both of order $k \ln(n)$ (up to a constant depending on γ for the lower bound). This shows that the lower bound given in Theorem 1 is sharp and that the test ϕ_{α} is rate optimal. Since the minimax rate of estimation with respect to the quadratic loss function $\| \ \|^2$ over $\mathcal{F}$ is of order $k \ln(en/k)$ (see Birgé & Massart (1999), Theorem 3) we note that in this case the squared minimax separation rate and the minimax estimation rate over $\mathcal{F}$ are both of the same order.

- When $\sqrt{n} \leq k < n$,

the lower and the upper bound do not depend on k any longer and are both of order $\sqrt{n}$. Here again, the lower bound stated in Theorem 1 is sharp and the test ϕ_{α} rate optimal. The fact that the separation rate stabilizes around $\sqrt{n}$ for $k > \sqrt{n}$ contrasts with the estimation problem for which the estimation rate keeps growing almost linearly with respect to k as k becomes large.This phenomenon is due to the fact that for the problem of hypothesis testing we benefit from the prior assumption that f belongs to S_n, the squared rate of testing over S_n being of order $\sqrt{n}$. Consequently, in the regression framework (by taking $S_N = \mathbb{R}^N$) rates of testing are always better than $\sqrt{N}$ (up to a constant). We shall meet this phenomenon again but not mention it any more.

- When $k < \sqrt{n}$ and k is close to $\sqrt{n}$,

the lower and the upper bound differ from at most a $\ln(n)$ factor. For example when k is of order $\sqrt{n}/\ln(n)$, the lower bound presented in Theorem 1 is of order $\sqrt{n} \ln \ln(n)/\ln(n)$, the upper bound being of order $\sqrt{n}$. We conjecture that the lower bound is sharp and do not know whether the preceding testing procedure is suboptimal or not.

Finally, let us emphasize the gap (in terms of rates of testing) between the situation where the location of the non zero components of the signal is known (the squared rate is of order $\sqrt{D}$) and where the location is unknown (then the squared rate is at least D). This difference is worth mentioning since for the estimation problem the corresponding minimax rates differ only from (at most) a $\ln(n)$ factor.

3. Minimax rate of testing over an ellipsoid

In this section we assume that $\mathcal{F}$ is an ellipsoid, that is of the form

$$\mathcal{E}_{a,2}(R) = \left\{ f \in \ell_2(I), \ \sum_{k \in I} \frac{f_k^2}{a_k^2} \leq R^2 \right\},$$

where R denotes a positive number and the a_j's a non increasing sequence of positive numbers such that $a_1 = 1$ and $\lim_{k \to +\infty} a_k = 0$ when $I = \mathbb{N}^*$. The case of ℓ_p-bodies, which is an extension to the case $p \neq 2$, will be considered in the next section.

3.1. **Lower bounds.** The following holds

Proposition 3. *Let us set*

$$\rho_{a,2,R}^2 = \sup_{D \in I} \left[\rho_D^2 \wedge (R^2 a_D^2) \right], \tag{11}$$

where ρ_D is defined by (5). Then we have

$$\beta\left(\{f \in \mathcal{E}_{a,2}(R), \ \|f\| \geq \rho_{a,2,R}\}\right) \geq \delta.$$

If $\alpha + \delta \leq 59\%$ then

$$\rho_{a,2,R}^2 \geq \sup_{D \in I} \left[(\sqrt{D}\sigma^2) \wedge (R^2 a_D^2) \right].$$

Proof. We use the notations introduced at the beginning of Section 2. We set $\mathcal{F} = \mathcal{E}_{a,2}(R)$ and for each $D \in I$, $r_D^2 = \rho_D^2 \wedge (R^2 a_D^2)$. Let us fix some $D \in I$. Since the a_j's are non increasing and $r_D^2 \leq R^2 a_D^2$, $\sum_{j=1}^{D} f_j^2/a_j^2 \leq R^2$ for all $f \in S_D$ such that $\|f\| = r_D$. This shows the inclusion

$$\{f \in S_D, \ \|f\| = r_D\} \subset \{f \in \mathcal{F}, \ \|f\| \geq r_D\}.$$

Now, since $r_D \leq \rho_D$ we deduce from Proposition 1 that

$$\beta\left(\{f \in \mathcal{E}_{a,2}(R), \ \|f\| \geq r_D\}\right) \geq \delta,$$

and the result of Proposition 3 follows since D is arbitrary in I. $\qquad\square$

3.2. **Optimality of the lower bounds.** In this section we show that the result of Proposition 3 is sharp. To this aim let us introduce the quantity D^* defined by

$$D^* = \inf\left\{ D \in I, \ R^2 a_D^2 \leq \sqrt{D}\sigma^2 \right\},$$

with the convention that $\inf \emptyset = N$. The following result holds

Proposition 4. *If $\sigma < R$, the test ϕ_α^* defined by $\phi_\alpha^* = \phi_{D^*,\alpha}$ where $\phi_{D^*,\alpha}$ is given by (10), satisfies*

$$P_0[\phi_\alpha^* = 1] \leq \alpha \quad \text{and} \quad P_f[\phi_\alpha^* = 0] \leq \delta,$$

for all $f \in \mathcal{E}_{a,2}(R)$ such that

$$\|f\|^2 \geq C_1 \sup_{D \in I} \left[(\sqrt{D}\sigma^2) \wedge (R^2 a_D^2) \right].$$

One can take $C_1 = \sqrt{2}[1 + 2(\sqrt{5} + 4)] \ln(1/(\alpha\delta))$.

Comment: This result and Proposition 3 show that the quantity $\rho_{a,2,R}$ is of the same order as the minimax rate of testing over $\mathcal{E}_{a,2}(R)$. Note that the quantity $\rho_{a,2,R}$ is obtained by finding the best trade-off over I between the two terms $R^2 a_D^2$ and ρ_D^2 (which is of order $\sqrt{D}\sigma^2$). The quantity Ra_D represents the maximal ℓ_2-distance of a point of $\mathcal{E}_{a,2}(R)$ to S_D. It is non increasing with respect to D. In contrast, the quantity ρ_D which is (up to a constant) the minimax rate of testing over S_D, is non decreasing with respect to D. The situation is very similar to the situation encountered in the estimation problem. Let us explain why. For the sake of simplicity let us assume that $I = \{1, ..., N\}$. For each $f \in \mathcal{E}_{a,2}(R)$, one can estimate f from the data $(Y_i)_{i \in I}$ thanks to the projection estimator onto S_D given by $\hat{f}_D = (Y_1, ..., Y_D, 0, ..., 0)'$. Since this estimator satisfies

$$\sup_{f \in \mathcal{E}_{a,2}(R)} \mathbb{E} \left[\|f - \hat{f}_D\|^2 \right] \leq R^2 a_D^2 + D\sigma^2,$$

one gets that for some value of $D = D_*$ suitably chosen to balance the bias term $R^2 a_D^2$ and the variance term $D\sigma^2$, the minimax risk on $\mathcal{E}_{a,2}(R)$ is bounded from above, up to an universal constant, by (actually some additional minor conditions should be added),

$$\sup_{D \in I} \left[(D\sigma^2) \wedge (R^2 a_D^2) \right].$$

This quantity turns out to be the minimax rate of estimation over the ellipsoid in various cases (see Birgé & Massart (1999)). Then, the analogy with the problem of testing becomes clear. It is worth mentioning that just as the estimator $\hat{f}_{D_*}$ is minimax (up to a constant) for the problem of estimation, the test based on the test statistic $\|\hat{f}_{D^*}\|^2$ is rate optimal for the problem of hypothesis testing. Yet, in general $D^* \neq D_*$, the choice of D^* being similar to that prescribed for the quadratic functional estimation problem by model selection (see Laurent & Massart (1998)).

Instead of considering the ellipsoid $\mathcal{E}_{a,2}(R)$ we could also have dealt with the larger set $\mathcal{E}'_{a,2}(R)$ defined by

$$\mathcal{E}'_{a,2}(R) = \{f \in \ell_2(I),\ \forall D \in I,\ \mathrm{d}(f, S_D) \leq Ra_D\},$$

where $\mathrm{d}(f, S_D)$ denotes the ℓ_2-distance between f and S_D. Then the lower and the upper bound for the separation rate would have been the same (it is enough to see that the proof of Proposition 4 remains unchanged when replacing $\mathcal{E}_{a,2}(R)$ by $\mathcal{E}'_{a,2}(R)$). Of course in the regression framework, via some orthogonal transformation the same result holds when replacing the nested collection of linear spaces $(S_D)_{D=1,\ldots,N}$ by any other. Lastly, let us mention that the result easily extends to sets of the form

$$\{f \in \ell_2(I),\ \forall D \in I',\ \mathrm{d}(f, S_D) \leq Ra_D\},$$

with $I' \subset I$, by noticing that

$$\{f \in \ell_2(I),\ \forall D \in I',\ \mathrm{d}(f, S_D) \leq Ra_D\} = \mathcal{E}'_{a,2}(R)$$

when one defines the a_D's for $D \in I \setminus I'$ by the formula

$$a_D = \inf\{a_k, k \in I' \cap \{1, \ldots, D\}\}.$$

Moreover, it is easy to check that one has

$$\rho_{a,2,R}^2 = \sup_{D \in I'} \left[\rho_D^2 \wedge (R^2 a_D^2)\right].$$

The proof of Proposition 4 is deferred to Section 8.

