ADAPTIVE GOODNESS-OF-FIT TESTS IN A DENSITY MODEL

BY MAGALIE FROMONT AND BÉATRICE LAURENT

Université Paris-Sud

Given an i.i.d. sample drawn from a density f, we propose to test that f equals some prescribed density f_0 or that f belongs to some translation/scale family. We introduce a multiple testing procedure based on an estimation of the $\mathbb{L}_2$-distance between f and f_0 or between f and the parametric family that we consider. For each sample size n, our test has level of significance α. In the case of simple hypotheses, we prove that our test is adaptive: it achieves the optimal rates of testing established by Ingster (2000) over various classes of smooth functions simultaneously. As for composite hypotheses, we obtain similar results up to a logarithmic factor. We carry out a simulation study to compare our procedures with the Kolmogorov-Smirnov tests, or with goodness-of-fit tests proposed by Bickel and Ritov (1992) and by Kallenberg and Ledwina (1995).

1. Introduction. Suppose that we observe n independent and identically distributed real random variables $X_1, \dots, X_n$ with common unknown density f. Let f_0 be some specified density. In this paper we consider the problem of testing the null hypothesis "$f \in \mathcal{F}$" against "$f \notin \mathcal{F}$" where $\mathcal{F}$ equals either the singleton $\{f_0\}$ or the parametric family

$$\mathcal{F} = \left\{ \frac{1}{\sigma} f_0 \left(\frac{. - \mu}{\sigma} \right), (\mu, \sigma) \in K \right\}, \tag{1}$$

for some subset K of $\mathbb{R} \times]0, +\infty[$.

This problem has widely been studied since the famous Kolmogorov-Smirnov (1933) and Cramér-von Mises (1928) tests based on the empirical distribution function.

Assuming that f belongs to $\mathbb{L}_2(\mathbb{R})$, it is quite natural to construct a test based on the estimation of the squared $\mathbb{L}_2$-distance between f and $\mathcal{F}$. In order to test the simple hypothesis "$f = f_0$", we actually consider a suitable collection of estimators of $\int_{\mathbb{R}} (f - f_0)^2$ and decide to reject the null hypothesis if some estimator in the collection is larger than its $(1 - u_\alpha)$ quantile under the null hypothesis, u_α being calibrated so that the final test has level of significance α. We then generalize this procedure to test that f belongs to the translation/scale family given by (1).

From a theoretical point of view, we evaluate the performances of our tests in terms of uniform separation rates with respect to the $\mathbb{L}_2$-distance over classes of smooth functions. Given β in $]0, 1[$ and a class of functions $\mathcal{B} \subset \mathbb{L}_2(\mathbb{R})$, we define the uniform separation rate $\rho(\Phi_\alpha, \mathcal{B}, \beta)$ of a level α test Φ_α of the null hypothesis "$f \in \mathcal{F}$" over the class $\mathcal{B}$ as the smallest number ρ such that the test guarantees a power at least equal to $(1 - \beta)$ for all alternatives f in $\mathcal{B}$ at a distance ρ from $\mathcal{F}$. More precisely, denoting by $d_2(f, \mathcal{F})$ the $\mathbb{L}_2$-distance between f and $\mathcal{F}$ and by $\mathbb{P}_f$ the distribution of the observation $(X_1, \dots, X_n)$,

$$\rho(\Phi_\alpha, \mathcal{B}, \beta) = \inf \left\{ \rho > 0, \ \forall f \in \mathcal{B}, \ d_2(f, \mathcal{F}) \geq \rho \Rightarrow \mathbb{P}_f \left(\Phi_\alpha \text{ rejects} \right) \geq 1 - \beta \right\}. \tag{2}$$

Assuming that f belongs to $\mathcal{B}$, the uniform separation rate $\rho(\Phi_\alpha, \mathcal{B}, \beta)$ is asymptotically related to the minimax rate of testing ρ_n introduced by Ingster (1993) and referred to as *critical radius*. Indeed, by definition, $\rho_n \to 0$ as $n \to +\infty$ and satisfies :

AMS 2000 subject classifications. Primary 62G10; Secondary 62G20.

Key words and phrases. Goodness-of-fit test, adaptive test, uniform separation rates, model selection.

- For any sequence ρ'_n such that $\rho'_n/\rho_n = o_n(1)$,

$$\inf_{\Phi_n} \left\{ \sup_{f \in \mathcal{F}} \mathbb{P}_f(\Phi_n = 1) + \sup_{f \in \mathcal{B}, d_2(f,\mathcal{F}) \geq \rho'_n} \mathbb{P}_f(\Phi_n = 0) \right\} = 1 - o_n(1),$$

  where the infimum is taken over all tests Φ_n with values in $\{0,1\}$ rejecting the null hypothesis "$f \in \mathcal{F}$" when $\Phi_n = 1$.

- For any α, $\beta > 0$, there exist some constant $C > 0$ and some test Φ_n^* such that

$$\sup_{f \in \mathcal{F}} \mathbb{P}_f(\Phi_n^* = 1) \leq \alpha + o_n(1) \quad \text{and} \quad \sup_{f \in \mathcal{B}, d_2(f,\mathcal{F}) \geq C\rho_n} \mathbb{P}_f(\Phi_n^* = 0) \leq \beta + o_n(1). \tag{3}$$

Since the goodness-of-fit test to some specified density f_0 can be reduced to a test of uniformity on $[0,1]$ for the variables $F_0(X_i)$'s (where F_0 is the distribution function associated with the density f_0), many papers are devoted to the problem of testing uniformity on $[0,1]$. The main reference for the computation of minimax rates of testing for this problem is the series of papers due to Ingster (1993). In particular, under the prior assumption that f belongs to some Hölder class with smoothness parameter $s \in]0,1]$, Ingster establishes the minimax rate of testing $\rho_n = n^{-2s/(4s+1)}$. However, the tests proposed to ensure the achievement of this rate (namely the inequalities given in (3)) are structurally based on the prior assumption: this is a crucial problem for their practical application since the smoothness parameter s is unknown. Following the work of Spokoiny (1998) in the Gaussian white noise model, Ingster (2000) focuses on the problem of finding an adaptive (assumption free) test of uniformity on $[0,1]$ and proves that adaptation is not possible without some loss of efficiency of the order of an extra $\log\log n$ factor. The adaptive test that he presents is based on chi-square statistics. It only has asymptotic properties and it has not been studied in practice. Other methods having Neyman's test as starting point are proposed in order to avoid using any prior assumption on the smoothness of f. To test uniformity on $[0,1]$, Neyman (1937) suggests considering some orthonormal basis $\{\phi_l,\ l \geq 0\}$ of $\mathbb{L}_2([0,1])$ with $\phi_0 = \mathbb{1}_{[0,1]}$ and rejecting the null hypothesis "$f = \mathbb{1}_{[0,1]}$" if the estimator $\sum_{l=1}^{D}(\sum_{i=1}^{n}\phi_l(X_i)/n)^2$ of $\theta_D = \sum_{l=1}^{D}(\int_{[0,1]} f\phi_l)^2$ is large enough, where D is some given integer. Bickel and Ritov (1992), Ledwina (1994), Kallenberg and Ledwina (1995) introduce data-driven versions of Neyman's test where the parameter D is chosen via some penalized criterion. Inglot and Ledwina (1996) establish theoretical results for the test described in Kallenberg and Ledwina (1995). They prove that, for a large set of alternatives, the asymptotic efficiency of the test with respect to the Neyman-Pearson test equals 1. Fan (1996) also proposes a new version of Neyman's test based on wavelet thresholding to test that the mean of a Gaussian vector equals 0 with applications to goodness-of-fit tests in a density model. When we test uniformity on $[0,1]$, our method amounts to considering, for all integer D in some set $\mathcal{D}_n$, the unbiased estimator of θ_D defined by

$$\hat{\theta}_D = \frac{1}{n(n-1)} \sum_{l=1}^{D} \sum_{i \neq j = 1}^{n} \phi_l(X_i)\phi_l(X_j)$$

and to penalize this estimator by its $(1-u_\alpha)$ quantile under the null hypothesis. The main difference between our method and the testing procedures proposed by previous authors lies in the order of magnitude of the penalty term. While Ledwina (1994) and Kallenberg and Ledwina (1995) choose the parameter D by using Schwarz's Bayesian Information Criterion (BIC), Kallenberg (2002) gives a discussion on the choice of the penalties for data-driven Neyman's tests. But the criteria considered in these papers have been introduced to estimate the density f itself whereas our penalties correspond to the ones used to build adaptive estimators of $\int_{\mathbb{R}} f^2$ by model selection in Laurent (2003). This choice allows us to obtain optimal uniform separation rates with respect to the $\mathbb{L}^2$-distance.

As for testing a composite null hypothesis, Pouet (2002) proves that provided that f belongs to $\mathbb{L}_2([0,1])$ and some Hölder class, the minimax rate of testing is comparable to the rate for the simple hypothesis "$f = \mathbb{1}_{[0,1]}$". However, the test proposed by Pouet depends on the smoothness assumption on f, which is not satisfactory from an experimental point of view. Inglot, Kallenberg and Ledwina (1997) introduce a procedure using no prior information about the smoothness of f to test composite hypotheses like "$f \in \{f(x,\beta),\ \beta \in B\}$" with $B \subset \mathbb{R}^q$. This procedure, generalizing Kallenberg and Ledwina's one (1995), also consists in a combination of Neyman's smooth test and Schwarz's selection rule. Its construction is based on the consideration of a sequence of exponential families with increasing dimensions to describe departures from the null model. The "right" dimension is selected by an extended Schwarz's rule which is obtained by inserting the maximum likelihood estimator $\hat{\beta}$ of β under the null hypothesis in the original definition of Schwarz's BIC. The next step is the application of Neyman's smooth test using a quadratic score statistic in the selected dimension. Inglot, Kallenberg and Ledwina (1997) prove the consistency of the test at essentially any alternative.