4. Minimax rates of testing over an ℓ_p-body with $0 < p < 2$

In this section we consider the case where $\mathcal{F}$ is an ℓ_p-body, that is of the form

$$\mathcal{E}_{a,p}(R) = \left\{f \in \ell_2(I),\ \sum_{k \in I} \left|\frac{f_k}{a_k}\right|^p \leq R^p\right\},$$

where R and p denote some positive numbers and $a = (a_k)_{k \in I}$ some non increasing sequence such that $a_1 = 1$ and $\lim_{k \to +\infty} a_k = 0$ when $I = \mathbb{N}^*$. The case $p = 2$ has already been considered in the previous section.

4.1. **Lower bounds.** The following result holds.

Proposition 5. *Let us set*

$$\rho_{a,p,R}^2 \;=\; \sup_{D \in I} \left[\rho_{\lceil \sqrt{D} \rceil, D}^2 \wedge \left(R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p} \right) \right],$$

where $\rho_{\lceil \sqrt{D} \rceil, D}^2$ is defined by (6). Then we have that

$$\beta \left(\{ f \in \mathcal{E}_{a,p}(R), \ \|f\| \geq \rho_{a,p,R} \} \right) \geq \delta.$$

If $\alpha + \delta \leq 29\%$ then

$$\rho_{a,p,R}^2 \geq \sup_{D \in I} \left[\left(\lceil \sqrt{D} \rceil \sigma^2 \right) \wedge \left(R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p} \right) \right].$$

Comment: As for the case $p = 2$, we see that the lower bound derives from some best trade-off between two terms, this trade-off being realized for some D^* satisfying (roughly speaking)

$$\sqrt{D^*} = R^p a_{D^*}^p / \sigma^p.$$

For the sake of the simplicity of the forthcoming comment, we assume that $D^* \in I$. As $\sqrt{D}\sigma^2$ and $R^2 a_D^2 \sigma^{2-p}$ are also of the same order for the same value of D^*, we also have that $\rho_{\lceil \sqrt{D} \rceil, D}^2$ is of order

$$\sup_{D \in I} \left[\left(\sqrt{D}\sigma^2 \right) \wedge \left(R^p a_D^p \sigma^{2-p} \right) \right].$$

In the light of the related result obtained for $p = 2$, the last lower bound turns out to be more tractable to comment. Indeed on the one hand we recognize the quantity $\sqrt{D}\sigma^2$ which is of the same order as the minimax rate of testing over S_D. On the other hand, the quantity $R^p a_D^p \sigma^{2-p}$ can be interpreted as a "bias" term since it is the maximal distance to S_D of a point belonging to the set

$$\mathcal{E}_{a,p}(R) \cap \left\{ f \in \ell_2(I), \ \max_{i \in I} |f_i| \leq \sigma \right\}.$$

In other words, we use the linear space S_D to approximate the signals of the ℓ_p-body belonging to some hypercube.

Proof. We use the notations introduced in Section 2, set $\mathcal{F} = \mathcal{E}_{a,p}(R)$ and for each $D \in I$, $r_D^2 = \rho_{\lceil \sqrt{D} \rceil, D}^2 \wedge \left(R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p} \right)$. Let us now fix $D \in I$. For all $m \in \mathcal{M}(\lceil \sqrt{D} \rceil, D)$ and $f \in S_m \subset S_D$ such that

$\|f\| = r_D$, we have by Hölder's inequality

$$\sum_{j \in I} \left| \frac{f_j}{a_j} \right|^p = \sum_{j \in m} \left| \frac{f_j}{a_j} \right|^p \le |m|^{1-p/2} \left(\sum_{j \in m} \frac{f_j^2}{a_j^2} \right)^{p/2}$$

$$\le \frac{r_D^p \lceil \sqrt{D} \rceil^{1-p/2}}{a_D^p} \le R^p, \tag{12}$$

using that $r_D^2 \le R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p}$. We deduce from (12) the inclusion

$$\left\{ f \in \bigcup_{m \in \mathcal{M}(\lceil \sqrt{D} \rceil, D)} S_m, \ \|f\| = r_D \right\} \subset \{f \in \mathcal{F}, \|f\| = r_D\},$$

and as $r_D \le \rho_{\lceil \sqrt{D} \rceil, D}$, we derive from Theorem 1 that

$$\beta \left(\{ f \in \mathcal{E}_{a,p}(R), \ \|f\| \ge r_D \} \right) \ge \delta.$$

The result follows since D is arbitrary in I. To finish the proof of Proposition 5, it remains to check that $\rho_{\lceil \sqrt{D} \rceil, D}^2 \ge \lceil \sqrt{D} \rceil$ when $\alpha + \delta \le$ 29%. Since for $D \ge 1$, $D/\lceil \sqrt{D} \rceil^2 \ge 1/2$, we deduce from (6) that

$$\rho_{\lceil \sqrt{D} \rceil, D}^2 \ge \ln(1 + \mathcal{L}(\eta)/2 + \sqrt{\mathcal{L}(\eta) + \mathcal{L}(\eta)^2/4})) \lceil \sqrt{D} \rceil,$$

and the result follows since for $\alpha + \delta \le 29\%$,

$$\ln(1 + \mathcal{L}(\eta)/2 + \sqrt{\mathcal{L}(\eta) + \mathcal{L}(\eta)^2/4})) \ge 1.$$

$\square$

4.2. **Upper bounds.** Let us define D^* by

$$D^* = \inf \left\{ D \in I, \ R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p} \le \lceil \sqrt{D} \rceil \sigma^2 \right\},$$

with the convention, $\inf \emptyset = N$ and

$$\phi_{loc,\alpha/2} = \sup_{j > D^*, j \in I} \phi_{\{j\}, 2\alpha/(\pi^2(j-D^*)^2)},$$

where the tests $\phi_{\{j\}, 2\alpha/(\pi^2(j-D^*)^2)}$ are given by (8). Let us now set

$$\varrho_{a,p,R}^2 = \sup_{D \in I} \left[\left(\lceil \sqrt{D} \rceil \sigma^2 \right) \wedge \left(R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p} \right) \right].$$

The first result considers the case of the regression framework.

Proposition 6. *Assume that $I = \{1, ..., N\}$ and that $\sigma < R$. Let us define the test ϕ_α^* by*

$$\phi_\alpha^* = \phi_{loc,\alpha/2} \vee \phi_{D^*,\alpha/2}. \tag{13}$$

The test ϕ_α^* satisfies

$$P_0[\phi_\alpha^* = 1] \leq \alpha \quad \text{and} \quad P_f[\phi_\alpha^* = 0] \leq \delta, \tag{14}$$

for all $f \in \mathcal{E}_{a,p}(R)$ such that

$$\|f\|^2 \geq C \left(\ln(2 + N)\right)^{1-p/2} \varrho_{a,p,R}^2. \tag{15}$$

One can take $C = 8(\sqrt{5} + 4)\ln(e\pi/(\alpha\delta))$.

Comment: This result shows that in the regression framework the rate $\rho_{a,p,R}^2$ is optimal up to a possible $\ln(N)$ factor. Note that the test presented above actually mixes several tests. The presence of local tests, namely the $\phi_{\{j\},2\alpha/(\pi^2(j-D^*)^2)}$'s, allows to reject the null hypothesis when one value of the $|Y_j|$ is large enough.

The next Proposition shows that the rate $\rho_{a,p,R}^2$ is optimal under the following (restrictive) condition:

(H) The sequence $(\theta_j)_{j \in I}$ defined by

$$\theta_j = \sup_{j' \in I, j+j' \in I} \frac{a_{j+j'}}{a_{j'}}$$

satisfies

$$\Sigma = \sum_{j \in I} \theta_j^p \ln(2 + j)^{1-p/2} < +\infty.$$

Proposition 7. *Assume $\sigma < R$ and that (H) holds. The test ϕ_α^* defined by (13) satisfies*

$$P_0[\phi_\alpha^* = 1] \leq \alpha \quad \text{and} \quad P_f[\phi_\alpha^* = 0] \leq \delta,$$

for all $f \in \mathcal{E}_{a,p}(R)$ such that

$$\|f\|^2 \geq C' \varrho_{a,p,R}^2. \tag{16}$$

One can take $C' = (\Sigma \vee 1)8(\sqrt{5} + 4)\ln(e\pi/(\alpha\delta))$.

Comment: Condition **(H)** is fulfilled when, for example, the a_j's are of the form $\theta e^{-\lambda j}$ for some $\lambda, \theta > 0$. Unfortunately, when the a_k's are of the form k^{-s} for some $s > 0$, Condition **(H)** is not fulfilled. Yet, in this case the lower bound obtained in Proposition 5 is known to be sharp as we shall see in the next section.