The approach considered in the present paper has been initiated by Baraud, Huet and Laurent (2003a), (2003b) and (2001) for the problem of testing linear or qualitative hypotheses in the Gaussian regression model. The properties of the testing procedures proposed here are non-asymptotic. For each n, the tests have the desired level of significance and we characterize some sets of alternatives over which they have a prescribed power. For the goodness-of-fit test to some specified density, we state in Section 2 that our procedure is adaptive over some collection of classes of smooth functions in the sense that it achieves the optimal "adaptive" rate of testing established by Ingster (2000) over all the classes of the collection simultaneously. We also investigate in Section 4 the test from a practical point of view by Monte-Carlo experiments. The results show that our procedure is competitive with the ones due to Bickel and Ritov (1992) or Kallenberg and Ledwina (1995). For the test of the hypothesis "$f \in \mathcal{F}$", where $\mathcal{F}$ is the translation/scale family defined by (1), we get in Section 3 uniform separation rates over classes of smooth alternatives of the same order (up to a logarithmic factor) as the rates obtained when testing the simple hypothesis "$f = f_0$". We finally implement the procedure to test exponentiality in Section 4: we can notice that it particularly gives good results under oscillating alternatives in comparison with the Kolmogorov-Smirnov test. The proofs of the results stated in the paper are detailed in Section 5.

2. A goodness-of-fit test. Let $X_1, \dots, X_n$ be i.i.d. random variables with common density f with respect to the Lebesgue measure on $\mathbb{R}$. Let f_0 be some given density in $\mathbb{L}_2(\mathbb{R})$ and let α in $]0,1[$. Assuming that f belongs to $\mathbb{L}_2(\mathbb{R})$, we construct a level α test of the null hypothesis "$f = f_0$" against the alternative "$f \neq f_0$" from the observation $(X_1, \dots, X_n)$.

In the following, $\|.\|_2$ and $\langle ., . \rangle$ respectively denote the usual norm and scalar product in $\mathbb{L}_2(\mathbb{R})$. For any bounded function g, $\|g\|_\infty = \sup_{x\in\mathbb{R}}|g(x)|$.

2.1. *Description of the test.* Our test is based on an estimation of $\|f - f_0\|_2^2$ that is of $\|f\|_2^2 + \|f_0\|_2^2 - 2\langle f, f_0\rangle$. Since $\langle f, f_0\rangle$ is usually estimated by the empirical estimator $\sum_{i=1}^n f_0(X_i)/n$, the key point is the estimation of $\|f\|_2^2$. As in Laurent (2003), we introduce some at most countable collection $\{S_m,\ m \in \mathcal{M}\}$ of linear subspaces of $\mathbb{L}_2(\mathbb{R})$. For all $m \in \mathcal{M}$, let $\{p_l, l \in \mathcal{L}_m\}$ be some orthonormal basis of S_m. The variable

$$\hat{\theta}_m = \frac{1}{n(n-1)} \sum_{l\in\mathcal{L}_m} \sum_{i\neq j=1}^{n} p_l(X_i)p_l(X_j) \tag{4}$$

is an unbiased estimator of $\|\Pi_{S_m}(f)\|_2^2$, where Π_{S_m} denotes the orthogonal projection onto S_m.

Then, $\|f-f_0\|_2^2$ can be estimated by

$$\widehat{T}_m=\hat{\theta}_m+\|f_0\|_2^2-\frac{2}{n}\sum_{i=1}^n f_0(X_i),$$

for any $m\in\mathcal{M}$. Denoting by $t_m(u)$ the $(1-u)$ quantile of the law of $\widehat{T}_m$ under the hypothesis "$f=f_0$" and considering

$$u_\alpha=\sup\left\{u\in]0,1[,\ \mathbb{P}_{f_0}\left(\sup_{m\in\mathcal{M}}\left(\widehat{T}_m-t_m(u)\right)>0\right)\leq\alpha\right\},\tag{5}$$

we introduce the test statistic T_α defined by

$$T_\alpha=\sup_{m\in\mathcal{M}}\left(\widehat{T}_m-t_m(u_\alpha)\right).$$

Our test consists in rejecting the null hypothesis if T_α is positive.

In practice, the values of u_α and the quantiles $\{t_m(u_\alpha),m\in\mathcal{M}\}$ are estimated by Monte-Carlo experiments under f_0 as explained in Section 4.

This method amounts to a multiple testing procedure. Indeed, for each m in $\mathcal{M}$, we construct a level u_α test of the null hypothesis "$f=f_0$" by rejecting this hypothesis if $\widehat{T}_m$ is larger than its $(1-u_\alpha)$ quantile under the hypothesis "$f=f_0$". We thus obtain a collection of tests and we decide to reject the null hypothesis if for some of the tests of the collection, this hypothesis is rejected.

2.2. *The power of the test.* Let us now precise the collection of linear subspaces $\{S_m,m\in\mathcal{M}\}$ that we use to define our testing procedure here. This collection is obtained by mixing spaces generated by constant piecewise functions, scaling functions and, in the case of compactly supported densities, trigonometric polynomials.
(i) For all $D\in\mathbb{N}^*$ and $k\in\mathbb{Z}$, let

$$I_{D,k}=\sqrt{D}\ \mathbb{1}_{[\frac{k}{D},\frac{k+1}{D}[}.$$

For all $D\in\mathbb{N}^*$, we define $S_{(1,D)}$ as the space generated by the functions $\{I_{D,k},\ k\in\mathbb{Z}\}$ and

$$\hat{\theta}_{(1,D)}=\frac{1}{n(n-1)}\sum_{k\in\mathbb{Z}}\sum_{i\neq j=1}^n I_{D,k}(X_i)I_{D,k}(X_j).$$

(ii) Let us consider a pair of compactly supported orthonormal wavelets (φ,ψ) such that for all $J\in\mathbb{N}$, $\{\varphi_{J,k}=2^{J/2}\varphi(2^J.-k),k\in\mathbb{Z}\}\cup\{\psi_{j,k}=2^{j/2}\psi(2^j.-k),j\in\mathbb{N},j\geq J,k\in\mathbb{Z}\}$ is an orthonormal basis of $\mathbb{L}_2(\mathbb{R})$. For all $J\in\mathbb{N}$ and $D=2^J$, we define $S_{(2,D)}$ as the space generated by the scaling functions $\{\varphi_{J,k},k\in\mathbb{Z}\}$ and

$$\hat{\theta}_{(2,D)}=\frac{1}{n(n-1)}\sum_{k\in\mathbb{Z}}\sum_{i\neq j=1}^n \varphi_{J,k}(X_i)\varphi_{J,k}(X_j).$$

(iii) Let us consider the Fourier basis of $\mathbb{L}^2([0,1])$ given by:

$$\begin{aligned}
g_0(x)&=\ \mathbb{1}_{[0,1]}(x),\\
g_{2p-1}(x)&=\sqrt{2}\cos(2\pi px)\ \mathbb{1}_{[0,1]}(x),\ \text{for all } p\geq 1,\\
g_{2p}(x)&=\sqrt{2}\sin(2\pi px)\ \mathbb{1}_{[0,1]}(x),\ \text{for all } p\geq 1.
\end{aligned}$$

For all $D\in\mathbb{N}^*$, we define $S_{(3,D)}$ as the space generated by the functions $\{g_l,l=0,\ldots,D\}$ and

$$\hat{\theta}_{(3,D)}=\frac{1}{n(n-1)}\sum_{l=0}^D\sum_{i\neq j=1}^n g_l(X_i)g_l(X_j).$$

We want here to notice that the constants involved in the following may depend on the chosen scaling function φ but we will not always precise it.

Introduce $\mathbb{D}_1 = \mathbb{D}_3 = \mathbb{N}^*$ and $\mathbb{D}_2 = \{2^J,\ J \in \mathbb{N}\}$. For $l \in \{1,2,3\}, D \in \mathbb{D}_l$, $\Pi_{S_{(l,D)}}$ denotes the orthogonal projection onto $S_{(l,D)}$ in $\mathbb{L}_2(\mathbb{R})$.

For all $l \in \{1,2,3\}$, we take $\mathcal{D}_l \subset \mathbb{D}_l$ with $\cup_{l\in\{1,2,3\}}\mathcal{D}_l \neq \emptyset$ and $\mathcal{D}_3 = \emptyset$ if the X_i's are not included in $[0,1]$. Let

$$\mathcal{M} = \{(l,D), l \in \{1,2,3\}, D \in \mathcal{D}_l\}.$$

For all $m \in \mathcal{M}$, we set

$$\widehat{T}_m = \hat{\theta}_m + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^n f_0(X_i).$$

The test statistic that we consider is

$$T_\alpha = \sup_{m\in\mathcal{M}}\left(\widehat{T}_m - t_m(u_\alpha)\right), \tag{6}$$

where $t_m(u_\alpha)$ is defined in Section 2.1.

The aim of the following theorem is to describe classes of alternatives over which the corresponding test has a prescribed power.

THEOREM 1. *Let $X_1,\dots,X_n$ be i.i.d. real random variables with common density f and let f_0 be some given density. Let T_α be the test statistic defined by (6). Assume that f_0 and f belong to $\mathbb{L}_\infty(\mathbb{R})$ and fix some β in $]0,1[$. For all $\varepsilon \in]0,2[$, there exist some positive constants $C_1(\beta)$ and $C_2(\beta,\varepsilon,\|f\|_\infty,\|f_0\|_\infty)$ such that, setting for all $m=(l,D)\in\mathcal{M}$,*

$$V_m(\beta) = \frac{C_1(\beta)}{n}\left((\sqrt{\|f\|_\infty} + \|f\|_\infty)\sqrt{D} + \frac{D}{n}\right) + \frac{C_2(\beta,\varepsilon,\|f\|_\infty,\|f_0\|_\infty)}{n},$$

if f satisfies

$$\|f-f_0\|_2^2 > (1+\varepsilon)\inf_{m\in\mathcal{M}}\left\{\|f-\Pi_{S_m}(f)\|_2^2 + t_m(u_\alpha) + V_m(\beta)\right\}, \tag{7}$$

then

$$\mathbb{P}_f\left(T_\alpha \le 0\right) \le \beta.$$

Comments:
(i) Let us see what is the advantage of considering a multiple testing procedure. We deduce from Theorem 1 that if we fix some element m in $\mathcal{M}$ and if we focus on the test that rejects the null hypothesis when $\widehat{T}_m$ is larger than its $(1-\alpha)$ quantile under the hypothesis "$f=f_0$" denoted by $t_m(\alpha)$, then the error of second kind of the test is smaller than β for all f such that

$$\|f-f_0\|_2^2 > (1+\varepsilon)\left\{\|f-\Pi_{S_m}(f)\|_2^2 + t_m(\alpha) + V_m(\beta)\right\}.$$

For the multiple testing procedure, the right hand side of the above inequality is replaced by its infimum over all m in $\mathcal{M}$, at the price that $t_m(\alpha)$ is replaced by $t_m(u_\alpha)$. When we evaluate the uniform separation rates, we show that for the collections $\{S_m,\ m\in\mathcal{M}\}$ that we have chosen, the quantities $t_m(\alpha)$ and $t_m(u_\alpha)$ just differ by a logarithmic factor. Therefore, the multiple testing procedure behaves almost as well as the best test among the considered collection of tests.
(ii) We prove in Section 5 that if $\mathcal{M}$ is finite, then for every $m=(l,D)$ in $\mathcal{M}$, $t_m(u_\alpha)$ is precisely of order $\sqrt{D\log(|\mathcal{M}|/\alpha)}/n$, where $|\mathcal{M}|$ denotes the cardinality of $\mathcal{M}$. This allows us to establish uniform separation rates over classes of alternatives for which the bias terms

$\|f - \Pi_{S_m}(f)\|_2^2$ are well controlled.
(iii) The key point of the proof of Theorem 1 is an exponential inequality for U-statistics of order 2 due to Houdré and Reynaud-Bouret (2002). The same result could also be obtained with an inequality due to Giné, Latala and Zinn (2000).

2.3. *Uniform separation rates.* Our purpose here is to evaluate the uniform separation rates of the test proposed above over several classes of alternatives. For $s > 0$, $R > 0$, $M > 0$ and $l \in \{1, 2, 3\}$, we introduce:

$$\mathcal{B}_s^{(l)}(R, M) = \left\{ f \in \mathbb{L}_2(\mathbb{R}),\ \forall D \in \mathbb{D}_l,\ \|f - \Pi_{S_{(l,D)}}(f)\|_2^2 \leq R^2 D^{-2s},\ \|f\|_\infty \leq M \right\}.$$

These sets of functions include some Hölder balls or Besov bodies. To be more precise, we consider, for all $s \in]0, 1]$ and $R > 0$, the Hölder ball $\mathcal{H}_s(R)$ defined by

$$\mathcal{H}_s(R) = \{ f : [0, 1] \to \mathbb{R}, \quad \forall x, y \in [0, 1],\ |f(x) - f(y)| \leq R|x - y|^s \}.$$

Let for $j \in \mathbb{N},\ k \in \mathbb{Z}, \beta_{j,k}(f) = \langle f, \psi_{j,k} \rangle$. For all $s > 0$ and $R > 0$, we define the Besov body $\mathcal{B}_{s,2,\infty}(R)$ as

$$\mathcal{B}_{s,2,\infty}(R) = \left\{ f \in \mathbb{L}_2(\mathbb{R}),\ \forall j \in \mathbb{N}, \sum_{k \in \mathbb{Z}} \beta_{j,k}^2(f) \leq R^2 2^{-2js} \right\}.$$

Then, one can see by straightforward computations that for $s \in]0, 1], R > 0, M > 0$,

$$\mathcal{H}_s(R) \cap \{f,\ \|f\|_\infty \leq M\} \subset \mathcal{B}_s^{(1)}(R, M)\ \cap\ \mathcal{B}_s^{(3)}(R/\sqrt{2(4^s - 1)}, M),$$

and for $s > 0, R > 0, M > 0$,

$$\mathcal{B}_{s,2,\infty}(R) \cap \{f,\ \|f\|_\infty \leq M\} \subset \mathcal{B}_s^{(2)}(R/\sqrt{1 - 4^{-s}}, M).$$

The following corollary gives upper bounds for the uniform separation rates of our testing procedure over the classes $\mathcal{B}_s^{(l)}(R, M)$.

Corollary 1. *Let T_α be the test statistic defined by (6). Assume that $n \geq 16$ and that for all $l \in \{1, 2, 3\}$, $\mathcal{D}_l$ equals $\{2^J, 0 \leq J \leq \log_2(n^2/(\log\log n)^3)\}$ or $\emptyset$. Fix some β in $]0, 1[$. For all $s > 0$, $M > 0$, $R > 0$ and $l \in \{1, 2, 3\}$ such that $\mathcal{D}_l \neq \emptyset$, there exists some positive constant $C = C(s, \alpha, \beta, M, \|f_0\|_\infty)$ such that if f belongs to the set $\mathcal{B}_s^{(l)}(R, M)$ and satisfies*

$$\|f - f_0\|_2^2 > C \left[R^{\frac{2}{4s+1}} \left(\frac{\sqrt{\log\log n}}{n} \right)^{\frac{4s}{4s+1}} + R^2 \left(\frac{(\log\log n)^3}{n^2} \right)^{2s} + \frac{(\log\log n)\log n}{n} \right],$$

then

$$\mathbb{P}_f\left(T_\alpha \leq 0\right) \leq \beta.$$

In particular, if

$$(\log\log n)^{s+1/2}(\log n)^{2s+1/2}/\sqrt{n} \leq R \leq n^{2s}/\left(\log\log n\right)^{3s+1/2}, \tag{8}$$

there exists some positive constant $C'(s, \alpha, \beta, M, \|f_0\|_\infty)$ such that the uniform separation rate of the test $\mathbb{1}_{T_\alpha > 0}$ over $\mathcal{B}_s^{(l)}(R, M)$ satisfies

$$\rho(\ \mathbb{1}_{T_\alpha > 0}, \mathcal{B}_s^{(l)}(R, M), \beta) \leq C'(s, \alpha, \beta, M, \|f_0\|_\infty) R^{\frac{1}{4s+1}} \left(\frac{\sqrt{\log\log n}}{n} \right)^{\frac{2s}{4s+1}}.$$

Comments:
For the problem of testing "$f = \mathbb{1}_{[0,1]}$" against the alternative "$f = \mathbb{1}_{[0,1]} + g$ with $g \neq 0$ and $g \in B_s(R)$" where $B_s(R)$ is a class of smooth functions (like some Hölder, Sobolev or Besov ball in $\mathbb{L}_2([0,1])$) with unknown smoothness parameter s, Ingster (2000) establishes that the adaptive minimax rate of testing is of order $(\sqrt{\log\log n}/n)^{\frac{2s}{4s+1}}$. From Corollary 1, we thus deduce that the procedure that we propose is adaptive in the sense that it is rate optimal over all the classes $\mathcal{B}_s^{(l)}(R, M)$ simultaneously, provided that R belongs to the range given by (8).

3. Testing a parametric family. Let $X_1, \dots, X_n$ be i.i.d. real random variables with common density f. In this section, we consider the problem of testing that f belongs to some translation/scale family of the form

$$\mathcal{F} = \left\{ \frac{1}{\sigma} f_0 \left(\frac{. - \mu}{\sigma} \right), (\mu, \sigma) \in K \right\},$$

where f_0 is some given density and K is some subset of $\mathbb{R} \times]0, +\infty[$. The families of Gaussian, uniform or exponential densities and translation models are typical examples of such translation/scale families.

3.1. *Description of the test.* The testing procedure that we introduce here is essentially based on the idea that if f belongs to $\mathcal{F}$, there exists (μ, σ) in K such that the density of the variables $(X_i - \mu)/\sigma$'s is f_0. As in Section 2.1, we take some at most countable collection $\{S_m, m \in \mathcal{M}\}$ of linear subspaces of $\mathbb{L}^2(\mathbb{R})$. For all m in $\mathcal{M}$, we consider an orthonormal basis $\{p_l, l \in \mathcal{L}_m\}$ of S_m composed of right-continuous functions and we denote by $\tilde{T}_m(X_1, \dots, X_n)$ the variable defined as

$$\inf_{(\mu,\sigma)\in K} \left\{ \frac{1}{n(n-1)} \sum_{l \in \mathcal{L}_m} \sum_{i \neq j = 1}^{n} p_l \left(\frac{X_i - \mu}{\sigma} \right) p_l \left(\frac{X_j - \mu}{\sigma} \right) + \|f_0\|_2^2 - \frac{2}{n} \sum_{i=1}^{n} f_0 \left(\frac{X_i - \mu}{\sigma} \right) \right\}. \quad (9)$$

Since the functions p_l are right-continuous, the infimum over (μ, σ) in K can be replaced by the infimum over (μ, σ) in $K \cap \mathbb{Q}^2$ so that $\tilde{T}_m(X_1, \dots, X_n)$ is a random variable.
We reject the null hypothesis "$f \in \mathcal{F}$" if

$$\tilde{T}_\alpha = \sup_{m \in \mathcal{M}} \left(\tilde{T}_m(X_1, \dots, X_n) - \tilde{q}_{m,\alpha} \right)$$

is positive, where $\{\tilde{q}_{m,\alpha},\ m \in \mathcal{M}\}$ is a family of positive numbers such that

$$\sup_{f \in \mathcal{F}} \mathbb{P}_f \left(\tilde{T}_\alpha > 0 \right) \leq \alpha. \quad (10)$$

Let us explain how we choose $\{\tilde{q}_{m,\alpha},\ m \in \mathcal{M}\}$. We distinguish two cases.
(i) The first one corresponds to the case where for every m in $\mathcal{M}$, the variable $\tilde{T}_m(X_1, \dots, X_n)$ defined by (9) satisfies

$$\forall\, (\mu, \sigma) \in K,\ \tilde{T}_m(X_1, \dots, X_n) = \tilde{T}_m \left(\frac{X_1 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma} \right). \quad (11)$$

This equality holds if, for instance, $K = \mathbb{R} \times]0, +\infty[$, $K = \{0\} \times]0, +\infty[$ or $K = \mathbb{R} \times \{1\}$. In this case, we take $\tilde{q}_{m,\alpha} = \tilde{t}_m(\tilde{u}_\alpha)$, where $\tilde{t}_m(u)$ is the $(1-u)$ quantile of $\tilde{T}_m(X_1, \dots, X_n)$ under the hypothesis "$f = f_0$", and $\tilde{u}_\alpha$ is taken as

$$\tilde{u}_\alpha = \sup \left\{ u \in]0,1[,\ \mathbb{P}_{f_0} \left(\sup_{m \in \mathcal{M}} \left(\tilde{T}_m(X_1, \dots, X_n) - \tilde{t}_m(u) \right) > 0 \right) \leq \alpha \right\}.$$

The quantities $\tilde{u}_\alpha$ and $\{\tilde{t}_m(\tilde{u}_\alpha), m \in \mathcal{M}\}$ can be estimated by Monte-Carlo experiments. Let us see how this choice leads to inequality (10). Under the null hypothesis, there exists