The proofs of Proposition 6 and 7 are deferred to Section 8.

5. Besov bodies

This section is devoted to the statement of lower bounds for the minimax rate of testing over Besov bodies. Let us first recall what a Besov body is (as introduced by Donoho & Johnstone (1998)). In the sequel $I = \mathbb{N}^*$. Let $R > 0$, $p > 0$, $q \in]0, +\infty]$ and $s' > (1/p - 1/2)_+$. Setting $s = s' - (1/p - 1/2)_+$ we define the Besov body $\mathcal{B}_{s',p,q}(R)$ by

$$\mathcal{B}_{s',p,q}(R) = \left\{ f \in \ell_2(I), \ \sum_{j \geq 0} \left[2^{js} \left(\sum_{k=2^j}^{2^{j+1}-1} |f_k|^p \right)^{1/p} \right]^q \leq R^q \right\},$$

when $q < +\infty$ and

$$\mathcal{B}_{s',p,\infty}(R) = \left\{ f \in \ell_2(I), \ \sup_{j \geq 0} 2^{js} \left(\sum_{k=2^j}^{2^{j+1}-1} |f_k|^p \right)^{1/p} \leq R \right\}.$$

Clearly, when $p \leq q$ the inclusion $\mathcal{B}_{s',p,p}(R) \subset \mathcal{B}_{s',p,q}(R)$ holds.

5.1. **From Besov to ℓ_p-bodies.** Originally the Gaussian white noise model was the statistical framework chosen to study the problem of minimax hypothesis testing (we have already mentioned the work of Ingster, Lepski & Spokoiny ...). The use of a suitable wavelet basis allows to translate both the problem at hand, from the Gaussian white noise model to the Gaussian sequence model, and the property that the function belongs to some usual functional space (such as a Besov space) to the property that the sequence of its coefficients onto the wavelet basis belongs to some related sequence space (namely, a Besov body). This translation is described in Spokoiny (1996). In order to make further connections of our results with previous works, we now establish some connections between Besov and ℓ_p-bodies.

Proposition 8. *For all $s, p > 0$, let us denote by $\mathcal{E}_{s,p}(R)$ the ℓ_p-body defined by*

$$\mathcal{E}_{s,p}(R) = \left\{ f \in \ell_2(I), \ \sum_{k \in I} k^{ps} |f_k|^p \leq R^p \right\}.$$

We have

$$\mathcal{B}_{s',p,p}(2^{-s}R) \subset \mathcal{E}_{s,p}(R) \subset \mathcal{B}_{s',p,p}(R),$$

where $s' = s + (1/p - 1/2)_+$.

This proposition shows that from the minimax point of view, the ℓ_p-body $\mathcal{E}_{s,p}(R)$ and the Besov body $\mathcal{B}_{s',p,p}(R)$ behave essentially in the same way. In the next section we shall restrict our study to those ℓ_p-bodies. To keep our notation coherent we write $\rho_{s,2,R}$ for $\rho_{a,2,R}$ when the a_k's are of the form k^{-s}.

Proof. We have that

$$\sum_{j\geq 0} 2^{jps} \sum_{k=2^j}^{2^{j+1}-1} |f_k|^p \;\leq\; \sum_{j\geq 0} \sum_{k=2^j}^{2^{j+1}-1} k^{ps}|f_k|^p \leq \sum_{k\geq 1} k^{ps}|f_k|^p,$$

which shows that $\mathcal{E}_{s,p}(R) \subset \mathcal{B}_{s',p,p}(R)$. Conversely,

$$\sum_{k\geq 1} k^{ps}|f_k|^p \;=\; \sum_{j\geq 0} \sum_{k=2^j}^{2^{j+1}-1} k^{ps}|f_k|^p \leq 2^{ps} \sum_{j\geq 0} 2^{jps} \sum_{k=2^j}^{2^{j+1}-1} |f_k|^p$$

which shows that $\mathcal{B}_{s',p,p}(2^{-s}R) \subset \mathcal{E}_{s,p}(R)$. $\qquad\square$

5.2. **The result for $p = 2$.** The asymptotic version of this result is known from Ermakov (1991).

Corollary 1. *Let $s > 0$. Assume that $\sigma^2 < R^2$ and that $\alpha + \delta \leq 59\%$, then we have for $I = \mathbb{N}^*$,*

$$\rho_{s,2,R}^2 \geq 2^{-2s} R^{2/(1+4s)} \sigma^{8s/(1+4s)}, \tag{17}$$

and for $I = \{1, ..., N\}$,

$$\rho_{s,2,R}^2 \geq 2^{-2s} \left[\left(R^{2/(1+4s)} \sigma^{8s/(1+4s)} \right) \wedge \left(\sqrt{N} \sigma^2 \right) \right]. \tag{18}$$

Comment: From an asymptotic point of view, by taking $\sigma^2 = 1/N$ in the Gaussian regression model we obtain that the right-hand side of (18) is of order $N^{-4s/(1+4s)}$ if $s > 1/4$ and of order $1/\sqrt{N}$ otherwise.

Proof. Applying Proposition 3 we get

$$\rho_{s,2,R}^2 \geq \sup_{D \in I} \left[(\sqrt{D}\sigma^2) \wedge (R^2 D^{-2s}) \right].$$

For all $x > 0$, $\sqrt{x}\sigma^2 \geq R^2 x^{-2s}$ if and only if

$$x \geq \left(R^2/\sigma^2 \right)^{2/(1+4s)} = x^* \geq 1.$$

If $D^* = \lceil x^* \rceil$ belongs to I, then $x^* \leq D^* \leq x^* + 1 \leq 2x^*$ and we get that

$$\rho_{s,2,R}^2 \;\geq\; R^2 (D^*)^{-2s} \geq 2^{-2s} R^2 (x^*)^{-2s} = 2^{-2s} R^{2/(1+4s)} \sigma^{8s/(1+4s)}.$$

If $D^* \notin I$, then $I = \{1, ..., N\}$ and $N < x^*$ which implies that

$$\rho_{s,2,R}^2 \geq \sqrt{N}\sigma^2.$$

$\square$

5.3. The result for $p < 2$.
The rates given below are optimal according to the results by Spokoiny (1996) on the related Besov bodies.

Corollary 2. *Let $s > 0$ and set $s'' = s - 1/4 + 1/(2p)$. Assume that $\sigma^2 < R^2$ and that $\alpha + \delta \leq 29\%$, then we have for $I = \mathbb{N}^*$,*

$$\rho_{s,p,R}^2 \geq 2^{-4s''} R^{2/(1+4s'')} \sigma^{8s''/(1+4s'')}, \tag{19}$$

and for $I = \{1, ..., N\}$,

$$\rho_{s,p,R}^2 \geq 2^{-4s''} \left[\left(R^{2/(1+4s'')} \sigma^{8s''/(1+4s'')} \right) \wedge \left(\sqrt{N}\sigma^2 \right) \right]. \tag{20}$$

Comment: From an asymptotic point of view, by taking $\sigma^2 = 1/N$ we obtain that the right-hand side of (20) is of order $N^{-4s''/(1+4s'')}$ when $s \geq 1/2 - 1/(2p)$.

Proof. Applying Proposition 5 we get that

$$
\begin{aligned}
\rho_{s,p,R}^2 &\geq \left[\sup_{D \in I} \left(R^2 D^{-2s} \lceil \sqrt{D} \rceil^{1-2/p} \right) \wedge \left(\lceil \sqrt{D} \rceil \sigma^2 \right) \right] \\
&\geq \left[\sup_{D \in I} \left(R^2 \lceil \sqrt{D} \rceil^{-4s''} \right) \wedge \left(\lceil \sqrt{D} \rceil \sigma^2 \right) \right].
\end{aligned}
$$

For $x > 0$, $x\sigma^2 \geq R^2 x^{-4s''}$ if and only if

$$x \geq (R^2/\sigma^2)^{1/(1+4s'')} = x^* \geq 1.$$

Let D^* be the smallest integer such that $\lceil \sqrt{D} \rceil \geq x^*$. One has $D^* \geq 1$ since $x^* \geq 1$ and $\lceil \sqrt{D^*} \rceil \leq x^* + 1 \leq 2x^*$ since

$$\lceil \sqrt{D^*} \rceil - 1 = \lceil \sqrt{D^*} - 1 \rceil \leq \lceil \sqrt{D^* - 1} \rceil \leq x^*.$$

If $D^* \in I$, then

$$\rho_{s,p,R}^2 \geq R^2 \lceil \sqrt{D^*} \rceil^{-4s''} \geq 2^{-4s''} R^{2/(1+4s'')} \sigma^{8s''/(1+4s'')}$$

Otherwise, $I = \{1, ..., N\}$ and $\lceil \sqrt{N} \rceil < x^*$ which implies that

$$\rho_{s,p,R}^2 \geq \lceil \sqrt{N} \rceil \sigma^2,$$

which completes the proof of Corollary 2.