(μ, σ) in K such that the density of the variables $(X_i - \mu)/\sigma$'s is f_0. From equation (11), one can then deduce:

$$\begin{aligned}
&\mathbb{P}_f\left(\sup_{m\in\mathcal{M}}\left(\tilde{T}_m(X_1,\ldots,X_n) - \tilde{t}_m(\tilde{u}_\alpha)\right) > 0\right)\\
&= \mathbb{P}_f\left(\sup_{m\in\mathcal{M}}\left(\tilde{T}_m\left(\frac{X_1-\mu}{\sigma},\ldots,\frac{X_n-\mu}{\sigma}\right) - \tilde{t}_m(\tilde{u}_\alpha)\right) > 0\right)\\
&= \mathbb{P}_{f_0}\left(\sup_{m\in\mathcal{M}}\left(\tilde{T}_m(X_1,\ldots,X_n) - \tilde{t}_m(\tilde{u}_\alpha)\right) > 0\right),
\end{aligned}$$

and according to the definition of $\tilde{t}_m(\tilde{u}_\alpha)$, this probability is at most α.
(ii) The second one corresponds to the case where (11) is not satisfied. This occurs for instance if K is a compact set of $\mathbb{R}\times]0,+\infty[$. Here, we take $\tilde{q}_{m,\alpha} = t_m(u_\alpha)$ where, as in Section 2.1, $t_m(u)$ is the $(1-u)$ quantile of the variable

$$\hat{T}_m(X_1,\ldots,X_n) = \frac{1}{n(n-1)}\sum_{l\in\mathcal{L}_m}\sum_{i\neq j=1}^{n} p_l(X_i)p_l(X_j) + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^{n} f_0(X_i)$$

under the assumption that the variables $X_1,\ldots,X_n$ are i.i.d. with common density f_0, and u_α is defined as

$$u_\alpha = \sup\left\{u\in]0,1[,\ \mathbb{P}_{f_0}\left(\sup_{m\in\mathcal{M}}\left(\hat{T}_m(X_1,\ldots,X_n) - t_m(u)\right) > 0\right) \leq \alpha\right\}.$$

Inequality (10) also holds in this case: if f belongs to $\mathcal{F}$, there exists (μ,σ) in K such that $f = f_0((.-\mu)/\sigma)/\sigma$. By definition of $\tilde{T}_m(X_1,\ldots,X_n)$, one has the following inequality:

$$\tilde{T}_m(X_1,\ldots,X_n) \leq \hat{T}_m\left(\frac{X_1-\mu}{\sigma},\ldots,\frac{X_n-\mu}{\sigma}\right).$$

Hence,

$$\mathbb{P}_f(\tilde{T}_\alpha > 0) \leq \mathbb{P}_f\left(\sup_{m\in\mathcal{M}}\left(\hat{T}_m\left(\frac{X_1-\mu}{\sigma},\ldots,\frac{X_n-\mu}{\sigma}\right) - t_m(u_\alpha)\right) > 0\right).$$

Since the variables $(X_i-\mu)/\sigma$'s have f_0 as common density, it follows from the definition of u_α that the above quantity is smaller than α.

We shall remark that the choice of $\{\tilde{q}_{m,\alpha}, m\in\mathcal{M}\}$ proposed in (ii) may lead to a conservative procedure. It is therefore preferable to use the procedure proposed in (i) whenever Condition (11) holds.

3.2. *The power of the test.* In the following, we use the same notations as in Section 2.2.
(i) For all D in $\mathbb{D}_1 = \mathbb{N}^*$ and $m = (1, D)$, we define

$$\hat{T}_m(X_1,\ldots,X_n) = \frac{1}{n(n-1)}\sum_{k\in\mathbb{Z}}\sum_{i\neq j=1}^{n} I_{D,k}(X_i)I_{D,k}(X_j) + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^{n} f_0(X_i).$$

(ii) Choose the scaling function φ such that it satisfies a Lipschitz condition like

$$\forall x, y\in\mathbb{R},\ |\varphi(x)-\varphi(y)| \leq C_\varphi|x-y|.$$

For all $D = 2^J$ in $\mathbb{D}_2 = \{2^J, J\in\mathbb{N}\}$ and $m = (2, D)$, we define

$$\hat{T}_m(X_1,\ldots,X_n) = \frac{1}{n(n-1)}\sum_{k\in\mathbb{Z}}\sum_{i\neq j=1}^{n} \varphi_{J,k}(X_i)\varphi_{J,k}(X_j) + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^{n} f_0(X_i).$$

We recall that we do not specify the dependency on φ in the involved constants.

Let $\mathcal{D}_1 \subset \mathbb{D}_1$ and $\mathcal{D}_2 \subset \mathbb{D}_2$ such that $\mathcal{D}_1 \cup \mathcal{D}_2 \neq \emptyset$ and let

$$\mathcal{M} = \{(l, D), l \in \{1,2\}, D \in \mathcal{D}_l\}.$$

For all $m \in \mathcal{M}$, we set

$$\tilde{T}_m(X_1, \ldots, X_n) = \inf_{(\mu,\sigma)\in K} \hat{T}_m\left(\frac{X_1-\mu}{\sigma}, \ldots, \frac{X_n-\mu}{\sigma}\right).$$

We consider the test statistic

$$\tilde{T}_\alpha = \sup_{m\in\mathcal{M}} \left(\tilde{T}_m(X_1, \ldots, X_n) - \tilde{q}_{m,\alpha}\right) \tag{12}$$

where $\{\tilde{q}_{m,\alpha},\ m \in \mathcal{M}\}$ is a family of positive numbers satisfying (10) .

Theorem 2. *Let $X_1, \ldots, X_n$ be i.i.d. real random variables with common density $f \in \mathbb{L}_\infty(\mathbb{R})$. Let*

$$\mathcal{F} = \left\{\frac{1}{\sigma} f_0\left(\frac{.-\mu}{\sigma}\right), (\mu,\sigma) \in K\right\},$$

where f_0 is some given bounded density and $K = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}, \overline{\sigma}]$, $\underline{\mu}, \overline{\mu}, \underline{\sigma}, \overline{\sigma}$ being real numbers such that $\underline{\sigma} > 0$. Suppose that the following hypotheses hold:

(h_1) There exists some constant $C_{f_0} > 0$ such that for all x, y in the support of f, $(\mu,\sigma) \in K$, $(\mu',\sigma') \in K$,

$$\left| f_0\left(\frac{x-\mu}{\sigma}\right) - f_0\left(\frac{y-\mu'}{\sigma'}\right)\right| \leq C_{f_0}\left|\frac{x-\mu}{\sigma} - \frac{y-\mu'}{\sigma'}\right|.$$

(h_2) There exist $\nu > 0$ and $c > 0$ such that for all $k \geq 2$,

$$\mathbb{E}\left(|X_i|^k\right) \leq \frac{k!}{2}\nu c^{k-2}.$$

Let $\tilde{T}_\alpha$ be the test statistic defined by (12). Assume that $\mathcal{D}_2 \neq \emptyset$ and that n is large enough so that $n \geq 3$ and $n^2(\overline{\mu} - \underline{\mu}) \wedge (\overline{\sigma} - \underline{\sigma}) \geq 2$. Fix some $\beta \in]0,1[$. For all ε in $]0,2[$, there exists some positive constant $C = C(\underline{\mu}, \overline{\mu}, \underline{\sigma}, \overline{\sigma}, \|f_0\|_\infty, C_{f_0}, \|f\|_\infty, \nu, c, \beta, \varepsilon)$ such that, setting

$$\tilde{V}_D(\beta) = C\Big(\frac{\sqrt{D}}{n}\sqrt{\log(n^2D)} + \frac{D}{n^2}\log^2(n^2D) + \frac{\log(n^2D)}{n}\Big),$$

if f satisfies

$$\inf_{(\mu,\sigma)\in K} \sigma\left\| f - \frac{1}{\sigma}f_0\left(\frac{.-\mu}{\sigma}\right)\right\|_2^2 \geq$$

$$(1+\varepsilon)\inf_{D\in\mathcal{D}_2}\Big\{\sup_{(\mu,\sigma)\in K}\|\sigma f(\sigma.+\mu) - \Pi_{S_{(2,D)}}(\sigma f(\sigma.+\mu))\|_2^2 + \tilde{V}_D(\beta) + \tilde{q}_{(2,D),\alpha}\Big\},$$

then

$$\mathbb{P}_f(\tilde{T}_\alpha \leq 0) \leq \beta.$$

Comments:

(i) In Theorem 2, we only consider the case where K is compact. This is due to technical reasons: we have to evaluate the supremum over (μ,σ) in K of $U-$statistics of order 2 depending on (μ,σ). By considering a finite grid on the compact set K, we reduce the problem to a control of a finite number of these $U-$ statistics and we can use the inequality due to Houdré and Reynaud-Bouret (2002) again to control each of them.

(ii) Condition (h_1) is satisfied by families of Gaussian densities (when $f_0(x) = e^{-x^2/2}/\sqrt{2\pi}$) whatever the support of f and also by families of exponential densities (when $f_0(x) = e^{-x}\,\mathbb{1}_{x\geq 0}$ and $K \subset \{0\}\times]0,+\infty[$) when the support of f is included in $[0,+\infty[$. As for families of uniform densities (when $f_0 = \mathbb{1}_{[0,1]}$), Condition (h_1) is not satisfied but the result still holds: to see this, we refer to a theorem stated in Fromont (2003), where Condition (h_1) in Theorem 2 is replaced by some $\mathbb{L}_2$-entropy with bracketing condition on $\mathcal{F}$.