$\square$

6. Simultaneous rates of testing

6.1. Coming back to the problem of detecting non zero coordinates.

In this section we come back, for a short time, to the problem of detecting non zero coordinates. In order to explain the problem let us introduce some notations. Let $(I_j)_{j \in \mathcal{J}}$ be some finite or countable family of finite disjoint subsets of I. For each $j \in \mathcal{J}$ let $n(j) = |I_j|$ and $k(j) \in \{1, ..., n(j)\}$. Now we set

$$\mathcal{M}_j = \{m \subset I_j, \ |m| = k(j)\}, \quad \mathcal{F}_j = \bigcup_{m \in \mathcal{M}_j} S_m$$

and

$$\tilde{\rho}_j = \begin{cases} \rho_{n(j)} & \text{defined by (5) when } k(j) = n(j); \\ \rho_{k(j),n(j)} & \text{defined by (6) otherwise.} \end{cases}$$

We have seen in Section 2 that for each j, the quantity $\tilde{\rho}_j = \tilde{\rho}_j(\eta)$ is of the same order as the minimax separation rate over $\mathcal{F}_j$ (up to a possible $\ln(n(j))$ factor for some cases). From now on, the dependency of $\tilde{\rho}_j = \tilde{\rho}_j(\eta)$ with respect to η is emphasized.

The following result holds.

Proposition 9. *For any sequence of positive weights p_j such that*

$$\sum_{j \in \mathcal{J}} p_j \leq 1,$$

we have

$$\beta \left(\bigcup_{j \in \mathcal{J}} \{f \in \mathcal{F}_j \text{ and } \|f\| = r_j\} \right) \geq \delta,$$

as soon as for all $j \in \mathcal{J}$, $r_j \leq \tilde{\rho}_j(\eta/\sqrt{p_j})$.

The proof is postponed to Section 7.4.

Comment: Since the quantity $\tilde{\rho}_j(\eta/\sqrt{p_j})$ is of order $\tilde{\rho}_j$ times a power of $\ln(1/p_j)$, the result of Proposition 9 means that in the problem of testing 0 against this multiple alternative, a loss of efficiency over at least one of the alternatives is unavoidable. For example, when $|\mathcal{J}|$ is finite, by taking $p_j = 1/|\mathcal{J}|$ for all $j \in \mathcal{J}$ one derives that a loss of efficiency by a factor (a power of) $\ln(|\mathcal{J}|)$ over one of the $\mathcal{F}_j$'s is unavoidable. From an asymptotic point of view this phenomenon is worth mentioning when the cardinality of $\mathcal{J}$ depends upon σ (or N in the regression framework). Let us also mention that the loss of efficiency may not affect all the the alternatives (this fact is seldom

emphasized in the literature), we refer for further details to the work of Baraud, Huet, Laurent (1999) in the regression framework.

In the sequel we derive some lower bounds for the problem testing "$f = 0$" against a multiple alternative such as a collection of nested linear spaces or a collection of nested ellipsoids. Extensions to more geneneral ℓ_p-bodies possible to the price of more technicalities.

6.2. **The case of nested linear spaces.** We shall restrict our study to the case of the linear spaces S_D's defined at the beginning of Section 2. However, when $I = \{1, ..., N\}$ this result holds true for any (substantial) nested collection of linear subspaces of $\mathbb{R}^N$.

Corollary 3. *Let us set*

$$\bar{\rho}_D^2 = C\sqrt{\ln\ln(D+1)}\sqrt{D}\sigma^2. \tag{21}$$

with $C = \sqrt{2}\left[(\eta\pi/\sqrt{6}) \wedge 1\right]$. We have that

$$\beta\left(\bigcup_{D \in I} \{f \in S_D, \ \|f\| = r_D\}\right) \geq \delta,$$

if for all $D \in I$, $r_D \leq \bar{\rho}_D$.

Proof. We take $\sigma^2 = 1$. For all $j \geq 0$ such that $2^{j+1} - 1 \in I$ (i.e. for all $j \leq J$ with $J = +\infty$ if $I = \mathbb{N}^*$, $J = J(N) = \ln(N+1)/\ln(2) - 1$ when $I = \{1, ..., N\}$), let $\bar{S}_j$ be the linear span of the e_k's for $k \in \{2^j, ..., 2^{j+1} - 1\}$. Note that $\dim(\bar{S}_j) = 2^j$ and that $\bar{S}_j \subset S_D$ for $D = D(j) = 2^{j+1} - 1$. Setting for $\mathcal{F} \subset \ell_2(I)$ and $r > 0$,

$$\mathcal{F}[r] = \{f \in \mathcal{F}, \ \|f\| = r\},$$

we get that

$$\bigcup_{j=0}^{J} \bar{S}_j[r_{D(j)}] \subset \bigcup_{j=0}^{J} S_{D(j)}[r_{D(j)}] \subset \bigcup_{D \in I} S_D[r_D].$$

We now use Proposition 9 with $p_j = 6/[\pi^2(j+1)^2]$ for $j \in \mathbb{N}$ and we get

$$\beta\left(\bigcup_{D \in I} \{f \in S_D, \ \|f\| = r_D\}\right) \geq \delta,$$

as soon as for those $D = D(j)$,

$$r_D^2 \leq \sqrt{2\ln(1 + \eta^2/p_j)}\sqrt{D} = \sqrt{2\ln(1 + \eta^2\pi^2(j+1)^2/6)}2^{j/2}. \tag{22}$$

Thus, it remains to check (22). Using that

$$j + 1 = \ln(D + 1)/\ln(2) \geq \ln(D + 1),$$

$2^{j/2} \geq \sqrt{D/2}$ and the convexity inequality

$$\ln(1 + ux) \geq u \ln(1 + x), \tag{23}$$

which holds for all $x > 0$ and $u \in [0, 1]$, we obtain that

$$
\begin{aligned}
\sqrt{2 \ln(1 + \eta^2 \pi^2 (j + 1)^2 / 6)} 2^{j/2} & \\
\geq \ & \left[(\eta \pi / \sqrt{6}) \wedge 1 \right] \sqrt{\ln(1 + \ln^2(D + 1))} \sqrt{D} \\
\geq \ & \sqrt{2} \left[(\eta \pi / \sqrt{6}) \wedge 1 \right] \sqrt{\ln \ln(D + 1))} \sqrt{D} \\
= \ & \bar{\rho}_D^2.
\end{aligned}
$$

Since by assumption $\bar{\rho}_D^2 \geq r_D^2$, (22) is proved and the result follows. □

6.3. Collection of nested ellipsoids.
In this section we consider the case of a collection of ellipsoids of the form $\{\mathcal{E}_{a,2}(R), \ R \in \mathbb{R}_+\}$.

Corollary 4. *For each $R > 0$, let us set*

$$\bar{\rho}_{a,2,R}^2 = \sup_{D \in I} \left[\bar{\rho}_D^2 \wedge \left(R^2 a_D^2 \right) \right],$$

where $\bar{\rho}_D$ is given by (21). Then we have that

$$\beta \left(\bigcup_{R > 0} \{ f \in \mathcal{E}_{a,2}(R), \ \ \|f\| \geq \bar{\rho}_{a,2,R} \} \right) \geq \delta.$$

Comment: The problem of finding a test that achieves (up) to a constant the minimax separation rate simultaneously over a family of alternatives is usually called the problem of adaptation. In contrast with the problem of estimation, in the problem of hypothesis testing adaptation is in general impossible. This result was proved by Spokoiny (1996) for the case of a family of Besov bodies. In the case considered here we deal with the family of nested ellipsoids $\{\mathcal{E}_{a,2}(R), \ R \in \mathbb{R}_+\}$. This amounts to adapting over the radius R in $\mathbb{R}_+^*$. In the literature, one usually tries to adapt over both $R \in \mathbb{R}_+^*$ and a among some non trivial class of sequences of positive numbers, but since we are interested in lower bounds, the problem of adaptation over R only is enough. As Spokoiny, by this result we obtain that the problem of finding adaptive tests is possible only if one tolerates a loss a efficiency (which is of order a $\ln \ln(N)$ factor in the regression framework).