3.3. *Uniform separation rates.* As in Section 2.3, Theorem 2 allows us to evaluate the uniform separation rates of our test over classes of smooth functions. For all $s > 0$, $R > 0$, $M > 0$, we consider the set:

$$\tilde{\mathcal{B}}_s(R,M) = \{f \in \mathbb{L}_2(\mathbb{R}),\ \|f\|_\infty \leq M,\ \forall D \in \mathbb{D}_2,\ \forall(\mu,\sigma) \in K,$$
$$\|\sigma f(\sigma . + \mu) - \Pi_{S_{(2,D)}}(\sigma f(\sigma . + \mu))\|_2^2 \leq R^2 \sigma^{1+2s} D^{-2s}\}.$$

Such a set contains, among others, the functions f belonging to some Besov ball and satisfying the inequality $\|f\|_\infty \leq M$. To see this, in the notation of Devore and Lorentz (1993), we introduce, for all $h > 0$ and $r \in \mathbb{N}^*$,

$$\Delta_h^r(f,x) = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} f(x+kh).$$

The r-th modulus of smoothness of f in $\mathbb{L}_2(\mathbb{R})$ is defined by

$$\omega_r(f,t)_2 = \sup_{0 \leq h \leq t} \|\Delta_h^r(f,.)\|_2.$$

Then f belongs to the Besov ball $B_{s,2,\infty}(R)$ if for $r = [s]+1$,

$$\sup_{t>0} t^{-s} \omega_r(f,t)_2 \leq R.$$

One can easily see that

$$\omega_r(\sigma f(\sigma . + \mu), t)_2 = \sigma^{1/2} \omega_r(f, \sigma t)_2.$$

Let us now recall an inequality due to Devore, Jawerth and Popov (1992): if the wavelet ψ satisfies that for all $j < r$, $\int x^j \psi(x)dx = 0$, then, for every function g in $\mathbb{L}_2(\mathbb{R})$, for all $j \geq 0$,

$$\sum_{k \in \mathbb{Z}} \beta_{j,k}^2(g) \leq C \omega_r^2(g, 2^{-j})_2,$$

where C is an absolute constant. Hence, if f belongs to the Besov ball $B_{s,2,\infty}(R)$, for all $J \geq 0$, for all $(\mu,\sigma) \in K$,

$$\begin{aligned}\|\sigma f(\sigma . + \mu) - \Pi_{S_{(2,2^J)}}(\sigma f(\sigma . + \mu))\|_2^2 &= \sum_{j \geq J} \sum_{k \in \mathbb{Z}} \beta_{j,k}^2(\sigma f(\sigma . + \mu)) \\ &\leq C(1-4^{-s})^{-1} R^2 \sigma^{1+2s} 2^{-2Js}.\end{aligned}$$

If, in addition, $\|f\|_\infty \leq M$, then $f \in \tilde{\mathcal{B}}_s(C^{1/2}(1-4^{-s})^{-1/2}R, M)$.

The following corollary gives upper bounds for the uniform separation rates of the testing procedure over the classes $\tilde{\mathcal{B}}_s(R,M)$.

COROLLARY 2. *Assume that the conditions of Theorem 2 are satisfied. Let $\tilde{T}_\alpha$ be the test statistic defined by (12) with $\tilde{q}_{m,\alpha}$ equal to $t_m(u_\alpha)$ as explained in Section 3.1 (case (ii)). Choose $\mathcal{D}_1 \subset \{2^J, 0 \leq J \leq \log_2(n^2)\}$ and $\mathcal{D}_2 = \{2^J, 0 \leq J \leq \log_2(n^2/\log^3 n)\}$. Let $\beta \in]0,1[$. For all $s > 0$, $M > 0$, $R > 0$, there exists some positive constant $C = C(\underline{\mu}, \overline{\mu}, \underline{\sigma}, \overline{\sigma}, \|f_0\|_\infty, C_{f_0}, M, \nu, c, \alpha, \beta, s)$ such that, if f belongs to the set $\tilde{\mathcal{B}}_s(R,M)$ and satisfies*

$$\inf_{(\mu,\sigma)\in K} \left\| f - \frac{1}{\sigma} f_0\left(\frac{. - \mu}{\sigma}\right)\right\|_2^2 > C\left[R^{\frac{2}{4s+1}} \left(\frac{\sqrt{\log n}}{n}\right)^{\frac{4s}{4s+1}} + R^2 \left(\frac{(\log n)^3}{n^2}\right)^{2s} + \frac{\log n}{n}\right],$$

then

$$\mathbb{P}_f\left(\tilde{T}_\alpha \le 0\right) \le \beta.$$

Comments:
When R satisfies $(\log n)^{s+1/2}/\sqrt{n} \le R \le n^{2s}/(\log n)^{3s+1/2}$, the uniform separation rate over the class of functions belonging to $\tilde{\mathcal{B}}_s(R,M)$ and satisfying (h_1) and (h_2) is bounded from above by

$$C'(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\|f_0\|_\infty,C_{f_0}M,\nu,c,\alpha,\beta,s)R^{\frac{1}{4s+1}}\left(\frac{\sqrt{\log n}}{n}\right)^{\frac{2s}{1+4s}}.$$

This corresponds, up to a logarithmic factor, to the rate over the classes $\mathcal{B}_s^{(l)}(R,M)$ for the test of simple hypotheses obtained in Corollary 1.

4. Simulation study.

4.1. *Test of uniformity on* $[0,1]$. We first present simulation results for the problem of testing that the distribution of some i.i.d. random variables $X_1,\ldots,X_n$ with values in $[0,1]$ is uniform on $[0,1]$. In order to implement our procedure, we have to choose the set $\mathcal{M}=\{(l,D), l\in\{1,2,3\}, D\in\mathcal{D}_l\}$ that occurs in the definition (6) of the test statistic T_α. We examine two cases. In the first case, we only consider trigonometric polynomials. We take $\mathcal{D}_1=\mathcal{D}_2=\emptyset$ and $\mathcal{D}_3=\{1,2,\ldots D_{tr}\}$. Setting $g_0=\mathbb{1}_{[0,1]}$ and for all $p\ge 1$, $g_{2p-1}(x)=\sqrt{2}\cos(2p\pi x)\,\mathbb{1}_{[0,1]}(x)$, and $g_{2p}(x)=\sqrt{2}\sin(2p\pi x)\,\mathbb{1}_{[0,1]}(x)$, the test statistic is based on orthogonal projections onto the spaces spanned by the functions $\{g_l, l=0,\ldots D\}$ for D in $\mathcal{D}_3$. The second case consists in mixing trigonometric polynomials and constant piecewise functions: we take $\mathcal{D}_1=\{2,3,\ldots D_{ct}\}$ and $\mathcal{D}_3=\{1,\ldots D_{tr}\}$. The tests corresponding to those two cases are respectively denoted by $\mathcal{T}_{tr}$ and $\mathcal{T}_{tr/ct}$. We compare their powers with the powers of the tests proposed by Kallenberg and Ledwina (1995) (denoted by $\mathcal{T}_{KL}$), Bickel and Ritov (1992) (denoted by $\mathcal{T}_{BR}$) and Kolmogorov and Smirnov (denoted by $\mathcal{T}_{KS}$). As explained in the introduction, Kallenberg and Ledwina propose a test of uniformity on $[0,1]$ which is a data-driven version of Neyman's test (1937). They consider the orthonormal system $\{\phi_l,\ l\ge 0\}$ of $\mathbb{L}^2([0,1])$, where $\phi_0=\mathbb{1}_{[0,1]}$ and the ϕ_l's for $l\ge 1$ are the orthonormal Legendre polynomials on $[0,1]$. They decide to reject the null hypothesis if the statistic

$$T_D=\sum_{l=1}^{D}\left(n^{-1/2}\sum_{i=1}^{n}\phi_l(X_i)\right)^2$$

is large, D being chosen via Schwarz's Bayesian Information Criterion. The critical value is estimated by simulations. The test proposed by Bickel and Ritov is based on the statistic

$$T_n=\max_{1\le D\le d(n)}\left(\frac{T_{n,D}-D}{\sqrt{2D}}\right)$$

where

$$T_{n,D}=\frac{1}{n}\sum_{l=1}^{D}\sum_{i,j=1}^{n}2\cos(l\pi X_i)\cos(l\pi X_j).$$

We focus on the alternatives studied in the paper by Kallenberg and Ledwina (1995), for which the power of Bickel and Ritov's test is also given. These alternatives are the following ones:

$$f_{(\rho,j)}(x)=1+\rho\cos(j\pi x),$$
$$g_{(p,q,\varepsilon)}(x)=1-\varepsilon+\varepsilon\beta_{p,q}(x),$$
$$h_{(\rho,j)}(x)=1+\rho\phi_j(x),$$

where $\beta_{p,q}$ is the Beta density with parameter (p,q) and $\{\phi_j, j \geq 1\}$ is the family of orthonormal Legendre polynomials on $[0,1]$.

We have chosen a level $\alpha = 5\%$. The value of u_α and the quantiles $\{t_m(u_\alpha), m \in \mathcal{M}\}$ are estimated by 40000 simulations. We use 20000 simulations for the estimation of the $(1-u)$ quantiles $t_m(u)$ of the variables

$$\hat{T}_m = \hat{\theta}_m + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^n f_0(X_i)$$

under the hypothesis "$f = f_0$" for u varying on a regular grid of $]0,\alpha[$ and 20000 simulations for the estimation of the probabilities

$$\mathbb{P}_{f_0}\left(\sup_{m\in\mathcal{M}}\left(\hat{T}_m - t_m(u)\right) > 0\right).$$

The two following tables present the estimated powers for the tests $\mathcal{T}_{tr}, \mathcal{T}_{tr/ct}, \mathcal{T}_{KL}$, $\mathcal{T}_{BR}$, and $\mathcal{T}_{KS}$ under various alternatives for a number of observations equal to 50 or 100. The powers of the tests $\mathcal{T}_{tr}$, $\mathcal{T}_{tr/ct}$, and $\mathcal{T}_{KS}$ are estimated by 5000 experiments and the levels by 20000 experiments.

For a number of observations equal to 50, we choose $D_{tr} = 6$, $D_{ct} = 6$, and we obtain the following results:

Alternatives $f_{(\rho,j)}$					
(ρ,j)	$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KL}$	$\mathcal{T}_{BR}$	$\mathcal{T}_{KS}$
$(0.5,2)$	0.61	0.56	0.56	0.48	0.29
$(0.7,4)$	0.80	0.77	0.50	0.71	0.16
$(0.7,6)$	0.69	0.62	0.23	0.60	0.10
Alternatives $g_{(p,q,\varepsilon)}$					
(p,q,ε)	$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KL}$	$\mathcal{T}_{BR}$	$\mathcal{T}_{KS}$
$(3,3,1/2)$	0.55	0.49	0.53	0.40	0.14
$(10,20,0.25)$	0.46	0.49	0.36	0.41	0.33
$(2,2,0.8)$	0.62	0.55	0.63	0.44	0.15
$(2,4,0.5)$	0.57	0.60	0.55	0.58	0.64
Alternatives $h_{(\rho,j)}$					
(ρ,j)	$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KL}$	$\mathcal{T}_{BR}$	$\mathcal{T}_{KS}$
$(0.4,2)$	0.69	0.65	0.70	0.59	0.32
$(0.3,5)$	0.16	0.16	0.13	0.14	0.07
Estimated levels					
	$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$			$\mathcal{T}_{KS}$
	0.051	0.055			0.050

TABLE 1
Estimated powers of the test of uniformity on $[0,1]$ for $n = 50$ with $D_{tr} = 6$ and $D_{ct} = 6$

For a number of observations equal to 100, we choose $D_{tr=12}$, $D_{ct} = 10$, and we obtain:

Alternatives $f_{(\rho,j)}$					
(ρ, j)	$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KL}$	$\mathcal{T}_{BR}$	$\mathcal{T}_{KS}$
$(0.5, 2)$	0.87	0.85	0.87	0.84	0.53
$(0.7, 4)$	0.98	0.98	0.83	0.98	0.29
$(0.7, 6)$	0.97	0.96	0.46	0.95	0.19
Alternatives $g_{(p,q,\varepsilon)}$					
(p, q, ε)	$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KL}$	$\mathcal{T}_{BR}$	$\mathcal{T}_{KS}$
$(3, 3, 1/2)$	0.83	0.77	0.88	0.76	0.35
$(10, 20, 0.25)$	0.77	0.78	0.62	0.75	0.60
$(2, 2, 0.8)$	0.90	0.86	0.95	0.82	0.36
$(2, 4, 0.5)$	0.87	0.89	0.88	0.90	0.91
Alternatives $h_{(\rho,j)}$					
(ρ, j)	$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KL}$	$\mathcal{T}_{BR}$	$\mathcal{T}_{KS}$
$(0.4, 2)$	0.93	0.91	0.95	0.90	0.60
$(0.3, 5)$	0.33	0.31	0.23	0.33	0.09

Estimated levels		
$\mathcal{T}_{tr}$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$
0.050	0.048	0.054

TABLE 2

Estimated powers of the test of uniformity on $[0,1]$ for $n = 100$ with $D_{tr} = 12$ and $D_{ct} = 10$

Comments:

In this simulation study, the alternatives that we consider are of three kinds.
The alternatives $f_{(\rho,j)}$ correspond to the uniform density contaminated by a cosine function. They are favorable to our tests and Bickel and Ritov's one since those tests are based on trigonometric polynomials. It is therefore natural to compare our power results with the results of $\mathcal{T}_{BR}$. The main difference between the two procedures lies in the fact that $\hat{\theta}_{(3,D)}$ is an unbiased estimator of the squared $\mathbb{L}_2$-norm of the orthogonal projection of f onto $S_{(3,D)}$, whereas $(T_{n,D} - D)/(n-1)$ is an unbiased estimator of $\sum_{l=1}^{D}(\int f(x)\sqrt{2}\cos(l\pi x)dx)^2$ only under the null hypothesis. The consequent bias term under some alternative f may be of order D/n, which does not allow to establish optimal uniform separation rates for the test proposed by Bickel and Ritov. This explains why the power of $\mathcal{T}_{tr}$ improves the power of $\mathcal{T}_{BR}$ in all cases.
The alternatives $h_{(\rho,j)}$ are more favorable to the test due to Kallenberg and Ledwina since this test is based on Legendre polynomials. However, under the alternative $h_{(0.3,5)}$, the test $\mathcal{T}_{tr}$ improves the results of $\mathcal{T}_{KL}$ and under the alternative $h_{(0.4,2)}$, the estimated power of $\mathcal{T}_{tr}$ is comparable to the one of $\mathcal{T}_{KL}$.
Since the functions $g_{(p,q,\varepsilon)}$ correspond to some "neutral" alternatives, we can focus on them. When the number of observations is equal to 100, the powers of our tests $\mathcal{T}_{tr}$ and $\mathcal{T}_{tr/ct}$ are at least equivalent to the powers of $\mathcal{T}_{KL}$ and $\mathcal{T}_{BR}$ for half of the considered cases. As for the other cases, the procedures $\mathcal{T}_{tr}$ and $\mathcal{T}_{tr/ct}$ are still more powerful than $\mathcal{T}_{BR}$. For small sample sizes ($n = 50$), our test $\mathcal{T}_{tr}$ is always at least as powerful as the tests $\mathcal{T}_{KL}$ and $\mathcal{T}_{BR}$.

4.2. *Other goodness-of-fit tests.* We are now interested in testing that the density f of the random variables $X_1, \ldots, X_n$ is a given density f_0, with $f_0 \neq \mathbb{1}_{[0,1]}$. To test such a hypothesis, we have two possible procedures: the first one consists in testing directly from the sample $X_1, \ldots, X_n$ the null hypothesis "$f = f_0$" as explained in Section 2.1. The second one consists in testing that the common distribution of the variables $F_0(X_1), \ldots, F_0(X_n)$, where F_0 is the distribution function associated with the density f_0, is uniform on $[0,1]$. This approach is the one which is proposed in most papers. Whereas the two procedures are equivalent for the Kolmogorov-Smirnov test, they are not for our method based on the estimation of a $\mathbb{L}_2$-distance. Indeed, in our case, the first test is based on the estimation of $\|f - f_0\|_2^2$. Since the density of $F_0(X_1)$ is given by $h(x) = f(F_0^{-1}(x))/f_0(F_0^{-1}(x))$ when F_0 is one to one, the second test is based on the estimation of $\int_{[0,1]}(h(x)-1)^2 \mathrm{d}x = \|(f - f_0)/\sqrt{f_0}\|_2^2$. To compare in practice these two procedures, we have chosen to test that the density f is Gaussian with mean 0 and with variance 1 firstly, with variance 0.01 secondly.

The choices we have made in order to implement our procedures are the following ones: for the direct test from $X_1, \ldots, X_n$ denoted by $\mathcal{T}_d$, the set $\mathcal{M} = \{(l, D), l \in \{1,2,3\}, D \in \mathcal{D}_l\}$ is taken such that $\mathcal{D}_2 = \mathcal{D}_3 = \emptyset$ and $\mathcal{D}_1 = \{1, \ldots, 10\}$ and for the second test from $F_0(X_1), \ldots, F_0(X_n)$, we apply the test $\mathcal{T}_{tr/ct}$ described in Section 4.1 with $D_{ct} = 10$ and $D_{tr} = 12$. We also present the power results of the Kolmogorov-Smirnov test that we still denote by $\mathcal{T}_{KS}$.

We have taken a number of observations $n = 100$ and a level $\alpha = 5\%$. The quantiles are estimated as above with 40000 simulations, the powers of the tests with 5000 simulations and the levels with 20000 simulations. The alternatives that we have considered are the following ones:

$$
\begin{aligned}
f_m(x) &= \frac{1}{2m}\, \mathbb{1}_{[-m,m]},\\
g_{(m,\sigma^2)}(x) &= \frac{1}{2\sqrt{2\pi}\sigma}\left(e^{-\frac{(x-m)^2}{2\sigma^2}} + e^{-\frac{(x+m)^2}{2\sigma^2}}\right),\\
h_p(x) &= \frac{p}{2}e^{px}\, \mathbb{1}_{x<0} + \frac{p}{2}e^{-px}\, \mathbb{1}_{x\geq 0}.
\end{aligned}
$$

Test of normality $\mathcal{N}(0,1)$				**Test of normality $\mathcal{N}(0,0.01)$**			
Alternatives f_m							
m	$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$	m	$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$
2	0.96	0.92	0.62	0.17	0.93	0.64	0.24
1.8	0.66	0.66	0.36	0.16	0.87	0.71	0.14
$\sqrt{\pi/2}$	0.71	1	0.07	0.12	0.99	1	0.14
Alternatives $g_{(m,\sigma^2)}$							
(m,σ^2)	$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$	(m,σ^2)	$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$
$(1,1)$	0.80	0.98	0.77	$(0.1, 0.01)$	1	0.98	0.77
$(0.5,2)$	0.66	0.98	0.70	$(0.05, 0.015)$	0.91	0.77	0.35
$(1,2)$	0.97	1	0.97	$(0.05, 0.02)$	1	0.97	0.68
Alternatives h_p							
p	$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$	p	$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$
$2/\sqrt{2\pi}$	0.24	0.95	0.42	$20/\sqrt{2\pi}$	0.96	0.95	0.41
$3/2\sqrt{2\pi}$	0.85	1	0.96	$15/\sqrt{2\pi}$	1	1	0.96
Estimated levels							
	$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$		$\mathcal{T}_d$	$\mathcal{T}_{tr/ct}$	$\mathcal{T}_{KS}$
	0.052	0.051	0.053		0.053	0.055	0.053

TABLE 3
Estimated powers for $n = 100$

Comments:
The first objective of this simulation study is to compare our tests with the Kolmogorov-Smirnov test. The estimated power of the most powerful of our tests is larger than the one of $\mathcal{T}_{KS}$ for most of the considered alternatives. In these cases, the power difference is really significant as we can see for example for the alternatives f_m.
The second objective is to compare the tests $\mathcal{T}_d$ and $\mathcal{T}_{tr/ct}$. We can notice that we reject the null hypothesis more often with the test of uniformity from $F_0(X_1), \ldots, F_0(X_n)$ when we

test that the density is Gaussian with variance 1, whereas the direct test performs better when we test that the density is Gaussian with variance 0.01. As explained above, this is due to the fact that $\|(f-f_0)/\sqrt{f_0}\|_2^2$ is larger than $\|f-f_0\|_2^2$ when f_0 is the standard Gaussian density but smaller when f_0 is the Gaussian density with variance 0.01.

4.3. *Testing a parametric family.* We now implement the testing procedure described in Section 3 in order to test that the density f of the observations $X_1,\ldots,X_n$ is an exponential density or, in other words, that f belongs to the set of densities

$$\mathcal{F}=\left\{f,\ f(x)=\sigma^{-1}e^{-x/\sigma}\ \mathbb{1}_{x\geq 0},\ \sigma>0\right\}.$$

To simplify the implementation, we base our test statistic on constant piecewise functions instead of scaling functions. For all $D\in\{2,\ldots,10\}$, we define

$$\tilde{T}_{(1,D)}(X_1,\ldots,X_n)=\inf_{\sigma>0}\left\{\frac{D}{n(n-1)}\sum_{k\geq 0}\sum_{i\neq j=1}^{n}\mathbb{1}_{\left\{\frac{X_i}{\sigma}\in\left[\frac{k}{D},\frac{k+1}{D}\right[\right\}}\mathbb{1}_{\left\{\frac{X_j}{\sigma}\in\left[\frac{k}{D},\frac{k+1}{D}\right[\right\}}\right.$$
$$\left.+\|f_0\|_2^2-\frac{2}{n}\sum_{i=1}^{n}f_0\left(\frac{X_i}{\sigma}\right)\right\},$$

where $f_0(x)=e^{-x}\ \mathbb{1}_{x\geq 0}$. It is easy to see that for all $\sigma>0$,

$$\tilde{T}_{(1,D)}(X_1/\sigma,\ldots,X_n/\sigma)=\tilde{T}_{(1,D)}(X_1,\ldots,X_n).$$

Let $\mathcal{M}=\{(1,D),D=2,\ldots,10\}$. As explained in Section 3, we obtain a level α test as follows: denoting by $\tilde{t}_m(u)$ the $(1-u)$ quantile of $\tilde{T}_m(X_1,\ldots,X_n)$ under the hypothesis "$f=f_0$", and setting

$$\tilde{u}_\alpha=\sup\left\{u\in]0,1[,\ \mathbb{P}_{f_0}\left(\sup_{m\in\mathcal{M}}\left(\tilde{T}_m(X_1,\ldots,X_n)-\tilde{t}_m(u)\right)>0\right)\leq\alpha\right\},$$

we reject the null hypothesis if

$$\sup_{m\in\mathcal{M}}\left(\tilde{T}_m(X_1,\ldots,X_n)-\tilde{t}_m(\tilde{u}_\alpha)\right)>0.$$

The quantities $\tilde{u}_\alpha$ and $\{\tilde{t}_m(\tilde{u}_\alpha),m\in\mathcal{M}\}$ are estimated by Monte Carlo experiments: 20000 simulations are used to estimate the values of $\{\tilde{t}_m(u),m\in\mathcal{M}\}$ for u varying on a regular grid of $]0,\alpha[$, and 20000 simulations are used to estimate the value of $\tilde{u}_\alpha$.