Proof. We use the same notations as in the proof of Proposition 3. Let $D(R) \in I$ which achieves the supremum of $\bar{\rho}_D^2 \wedge (R^2 a_D^2) = \bar{r}_D^2$ over I (the existence of $D(R)$ is obvious when I is finite and is a consequence of the monotonicity of $\bar{\rho}_D$ and $R^2 a_D^2$ otherwise). Arguing as in the proof of Proposition 3 we have for each R,

$$\left\{ f \in S_{D(R)}, \ \|f\| = r_{D(R)} \right\} \subset \left\{ f \in \mathcal{E}_{a,2}(R), \ \|f\| \geq r_{D(R)} \right\},$$

and as $D(R)$ describes I when R varies, we obtain that

$$
\begin{aligned}
\bigcup_{D \in I} \{ f \in S_D, \ \|f\| = r_D \} \ &= \ \bigcup_{R > 0} \{ f \in S_{D(R)}, \ \|f\| = r_{D(R)} \} \\
&\subset \ \bigcup_{R > 0} \left\{ f \in \mathcal{E}_{a,2}(R), \ \|f\| \geq r_{D(R)} \right\}.
\end{aligned}
$$

Then the result follows from Corollary 3. $\qquad\qquad\square$

7. Proof of Theorem 1 and Propositions 1 and 9

7.1. **A general method to obtain lower bounds.** The proof is based on a Bayesian approach which is classical (see Lehmann (1997) Chapter 6 for example). The starting point of the proof is similar to that described in Ingster (1993a,b,c) and borrows some classical inequalities on the norm in total variation that can be found in Le Cam (1986) (Chapter 4). . For the sake of completeness, let us describe the main ideas of the approach.

Let $\mathcal{F}$ be some subset $\ell_2(I)$ and ρ some positive number. Let μ_ρ be some probability measure on

$$\mathcal{F}[\rho] = \{ f \in \mathcal{F}, \ \|f\| = \rho \}.$$

Setting $P_{\mu_\rho} = \int P_f \, d\mu_\rho(f)$ and Φ_α the set of level-α tests, we have

$$
\begin{aligned}
\beta(\mathcal{F}[\rho]) \ &\geq \ \inf_{\phi_\alpha \in \Phi_\alpha} P_{\mu_\rho} [\phi_\alpha = 0] \\
&\geq \ 1 - \alpha - \sup_{A/ \ P_0(A) \leq \alpha} \left| P_{\mu_\rho}(A) - P_0(A) \right| \\
&\geq \ 1 - \alpha - \sup_{A \in \mathcal{A}} \left| P_{\mu_\rho}(A) - P_0(A) \right| \\
&= \ 1 - \alpha - \frac{1}{2} \| P_{\mu_\rho} - P_0 \|, \qquad\qquad (24)
\end{aligned}
$$

where $\| P_{\mu_\rho} - P_0 \|$ denotes the total variation norm between the probabilities P_{μ_ρ} and P_0.

Whenever P_{μ_ρ} is absolutely continuous with respect to P_0, the norm in total variation between these two probabilities is easy to compute. Setting

$$L_{\mu_\rho}(y) = \frac{dP_{\mu_\rho}}{dP_0}(y),$$

we get

$$
\begin{aligned}
\|P_{\mu_\rho} - P_0\| &= \int \left| L_{\mu_\rho}(y) - 1 \right| dP_0(y), \\
&= \mathbb{E}_0 \left[\left| L_{\mu_\rho}(Y) - 1 \right| \right], \\
&\leq \left(\mathbb{E}_0 \left[L_{\mu_\rho}^2(Y) \right] - 1 \right)^{1/2},
\end{aligned}
$$

and we deduce from (24) that

$$\beta(\mathcal{F}[\rho]) \geq 1 - \alpha - \frac{1}{2} \left(\mathbb{E}_0 \left[L_{\mu_\rho}^2(Y) \right] - 1 \right)^{1/2}.$$

Thus, it remains to find some $\rho^* = \rho^*(\eta)$ such that

$$\ln \left(\mathbb{E}_0 \left[L_{\mu_{\rho^*}}^2(Y) \right] \right) \leq \mathcal{L}(\eta), \tag{25}$$

to ensure that for all $\rho \leq \rho^*$,

$$\beta(\mathcal{F}[\rho]) \geq 1 - \alpha - \eta = \delta.$$

7.2. **Proof of Theorem 1.** By homogeneity, we assume that $\sigma^2 = 1$. Let $\hat{m}$ be some random variable uniformly distributed over $\mathcal{M}(k,n)$ and for each $m \in \mathcal{M}(k,n)$ let $\varepsilon^m = (\varepsilon_j^m)_{j \in m}$ be a sequence of Rademacher random variable (i.e. for each m, the ε_j^m's are i.i.d. random variables taking the values ± 1 with probability $1/2$). We assume that for all $m \in \mathcal{M}(k,n)$, ε^m and $\hat{m}$ are independent. Let ρ be given and μ_ρ the distribution of the random variable $\sum_{j \in \hat{m}} \lambda \varepsilon_j^{\hat{m}} e_j /$ where $\lambda = \rho/\sqrt{k}$. Clearly μ_ρ is supported by $\mathcal{F}[\rho]$. To prove the result, we apply the method described in the previous section with

$$
\begin{aligned}
L_{\mu_\rho}(Y) &= \mathbb{E}_{\varepsilon, \hat{m}} \left[\exp \left(-\frac{1}{2}\rho^2 + \lambda \sum_{j \in \hat{m}} \varepsilon_j^{\hat{m}} Y_j \right) \right] \\
&= \frac{1}{C_n^k} \sum_{m \in \mathcal{M}(k,n)} \mathbb{E}_\varepsilon \left[\exp \left(-\frac{1}{2}\rho^2 + \lambda \sum_{j \in m} \varepsilon_j^m Y_j \right) \right] \\
&= e^{-\rho^2/2} \frac{1}{C_n^k} \sum_{m \in \mathcal{M}(k,n)} \prod_{j \in m} \cosh\left(\lambda Y_j \right).
\end{aligned}
$$

Let us now compute $\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right]$. Introducing the notation

$$m\Delta m' = (m \cup m') \setminus (m \cap m')$$

for m, m' belonging to $\mathcal{M}(k,n)$ and we obtain that

$$
\begin{aligned}
&\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right] \\
&= \frac{e^{-\rho^2}}{(C_n^k)^2} \sum_{m,m'\in\mathcal{M}(k,n)} \mathbb{E}_0\left[\prod_{j\in m}\cosh\left(\lambda Y_j\right)\prod_{j\in m'}\cosh\left(\lambda Y_j\right)\right] \\
&= \frac{e^{-\rho^2}}{(C_n^k)^2} \sum_{m,m'\in\mathcal{M}(k,n)} \mathbb{E}_0\left[\prod_{j\in m\cap m'}\cosh^2\left(\lambda Y_j\right)\prod_{j\in m\Delta m'}\cosh\left(\lambda Y_j\right)\right] \\
&= \frac{e^{-\rho^2}}{(C_n^k)^2} \sum_{m,m'\in\mathcal{M}(k,n)} \left(\mathbb{E}_0\left[\cosh^2\left(\lambda Y_1\right)\right]\right)^{|m\cap m'|}\left(\mathbb{E}_0\left[\cosh\left(\lambda Y_1\right)\right]\right)^{|m\Delta m'|},
\end{aligned}
$$

by independence between the Y_j's. Using the fact that

$$\left(\mathbb{E}_0\left[\cosh(\lambda Y_1)\right]\right) = e^{\lambda^2/2}, \quad \mathbb{E}_0\left[\cosh^2(\lambda Y_1)\right] = e^{\lambda^2}\cosh(\lambda^2),$$

and noting that $|m\cap m'| + |m\Delta m'|/2 = k$, we derive

$$
\begin{aligned}
\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right] &= \frac{1}{(C_n^k)^2} \sum_{m,m'\in\mathcal{M}(k,n)} \left(\cosh\left(\lambda^2\right)\right)^{|m\cap m'|} \\
&= \sum_{j=1}^k \left(\cosh\left(\lambda^2\right)\right)^j p_{j,k,n},
\end{aligned}
$$

where

$$p_{j,k,n} = (C_n^k)^{-2}\left|\left\{(m,m')\in\mathcal{M}(k,n)^2/\ |m\cap m'| = j\right\}\right|.$$

If $j < 2k - n$ then obviously $p_{j,k,n} = 0$ and $p_{j,k,n} = C_k^{k-j}C_{n-k}^{k-j}/C_n^k$ otherwise. Hence, $p_{j,k,n} = P[X=j]$ where X is a random variable distributed according to a Hypergeometric distribution with parameters n, k and k/n. Thus, we derive that

$$\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right] = \mathbb{E}\left[\left(\cosh(\lambda^2)\right)^X\right]. \tag{26}$$

We know from Aldous (1985, p.173) that X has the same distribution as the random variable $\mathbb{E}[Z/\mathcal{B}_n]$ where Z is a binomial random variable of parameters $k, k/n$ and $\mathcal{B}_n$ some suitable σ-algebra. Thus, by a

convexity argument we infer from (26) that

$$
\begin{aligned}
\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right] &\leq \mathbb{E}\left[\left(\cosh(\lambda^2)\right)^Z\right] \\
&= \left(1 + \frac{k}{n}\left(\cosh(\lambda^2) - 1\right)\right)^k \\
&= \exp\left[k\ln\left(1 + \frac{k}{n}\left(\cosh(\lambda^2) - 1\right)\right)\right]. \quad (27)
\end{aligned}
$$