We compare the performances of our test with the Kolmogorov-Smirnov test described below. Let $\bar{X}_n=n^{-1}\sum_{i=1}^n X_i$ and let $\hat{F}_n$ be the empirical distribution function defined for all t in $\mathbb{R}$ by $\hat{F}_n(t)=n^{-1}\sum_{i=1}^n\ \mathbb{1}_{X_i\leq t}$. Let F_σ be the distribution function associated with the exponential density with parameter σ^{-1}: $F_\sigma(t)=(1-e^{-t/\sigma})\ \mathbb{1}_{t\geq 0}$. Under the hypothesis that the distribution of the X_i's is exponential with parameter σ^{-1}, the law of the statistic $D_n=\sup_{t\in\mathbb{R}}|\hat{F}_n(t)-F_{\bar{X}_n}(t)|$ is free of the parameter σ. Let $d_{n,1-\alpha}$ denote the $(1-\alpha)$ quantile of D_n under the assumption that the X_i's have f_0 as common density. The Kolmogorov-Smirnov test of exponentiality consists in rejecting the null hypothesis "$f\in\mathcal{F}$" if $D_n>d_{n,1-\alpha}$.

We consider several alternatives defined for $x>0$ by:

$$\begin{aligned}
&g_p(x)=\tfrac{1}{2}\left(e^{-x}+(1+\sin(p\pi x))\ \mathbb{1}_{0<x<1}\right)\ (p\text{ even}), && t(x)=e^{-(\log x)^2/2}/\left(x\sqrt{2\pi}\right),\\
&h_p(x)=\tfrac{1}{2}\left(e^{-x}+(1+\cos(p\pi x))\ \mathbb{1}_{0<x<1}\right), && v(x)=\sqrt{x}e^{-x/2}/\left(2^{3/2}\Gamma(3/2)\right),\\
&k_{(p,q,\varepsilon)}(x)=(1-\varepsilon)e^{-x}+\varepsilon\beta_{p,q}(x), && w(x)=1.5x^{0.5}e^{-x^{1.5}},\\
&l_{(p,q,\varepsilon)}(x)=(1-\varepsilon)e^{-x}+\varepsilon\gamma_{p,q}(x),
\end{aligned}$$

where $\beta_{p,q}$ and $\gamma_{p,q}$ respectively denote the Beta density and the Gamma density with parameters (p,q).

For each alternative, the power of the test is estimated by 5000 simulations. The levels are estimated by 20000 simulations. We choose $n = 100$ and $\alpha = 5\%$.

Table 4 presents the estimated powers of our test denoted by $\mathcal{T}'_\alpha$ and of the Kolmogorov-Smirnov test denoted by $\mathcal{T}'_{KS}$.

Alternatives g_p		
p	$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
4	0.89	0.74
Alternatives h_p		
p	$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
4	0.71	0.60
1	1	0.90
Alternatives $k_{(p,q,\varepsilon)}$		
(p,q,ε)	$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
(10,20,0.25)	0.91	0.65
Alternatives $l_{(p,q,\varepsilon)}$		
(p,q,ε)	$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
(2,5,0.5)	0.53	0.28
(2,5,0.75)	0.89	0.60

Alternative t	
$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
0.75	0.45
Alternative v	
$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
0.67	0.65
Alternative w	
$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
0.97	0.98

Estimated levels	
$\mathcal{T}'_\alpha$	$\mathcal{T}'_{KS}$
0.053	0.051

TABLE 4
Estimated powers of the test of exponentiality for $n = 100$

Comments:
We can see in Table 4 that our test is not always much more powerful than the Kolmogorov-Smirnov one under very smooth alternatives like v or w which respectively correspond to the chi-square with three degrees of freedom or Weibull with parameter 1.5 distributions. However, under oscillating alternatives like g_p, h_p, $k_{(p,q,\varepsilon)}$ and $l_{(p,q,\varepsilon)}$ for which the Kolmogorov-Smirnov test is known to fail, our test performs much better. We could furthermore expect better results for our procedure with regular scaling functions instead of constant piecewise functions that we have chosen since they are very easy to implement.

5. Proofs.

5.1. *Proof of Theorem 1.* The main tool of the proof is the canonical decomposition of the U-statistics $\hat{\theta}_m$ defined in Section 2.2. We introduce the processes U_n and P_n defined by

$$U_n(h) = \frac{1}{n(n-1)} \sum_{i \neq j = 1}^{n} h(X_i, X_j), \quad P_n(h) = \frac{1}{n} \sum_{i=1}^{n} h(X_i).$$

We also define $P(h) = \langle h, f \rangle$. Using the same notations as in Section 2.2, let

$$\begin{cases} i) \text{ for all } D \in \mathbb{N}^*, \mathcal{L}_{(1,D)} = \{D\} \times \mathbb{Z} \text{ and } \{p_l, l \in \mathcal{L}_{(1,D)}\} = \{I_{D,k}, k \in \mathbb{Z}\}, \\ ii) \text{ for all } J \in \mathbb{N}, \mathcal{L}_{(2,2^J)} = \{J\} \times \mathbb{Z} \text{ and } \{p_l, l \in \mathcal{L}_{(2,2^J)}\} = \{\varphi_{J,k}, k \in \mathbb{Z}\}, \\ iii) \text{ for all } D \in \mathbb{N}^*, \mathcal{L}_{(3,D)} = \{0, 1, \ldots D\} \text{ and } \{p_l, l \in \mathcal{L}_{(3,D)}\} = \{g_l, l = 0 \ldots D\}. \end{cases} \tag{13}$$

By setting, for all $m \in \mathcal{M}$,

$$H_m(x,y) = \sum_{l \in \mathcal{L}_m} (p_l(x) - a_l)(p_l(y) - a_l),$$

with $a_l = \langle f, p_l \rangle$, we obtain the decomposition:

$$\hat{\theta}_m = U_n(H_m) + (P_n - P)(2\Pi_{S_m}(f)) + \|\Pi_{S_m}(f)\|_2^2.$$

Let us fix some β in $]0,1[$. Recalling that

$$\mathbb{P}_f\left(T_\alpha \leq 0\right) = \mathbb{P}_f\left(\sup_{m\in\mathcal{M}}\left(\hat{\theta}_m + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^n f_0(X_i) - t_m(u_\alpha)\right) \leq 0\right),$$

we have:

$$\mathbb{P}_f\left(T_\alpha \leq 0\right) \leq \inf_{m\in\mathcal{M}} \mathbb{P}_f\left(\hat{\theta}_m + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^n f_0(X_i) - t_m(u_\alpha) \leq 0\right).$$

Since $\|f - \Pi_{S_m}(f)\|_2^2 = \|f\|_2^2 - \|\Pi_{S_m}(f)\|_2^2$,

$$\mathbb{P}_f\Big(T_\alpha \leq 0\Big) \leq \inf_{m\in\mathcal{M}} \mathbb{P}_f\Big(U_n(H_m) + (P_n - P)(2\Pi_{S_m}(f) - 2f) + (P_n - P)(2f - 2f_0) \\ + \|f - f_0\|_2^2 \leq \|f - \Pi_{S_m}(f)\|_2^2 + t_m(u_\alpha)\Big). \quad (14)$$

We then need to control $U_n(H_m)$, $(P_n - P)(2\Pi_{S_m}(f) - 2f)$ and $(P_n - P)(2f - 2f_0)$ for every m in $\mathcal{M}$.

- Control of $U_n(H_m)$.

We use the following lemma which derives from an exponential inequality for U-statistics of order 2 due to Houdré and Reynaud-Bouret (2002).

Lemma 1. *Let $X_1, \ldots, X_n$ be i.i.d. real random variables with common density $f \in \mathbb{L}_\infty(\mathbb{R})$. Let $\mathbb{D}_1 = \mathbb{D}_3 = \mathbb{N}^*$ and $\mathbb{D}_2 = \{2^J, J \in \mathbb{N}\}$. For all $m = (l, D)$ with $l \in \{1,2,3\}$ and $D \in \mathbb{D}_l$, introduce $\{p_l, l \in \mathcal{L}_m\}$ defined as in (13) and*

$$Z_m = \frac{1}{n(n-1)}\sum_{i\neq j=1}^n H_m(X_i, X_j),$$

with

$$H_m(x,y) = \sum_{l\in\mathcal{L}_m}\left(p_l(x) - \langle f, p_l\rangle\right)\left(p_l(y) - \langle f, p_l\rangle\right).$$

There exists some positive constant C (depending only on φ) such that, for all $l \in \{1,2,3\}$, $D \in \mathbb{D}_l$, $x > 0$,

$$\mathbb{P}\left(|Z_{(l,D)}| > \frac{C}{n}\left(\sqrt{Dx}\left(\|f\|_\infty + \sqrt{\|f\|_\infty}\right) + \|f\|_\infty x + \frac{Dx^2}{n}\right)\right) \leq 5.6\, e^{-x}. \quad (15)$$

The proof of this lemma is postponed to the Appendix.