For $\rho \leq \rho_{k,n}$, one has

$$
\lambda^2 \leq \lambda^2_{k,n} = \ln\left(1 + u + \sqrt{2u + u^2}\right),
$$

where $u = \mathcal{L}(\eta)n/k^2$. We deduce from (27) that for all $\rho \leq \rho_{k,n}$

$$
\begin{aligned}
\mathbb{E}_0\left[L^2_{\mu_\rho}(Y)\right] &\leq \exp\left[k\ln\left(1 + \frac{k}{n}\left(\cosh(\lambda^2_{k,n}) - 1\right)\right)\right] \\
&= \exp\left[k\ln\left(1 + \frac{k}{n}u\right)\right] \\
&\leq \exp\left[\frac{k^2}{n}u\right] = \exp\left[\mathcal{L}(\eta)\right] = 1 + \eta^2.
\end{aligned}
$$

To complete the proof of Theorem 1, it remains to check (7). Clearly we have that

$$
\rho^2_{k,n} \geq k\ln\left(1 + \left[(2\mathcal{L}(\eta)) \wedge \sqrt{2\mathcal{L}(\eta)}\right]\left[\frac{n}{k^2} \vee \sqrt{\frac{n}{k^2}}\right]\right),
$$

and thanks to the convexity inequality (23) we get that

$$
\rho^2_{k,n} \geq ((2\mathcal{L}(\eta)) \wedge 1)\, k\ln\left(1 + \frac{n}{k^2} \vee \sqrt{\frac{n}{k^2}}\right).
$$

The result follows since for $\alpha + \delta \leq 59\%$, $2\mathcal{L}(\eta) \geq 1$.

7.3. **Proof of Proposition 1.** We argue as previously taking $n = k = D$. Then the right-hand side of (27) merely becomes $(\cosh(\lambda^2))^D$. Since for all $x \in \mathbb{R}$,

$$
\cosh(x) \leq \exp(x^2/2)
$$

(compare the series) the result follows easily.

7.4. **Proof of Proposition 9.** It is enough to show the result under the assumption that $\sum_{j\in\mathcal{J}} p_j = 1$. Arguing as in the proof of Theorem 1 we know that for each $r_j \leq \tilde{\rho}_j(\eta/\sqrt{p_j})$ there exists some measure μ_j over

$$\mathcal{F}_j[r_j] = \{f \in \mathcal{F}_j, \ \|f\| = r_j\}$$

such that such that

$$\mathbb{E}_0\left[L^2_{\mu_j}(Y)\right] \leq 1 + \eta^2/p_j. \tag{28}$$

Let us now set $\mu = \sum_{j\in\mathcal{J}} p_j\mu_j$ which is a probability measure over $\bigcup_{j\in\mathcal{J}} \mathcal{F}_j[r_j]$. Denoting L_{μ_j} the density of $P_{\mu_j} = \int P_f d\mu_j(f)$ with respect to P_0 we have that

$$L_\mu(Y) = \frac{dP_\mu}{dP_0}(Y) = \sum_{j\in\mathcal{J}} p_j L_{\mu_j}(Y),$$

and thus

$$\mathbb{E}_0\left[L^2_\mu(Y)\right] = \sum_{j,j'\in\mathcal{J}} p_j p_{j'} \mathbb{E}_0\left[L_{\mu_j}(Y)L_{\mu_{j'}}(Y)\right].$$

Since for $j \neq j'$, $\mathcal{F}_j$ and $\mathcal{F}_{j'}$ are orthogonal the random variables $L_{\mu_j}(Y)$ and $L_{\mu_{j'}}(Y)$ are independent and thus,

$$\mathbb{E}_0\left[L^2_\mu(Y)\right] = 1 + \sum_{j\in\mathcal{J}} p_j^2\left(\mathbb{E}_0\left[L^2_{\mu_j}(Y)\right] - 1\right) \leq 1 + \eta^2$$

thanks to (28). This leads to the announced result via (25).

8. Proof of Propositions 2, 4, 6 and 7

8.1. **Some preliminary result.** The next result describes the performance of tests based on χ^2-statistics. It is a slight modification (the constants are sharper) of Theorem 1 and Proposition 1 in Baraud, Huet, Laurent (1999).

Theorem 2. *Let $\alpha, \delta \in [0, 1]$ and $\mathcal{F} \subset \ell_2(I)$. Let $\mathcal{M}$ be a class of finite subsets of I and $\bar{\alpha} = (\alpha_m)_{m\in\mathcal{M}}$ a sequence of non negative numbers such that $\sum_{m\in\mathcal{M}} \alpha_m \leq \alpha$. For each $f \in \mathcal{F}$, let us set*

$$\tilde{\rho}^2_{\mathcal{M},\bar{\alpha},\delta}(f)$$

$$= \inf_{m\in\mathcal{M}}\left\{\sum_{j\notin m} f_j^2 + 2\sqrt{5}\ln^{1/2}\left(\frac{1}{\alpha_m\delta}\right)\sqrt{|m|}\sigma^2 + 8\ln\left(\frac{1}{\alpha_m\delta}\right)\sigma^2\right\}$$

$$\leq \inf_{m\in\mathcal{M}}\left\{\sum_{j\notin m} f_j^2 + 2(\sqrt{5}+4)\ln\left(\frac{1}{\alpha_m\delta}\right)\sqrt{|m|}\sigma^2\right\} \tag{29}$$

Then, the test $\phi_{\mathcal{M},\bar{\alpha}}$ *defined by*

$$\phi_{\mathcal{M},\bar{\alpha}} = \sup_{m \in \mathcal{M}} \phi_{m,\alpha_m} \quad \text{where } \phi_{m,\alpha_m} \text{ is given by (8),}$$

satisfies

$$P_0[\phi_{\mathcal{M},\bar{\alpha}} = 1] \leq \alpha \quad \text{and} \quad P_f[\phi_{\mathcal{M},\bar{\alpha}} = 0] \leq \delta,$$

for all $f \in \mathcal{F}$ *such that* $\|f\| \geq \tilde{\rho}_{\mathcal{M},\bar{\alpha},\delta}(f)$.

Comment: Thanks to Theorem 2 the proofs of Propositions 2, 4, 6 and 7 reduce to obtaining some adequate upper bound on $\sup_{f \in \mathcal{F}} \tilde{\rho}_{\mathcal{M},\bar{\alpha},\delta}(f)$.

Proof. Inequality (29) is clear and the fact that the test $\phi_{\mathcal{M},\bar{\alpha}}$ is of level α merely derives from the following:

$$P_0[\phi_{\mathcal{M},\bar{\alpha}} = 1] \leq \sum_{m \in \mathcal{M}} P_0[\phi_{m,\alpha_m} = 1] = \sum_{m \in \mathcal{M}} \alpha_m \leq \alpha.$$

Let us now show the result on the power of the test. Without loss of generality we can take $\sigma^2 = 1$. For each $m \in \mathcal{M}$ we set $Z_{m,f}^2 = \sum_{j \in m} Y_j^2$ and $E_m^2 = \sum_{j \in m} f_j^2$. On the one hand we have that

$$
\begin{aligned}
P_f[\phi_{\mathcal{M},\bar{\alpha}} = 0] &= P_f[\forall m \in \mathcal{M}, \ Z_{m,f}^2 \leq t_{|m|,\alpha_m}] \\
&\leq \inf_{m \in \mathcal{M}} P_f[\ Z_{m,f}^2 \leq t_{|m|,\alpha_m}].
\end{aligned}
\tag{30}
$$

On the other hand, some deviation inequality on non central χ^2 random variables due to Birgé (1999) (Lemma 1) tells us that

$$P_f\left[Z_{m,f}^2 \leq |m| + E_m^2 - 2\sqrt{(|m| + 2E_m^2)\ln(1/\delta)}\right] \leq \delta.$$

Thus, the result is proved if we show that for some m in $\mathcal{M}$,

$$t_{|m|,\alpha_m} \leq |m| + E_m^2 - 2\sqrt{(|m| + 2E_m^2)\ln(1/\delta)}.\tag{31}$$

We now prove that (31) holds if m satisfies

$$E_m^2 = \|f\|^2 - \sum_{j \notin m} f_j^2 > 2\sqrt{5}\ln^{1/2}\left(\frac{1}{\alpha_m\delta}\right)\sqrt{|m|} + 8\ln\left(\frac{1}{\alpha_m\delta}\right).\tag{32}$$

We start with an inequality due to Laurent & Massart (1998) on central χ^2 random variables: we have

$$t_{|m|,\alpha_m} \leq |m| + 2\sqrt{|m|\ln(1/\alpha_m)} + 2\ln(1/\alpha_m).$$

Setting $x = \ln(1/\delta)$ and $y_m = \ln(1/\alpha_m)$, we need to check that

$$\frac{1}{2}E_m^2 \geq \sqrt{(|m| + 2E_m^2)x} + \sqrt{|m|y_m} + y_m.\tag{33}$$

Solving inequation (33) with respect to E_m^2 we obtain that (33) holds true as soon as

$$\frac{1}{2}E_m^2 \geq \sqrt{|m|y_m} + \sqrt{x}\sqrt{4x + 4y_m + 4\sqrt{|m|y_m} + |m|} + 2x + y_m. \quad (34)$$

Hence it remains to obtain a suitable upper bound for the right-hand side of (34). Using the inequalities $\sqrt{u+v} \leq \sqrt{u} + \sqrt{v}$, $2uv \leq u^2 + v^2$ and $\sqrt{u} + 2\sqrt{v} \leq \sqrt{5}\sqrt{u+v}$ which holds for all $u, v > 0$, we obtain that

$$\sqrt{|m|y_m} + \sqrt{x}\sqrt{4x + 4y_m + 4\sqrt{|m|y_m} + |m|} + 2x + y_m$$

$$\leq \quad \sqrt{|m|}\left(\sqrt{x} + \sqrt{y_m}\right) + 2\sqrt{x}\sqrt{x + y_m + \sqrt{|m|y_m}} + 2x + y_m$$

$$\leq \quad \sqrt{|m|}\left(\sqrt{x} + 2\sqrt{y_m}\right) + 4x + 2y_m$$

$$\leq \quad \sqrt{5}\sqrt{x + y_m}\sqrt{|m|} + 4(x + y_m),$$

the last expression being smaller than $E_m^2/2$ by (32). This ends the proof of (31). $\qquad\square$

8.2. **Proof of Proposition 2.** We set $\mathcal{M} = \mathcal{M}(k, n)$ and for each $m \in \mathcal{M}$,

$$\alpha_m = \alpha_{k,n} = \alpha/(2C_n^k) \geq \alpha/(2(en/k)^k).$$

We deduce from Theorem 2 that the test ϕ_α^* is of level α. Concerning the power of the test, we have that

$$\ln\left(\frac{1}{\alpha_m\delta}\right) \leq \ln\left(\frac{2}{\alpha\delta}\right) + k\ln\left(\frac{en}{k}\right) \leq \left[\ln\left(\frac{2}{\alpha\delta}\right) + 1\right]k\ln\left(\frac{en}{k}\right)$$

$$= \ln\left(\frac{2e}{\alpha\delta}\right)k\ln\left(\frac{en}{k}\right),$$

thus by setting

$$L = \ln\left(\frac{2e}{\alpha\delta}\right) \geq 1,$$

and choosing m among $\mathcal{M}(k, n)$ such that $f_j = 0$ for $j \notin m$ we deduce that for each f,

$$\tilde{\rho}_{\mathcal{M},\bar{\alpha},\delta}^2(f) \leq 2\sqrt{5}\sqrt{L}k\ln^{1/2}\left(\frac{en}{k}\right)\sigma^2 + 8Lk\ln\left(\frac{en}{k}\right)\sigma^2$$

$$\leq 2(\sqrt{5} + 4)Lk\ln\left(\frac{en}{k}\right)\sigma^2. \quad (35)$$

Now, by choosing $m = \{1, ..., n\}$ and arguing in the same way we get that

$$\tilde{\rho}_{\mathcal{M},\bar{\alpha},\delta}^2(f) \leq 2(\sqrt{5} + 4)L\sqrt{n}\sigma^2. \quad (36)$$

Inequalities (35), (36) and Theorem 2 lead to the desired result.

8.3. **Proof of Proposition 4.** It is straightforward to see that the test ϕ_α^* is of level α. In the sequel, we set

$$A_{a,2,R} = \left\{ D \in I, \ R^2 a_D^2 \leq \sqrt{D}\sigma^2 \right\}, \quad L = \ln(1/(\alpha\delta)) \geq 1,$$

$\mathcal{F} = \mathcal{E}_{a,2}(R)$, $\mathcal{M} = \{\{1, ..., D^*\}\}$ and $\bar{\alpha} = \alpha$. For the power of the test, we use Theorem 2 with some suitable upper bound for the quantity $\tilde{\rho}_{\mathcal{M},\bar{\alpha},\delta}^2(f)$ with $f \in \mathcal{F}$. To do so, we distinguish between two cases. Firstly, if $A_{a,2,R} = \emptyset$ then $D^* = N$ (note that the condition is possible only in the case of a finite I since the a_j's converge towards 0) and for all $D \in I$ $R^2 a_D^2 > \sqrt{D}\sigma^2$. This implies that

$$\forall f \in \mathcal{F}, \quad \sum_{j>D^*} f_j^2 = 0 \ \text{ and } \ \sup_{D\in I}\left[\left(\sqrt{D}\sigma^2\right) \wedge \left(R^2 a_D^2\right)\right] = \sqrt{N}\sigma^2$$

and thus for all $f \in \mathcal{F}$,

$$\begin{aligned}
\tilde{\rho}_{\mathcal{M},\bar{\alpha},\delta}^2(f) &\leq 2(\sqrt{5}+4)L\sqrt{N}\sigma^2 \\
&= 2(\sqrt{5}+4)L\sup_{D\in I}\left[\left(\sqrt{D}\sigma^2\right) \wedge \left(R^2 a_D^2\right)\right],
\end{aligned}$$

which proves the result in this case.

Secondly, if $A_{a,2,R} \neq \emptyset$ then there exists some $D^* \in I$ such that $R^2 a_{D^*}^2 \leq \sqrt{D^*}\sigma^2$ and by assumption we know that $D^* \geq 2$. For such a D^*, we have that

$$\forall f \in \mathcal{F}, \quad \sum_{j>D^*} f_j^2 \leq R^2 a_{D^*}^2,$$

and that

$$\begin{aligned}
\sqrt{D^*}\sigma^2 &\leq \sqrt{2}\sqrt{D^*-1}\sigma^2 = \sqrt{2}\left[\left(\sqrt{D^*-1}\sigma^2\right) \wedge \left(R^2 a_{D^*-1}\right)\right] \\
&\leq \sqrt{2}\sup_{D\in I}\left[\left(\sqrt{D}\sigma^2\right) \wedge \left(R^2 a_D^2\right)\right].
\end{aligned}$$

Thus for all $f \in \mathcal{F}$,

$$\begin{aligned}
\tilde{\rho}_{\mathcal{M},\bar{\alpha},\delta}^2(f) &\leq R^2 a_{D^*}^2 + 2(\sqrt{5}+4)L\sqrt{D^*}\sigma^2 \\
&\leq \left[1 + 2(\sqrt{5}+4)L\right]\sqrt{D^*}\sigma^2 \\
&\leq \sqrt{2}\left[1 + 2(\sqrt{5}+4)L\right]\sup_{D\in I}\left[\left(\sqrt{D}\sigma^2\right) \wedge \left(R^2 a_D^2\right)\right],
\end{aligned}$$

which concludes the proof.

8.4. **Proof of Propositions 6 and 7.** Let us set,

$$\mathcal{M} = \{1, ..., D^*\} \cup \left(\bigcup_{j > D^*, j \in I} \{j\} \right),$$

$\alpha_{\{1,...,D\}} = \alpha_{D^*} = \alpha/2$ and

$$\forall j \in I, \ j > D^* \ \alpha_{\{j\}} = \alpha_j = 2\alpha/(\pi^2(j - D^*)^2).$$

Thanks to Theorem 2, we obtain that the test $\phi_\alpha^* = \phi_{\mathcal{M},\bar{\alpha}}$ is clearly of level α. It remains to prove the result concerning the power of the test.

In the sequel, we set

$$\kappa = 2(\sqrt{5} + 4)\ln(2/(\alpha\delta)) \ \text{ and } \ \mathcal{F} = \mathcal{E}_{a,p}(R).$$

8.4.1. *Reduction of the problem.* Let us define the set $\tilde{A}_{a,p,R}$ by

$$\tilde{A}_{a,p,R} = \left\{ D \in I, \ R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p} \leq \lceil \sqrt{D} \rceil \sigma^2 \right\}.$$

Note that this set is non void when I is infinite since the a_D's converge towards 0.

We first prove Propositions 6 and 7 under one of the following conditions

(i) $\tilde{A}_{a,p,R} = \emptyset$ and (15) holds true.
(ii) f belongs to the space

$$\begin{aligned} \mathcal{F}_{loc} &= \left\{ f \in \ell_2(I), \ \exists j > D^*, \ |f_j|^2 \geq b_{j-D^*}^2 \sigma^2 \right\} \\ &= \left\{ f \in \ell_2(I), \ \|f\|^2 \geq \inf_{j > D^*} \left\{ \sum_{k \in I, k \neq j} f_k^2 + b_{j-D^*}^2 \sigma^2 \right\} \right\}. \end{aligned}$$

where for $j \in \mathbb{N}^*$, the b_j's are defined by

$$b_j^2 = 2(\sqrt{5} + 4)\ln(\pi^2 j^2/(2\alpha\delta)).$$

Let us now assume (i). Then $I = \{1, ..., N\}$, $D^* = N$ and

$$\varrho_{a,p,R}^2 = \lceil \sqrt{N} \rceil \sigma^2.$$

By applying Theorem 2 with $\phi_{N,\alpha/2} = \phi_{\{1,...,N\},\alpha/2}$ and arguing as in the proof of Proposition 4 we obtain that

$$P_f\left[\phi_\alpha^* = 0\right] \leq P_f\left[\phi_{N,\alpha/2} = 0\right] \leq \delta$$

for all $f \in \mathcal{F}$ such that $\|f\| \geq \tilde{\rho}_{\mathcal{M},\alpha/2,\delta}$, where we have taken $\mathcal{M} = \{\{1,...,N\}\}$. Since now,

$$\begin{aligned}
\tilde{\rho}^2_{\mathcal{M},\bar{\alpha},\delta}(f) &\leq \kappa\sqrt{N}\sigma^2 \\
&\leq C\left(\ln(2+N)\right)^{1-p/2}\varrho^2_{a,p,R},
\end{aligned}$$

the result follows under (15).

Let us now assume (ii). By setting

$$\mathcal{M}' = \{\{j\}/ \ j \in I, \ j > D^*\} \quad \text{and } \bar{\alpha}' = (\alpha_j)_{j \in \mathcal{M}'},$$

we have that $\phi_{loc,\alpha} = \phi_{\mathcal{M}',\bar{\alpha}'}$. We derive from (29) that for all $f \in \mathcal{F}_{loc}$,

$$\|f\|^2 \geq \inf_{j>D^*}\left\{\sum_{k\in I, k\neq j} f_k^2 + b^2_{j-D^*}\sigma^2\right\} \geq \tilde{\rho}^2_{\mathcal{M}',\bar{\alpha}',\delta}(f),$$

which leads to the result, i.e.

$$P_f\left[\phi^*_\alpha = 0\right] \leq P_f\left[\phi_{loc,\alpha} = 0\right] \leq \delta,$$

by applying Theorem 2.

Leaving out the cases (i) and (ii), we now assume that $\tilde{A}_{a,p,R} \neq \emptyset$ and that f belongs to the set

$$\mathcal{H} = \mathcal{F} \cap \left\{f \in \ell_2(I), \ \forall j > D^*, \ |f_j|^2 \leq b^2_{j-D^*}\sigma^2\right\}.$$

Thus, it remains to get some suitable upper bound on $\tilde{\rho}^2_{\mathcal{M},\bar{\alpha},\delta}(f)$ for $f \in \mathcal{H}$.

8.4.2. *End of the proof of Proposition 6.* For all $f \in \mathcal{H}$, we bound the bias term in the following way:

$$\sum_{j>D^*} f_j^2 \leq \sum_{j>D^*} a_j^p b^{2-p}_{j-D^*}\sigma^{2-p}\frac{|f_j|^p}{a_j^p} \tag{37}$$

$$\leq b^{2-p}_N R^p a^p_{D^*}\sigma^{2-p}. \tag{38}$$

Since $\tilde{A}_{a,p,R} \neq \emptyset$, we have that

$$\begin{aligned}
D^* &= \inf\{D \in I, \ R^2 a_D^2\lceil\sqrt{D}\rceil^{1-2/p} \leq \lceil\sqrt{D}\rceil\sigma^2\} \\
&= \inf\{D \in I, \ R^p a_D^p\sigma^{2-p} \leq \lceil\sqrt{D}\rceil\sigma^2\},
\end{aligned}$$

and by (38) we obtain that for all $f \in \mathcal{H}$,

$$\begin{aligned}
\tilde{\rho}^2_{\mathcal{M},\bar{\alpha},\delta}(f) &\leq \sum_{j>D^*} f_j^2 + \kappa\sqrt{D^*}\sigma^2 \\
&\leq \left(\kappa + b^{2-p}_N\right)\lceil\sqrt{D^*}\rceil\sigma^2.
\end{aligned}$$

By assumption $D^* \geq 2$ which implies that $\lceil \sqrt{D^*} \rceil \leq 2\lceil \sqrt{D^*-1} \rceil$ and thus by definition of D^* we get

$$
\begin{aligned}
\tilde{\rho}^2_{\mathcal{M},\bar{\alpha},\delta}(f) & \\
\leq\ & 2\left(\kappa + b_N^{2-p}\right) \lceil \sqrt{D^*-1} \rceil \sigma^2 \\
=\ & 2\left(\kappa + b_N^{2-p}\right) \left[\left(\lceil \sqrt{D^*-1} \rceil \sigma^2\right) \wedge \left(R^2 a_{D^*-1}^2 \lceil \sqrt{D^*-1} \rceil^{1-2/p}\right)\right] \\
\leq\ & 2\left(\kappa + b_N^{2-p}\right) \sup_{D \in I} \left[\left(\lceil \sqrt{D} \rceil \sigma^2\right) \wedge \left(R^2 a_D^2 \lceil \sqrt{D} \rceil^{1-2/p}\right)\right].
\end{aligned}
$$

To end the proof it remains to show that

$$
2\left(\kappa + b_N^{2-p}\right) \leq 8(\sqrt{5}+4)\ln\left(\frac{e\pi}{\alpha\delta}\right)\ln^{1-p/2}(2+N).
$$

This inequality is a straightforward consequence of the following ones: for all $j \geq 1$,

$$
\begin{aligned}
b_j^2 &= 4(\sqrt{5}+4)\left(\ln\left(\frac{\pi}{\sqrt{2\alpha\delta}}\right) + \ln(j)\right) \\
&\leq 2(\sqrt{5}+4)\ln\left(\frac{e^2\pi^2}{2\alpha\delta}\right)\ln(2+j). \qquad (39)
\end{aligned}
$$

8.4.3. *End of the proof of Proposition 7.* Using the Assumption (H) on the a_j's we infer from (37) that

$$
\begin{aligned}
\sum_{j>D^*} f_j^2 &\leq R^p \sigma^{2-p} \sum_{j\geq 1} a_{j+D^*}^p b_j^{2-p} \\
&\leq R^p a_{D^*}^p \sigma^{2-p} \sum_{j\geq 1} \theta_j^p b_j^{2-p}.
\end{aligned}
$$

Using now (39) we obtain that

$$
\sum_{j>D^*} f_j^2 \leq 2\Sigma(\sqrt{5}+4)\ln\left(\frac{e^2\pi^2}{2\alpha\delta}\right) R^p a_{D^*}^p \sigma^{2-p},
$$

and the result follows by arguing as previously.

REFERENCES

Aldous, D.J. (1985). *Exchangeability and related topics.* Ecole d'été de probabilités de Saint-Flour XIII, Lect. Notes Math. 1117, 1-198.

Baraud, Y., Huet, S., and Laurent, B. (1999). Adaptive tests of linear hypotheses by model selection. *Tech. rept. 2000-69 Université Paris XI.*

Birgé, L. (2001). An alternative point of view on Lepski's method. *State of the art in probability and statistics (Leiden, 1999)* (ed. Monogr., IMS Lecture Notes), **36** 113 − 133. Physica-Verlag, Beachwood.

Birgé, L., and Massart, P. (2001). Gaussian model selection. *AJ. Eur. Math. Soc.* **3** 203 − 268.

Donoho, D.L., and Johnstone, I.M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26** 879 − 921.

Ermakov, M.S. (1991). Minimax detection of a signal in a Gaussian white noise. *Theory Probab. Appl.* **35** 667 − 679.

Ingster, Yu. I. (1993a). Asymptotically minimax hypothesis testing for nonparametric alternatives I. *Math. Methods Statist.* **2** 85 − 114.

Ingster, Yu. I. (1993b). Asymptotically minimax hypothesis testing for nonparametric alternatives II. *Math. Methods Statist.* **3** 171 − 189.

Ingster, Yu. I. (1993c). Asymptotically minimax hypothesis testing for nonparametric alternatives III. *Math. Methods Statist.* **4** 249 − 268.

Ingster, Yu. I. (1998). Adaptation in minimax nonparametric hypothesis testing for ellipsoids and Besov bodies. Technical report 419, Weierstrass Institute, Berlin.

Ingster, Yu. I., and Suslina I. A. (1998). Minimax detection of a signal for Besov bodies and balls. *Problems Inform. Transmission* **34** 48 − 59.

Laurent, B., and Massart, P.(2000). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28** 1302 − 1338.

Le Cam, L. (1986). *Asymptotic methods in statistical decision theory.* Springer Series in Statistics. Springer-Verlag, New-York.

Lehmann, E.L. (1986). *Testing statistical hypotheses (Second Edition).* Springer Texts in Statistics. Springer, New-York.

Lepski, O. V., and Spokoiny, V. G. (1999). Minimax nonparametric hypothesis testing: the case of inhomogeneous alternative. *Bernoulli* **5** 333 − 358.

Spokoiny, V. G. (1996). Adaptive hypothesis testing using wavelets. *Ann. Statist.* **24***, 2477 − 2498.

Ecole Normale Supérieure, DMA, 45 rue d'Ulm, 75230 Paris Cedex 05

E-mail address: yannick.baraud@ens.fr