By setting $\lambda = \log(3/\beta)$ and $\tilde{\lambda} = \lambda + \log(5.6)$, Lemma 1 gives that there exists some constant $C > 0$ such that, for all $m = (l, D) \in \mathcal{M}$,

$$\mathbb{P}_f\Big(U_n(H_m) < -\frac{C}{n}\Big(\left(\|f\|_\infty + \sqrt{\|f\|_\infty}\right)\sqrt{D\tilde{\lambda}} + \|f\|_\infty\tilde{\lambda} + \frac{D\tilde{\lambda}^2}{n}\Big)\Big) \leq \frac{\beta}{3}. \quad (16)$$

We deduce from (14) and (16) that

$$\mathbb{P}_f\Big(T_\alpha \leq 0\Big) \leq \frac{\beta}{3} + \inf_{(l,D)\in\mathcal{M}}\Bigg\{\mathbb{P}_f\Bigg((P_n - P)(2\Pi_{S_{(l,D)}}(f) - 2f) + (P_n - P)(2f - 2f_0) + \|f - f_0\|_2^2 \\ \leq \|f - \Pi_{S_{l,D}}(f)\|_2^2 + t_{(l,D)}(u_\alpha) + \frac{C}{n}\left(\left(\|f\|_\infty + \sqrt{\|f\|_\infty}\right)\sqrt{D\tilde{\lambda}} + \|f\|_\infty\tilde{\lambda} + \frac{D\tilde{\lambda}^2}{n}\right)\Bigg)\Bigg\}. \quad (17)$$

• Control of $(P_n - P)(2\Pi_{S_m}(f) - 2f)$ and $(P_n - P)(2f - 2f_0)$.

We now use the following lemma due to Birgé and Massart (1998) which provides a special version of Bernstein's inequality.

LEMMA 2. *Let $X_1, \ldots, X_n$ be independent random variables satisfying the moments condition*

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(|X_i|^k\right) \leq \frac{k!}{2}\nu c^{k-2} \text{ for all } k \geq 2,$$

for some positive constants ν and c. Then, for any positive x,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}(X_i)) \geq \frac{\sqrt{2\nu x}}{\sqrt{n}} + \frac{cx}{n}\right) \leq e^{-x}.$$

In particular, if for all $i \in \{1, \ldots, n\}$, $|X_i| \leq b$ and $\mathbb{E}(X_i^2) \leq \nu$, the above inequality is satisfied with $c = b/3$.

It is easy to check that there exists some constant $C' > 0$ such that for all $l \in \{1, 2\}$, $D \in \mathbb{D}_l$,

$$|2\Pi_{S_{(l,D)}}(f)(X_i) - 2f(X_i)| \leq C'\|f\|_\infty.$$

Moreover, it is proved in Devore and Lorentz (1993) p. 269 that one can take C' such that for all $D \in \mathbb{D}_3$,

$$|2\Pi_{S_{(3,D)}}(f)(X_i) - 2f(X_i)| \leq C'\|f\|_\infty \log(D+1).$$

Since

$$\mathbb{E}\big(2\Pi_{S_m}(f)(X_i) - 2f(X_i)\big)^2 \leq 4\|f\|_\infty \|\Pi_{S_m}(f) - f\|_2^2,$$

we can deduce from Lemma 2 that for all $m = (l, D) \in \mathcal{M}$,

$$\mathbb{P}_f\left((P_n - P)(2\Pi_{S_m}(f) - 2f) < -2\sqrt{\|f\|_\infty}\|\Pi_{S_m}(f) - f\|_2\sqrt{\frac{2\lambda}{n}} - \frac{C'\lambda\|f\|_\infty \log(D+1)}{3n}\right) \leq \frac{\beta}{3},$$

By using the elementary inequality $2ab \leq 4a^2/\varepsilon + \varepsilon b^2/4$, we obtain that for $m = (l, D) \in \mathcal{M}$,

$$\mathbb{P}_f\left((P_n - P)(2\Pi_{S_m}(f) - 2f) + \frac{\varepsilon}{4}\|\Pi_{S_m}(f) - f\|_2^2 < -\left(\frac{8}{\varepsilon} + \frac{C'\log(D+1)}{3}\right)\frac{\|f\|_\infty\lambda}{n}\right) \leq \frac{\beta}{3}. \tag{18}$$

The control of $(P_n - P)(2f - 2f_0)$ is computed in the same way and we get:

$$\mathbb{P}_f\left((P_n - P)(2f - 2f_0) + \frac{\varepsilon}{4}\|f - f_0\|_2^2 < -\left(2(\frac{4}{\varepsilon} + \frac{1}{3})\|f\|_\infty + \frac{2}{3}\|f_0\|_\infty\right)\frac{\lambda}{n}\right) \leq \frac{\beta}{3}. \tag{19}$$

Finally, we deduce from (17), (18) and (19) that if there exists some $m = (l, D)$ in $\mathcal{M}$ such that

$$\begin{aligned}(1 - \frac{\varepsilon}{4})\|f - f_0\|_2^2 > (1 + \frac{\varepsilon}{4})\|f - \Pi_{S_m}(f)\|_2^2 + \frac{C}{n}\Big(\Big(\|f\|_\infty + \sqrt{\|f\|_\infty}\Big)\sqrt{D\tilde{\lambda}} + \|f\|_\infty\tilde{\lambda} + \frac{D\tilde{\lambda}^2}{n}\Big)\\ + \Big((\frac{16}{\varepsilon} + \frac{C'\log(D+1) + 2}{3})\|f\|_\infty + \frac{2}{3}\|f_0\|_\infty\Big)\frac{\lambda}{n} + t_m(u_\alpha),\end{aligned}$$

then

$$\mathbb{P}_f\left(T_\alpha \leq 0\right) \leq \beta.$$

This concludes the proof of Theorem 1.

5.2. *Proof of Corollary 1.* Assume that for all $l \in \{1,2,3\}$, $\mathcal{D}_l = \emptyset$ or $\{2^J, 0 \leq J \leq \log_2(n^2/(\log\log n)^3)\}$.

5.2.1. *An upper bound for* $t_m(u_\alpha)$, $m \in \mathcal{M}$.

PROPOSITION 1. *There exists some positive constant $C(\alpha)$, such that for all $m = (l, D) \in \mathcal{M}$,*

$$t_m(u_\alpha) \leq W_m(\alpha),$$

where

$$W_m(\alpha) = \frac{C(\alpha)}{n}\Big((\|f_0\|_\infty + \sqrt{\|f_0\|_\infty})\sqrt{D\log\log n} + \frac{D(\log\log n)^2}{n} + \|f_0\|_\infty(\log\log n)\log n\Big).$$

Proof. Recall that $t_m(u)$ denotes the $(1-u)$ quantile of the distribution of $\hat{T}_m$ under the hypothesis "$f = f_0$". We first notice that for $n \geq 16$, $|\mathcal{M}| \leq 3(1+\log_2 n^2)$. So, setting $\alpha_n = \alpha/(3(1+\log_2 n^2))$,

$$\begin{aligned}\mathbb{P}_{f_0}\Big(\sup_{m\in\mathcal{M}}(\hat{T}_m - t_m(\alpha_n)) > 0\Big) &\leq \sum_{m\in\mathcal{M}} \mathbb{P}_{f_0}\Big(\hat{T}_m - t_m(\alpha_n) > 0\Big)\\ &\leq \sum_{m\in\mathcal{M}} \frac{\alpha}{3(1+\log_2 n^2)}\\ &\leq \alpha.\end{aligned}$$

By definition of u_α, this implies that $\alpha_n \leq u_\alpha$ and for all $m \in \mathcal{M}$,

$$t_m(u_\alpha) \leq t_m(\alpha_n).$$

It thus remains to give an upper bound for $t_m(\alpha_n)$. Let $m = (l, D) \in \mathcal{M}$. We use the same notations as in the proof of Theorem 1 to obtain the decomposition

$$\hat{T}_m = U_n(H_m) + (P_n - P)(2\Pi_{S_m}(f)) - 2P_n(f_0) + \|f_0\|_2^2 + \|\Pi_{S_m}(f)\|_2^2.$$

Under the null hypothesis "$f = f_0$",

$$\hat{T}_m = U_n(H_m) + (P_n - P)(2\Pi_{S_m}(f_0) - 2f_0) - \|f_0\|_2^2 + \|\Pi_{S_m}(f_0)\|_2^2.$$

Since $\|\Pi_{S_m}(f_0) - f_0\|_2^2 = \|f_0\|_2^2 - \|\Pi_{S_m}(f_0)\|_2^2$, we obtain that, under "$f = f_0$",

$$\hat{T}_m = U_n(H_m) + (P_n - P)(2\Pi_{S_m}(f_0) - 2f_0) - \|\Pi_{S_m}(f_0) - f_0\|_2^2. \tag{20}$$

As in the proof of Theorem 1, we control $U_n(H_m)$ from Lemma 1 and $(P_n - P)(2\Pi_{S_m}(f_0) - 2f_0)$ from Lemma 2. We set $\lambda_n = \log(2/\alpha_n)$ and $\tilde{\lambda}_n = \lambda_n + \log(5.6)$.

On the one hand, Lemma 1 leads to

$$\mathbb{P}_{f_0}\Big(U_n(H_m) > \frac{C}{n}\Big(\sqrt{D\tilde{\lambda}_n}(\|f_0\|_\infty + \sqrt{\|f_0\|_\infty}) + \|f_0\|_\infty\tilde{\lambda}_n + \frac{D\tilde{\lambda}_n^2}{n}\Big)\Big) \leq \frac{\alpha_n}{2}. \tag{21}$$

On the other hand, since

$$|2(\Pi_{S_m}(f_0) - f_0)(X_i)| \leq C'\|f_0\|_\infty \log(D+1), \tag{22}$$

and

$$\mathbb{E}_{f_0}\,(2(\Pi_{S_m}(f_0) - f_0)(X_i))^2 \leq 4\|f_0\|_\infty\|\Pi_{S_m}(f_0) - f_0\|_2^2,$$

it follows from Lemma 2 that

$$\mathbb{P}_{f_0}\left((P_n - P)(2\Pi_{S_m}(f_0) - 2f_0) > 2\sqrt{\|f_0\|_\infty}\,\|\Pi_{S_m}(f_0) - f_0\|_2\sqrt{\frac{2\lambda_n}{n}} + \frac{C'\|f_0\|_\infty \lambda_n \log(D+1)}{3n}\right) \leq \frac{\alpha_n}{2}.$$

By using the inequality $2ab \leq a^2 + b^2$, and the fact that for $n \geq 16$, $\log(D+1) \leq \log(n^2+1)$, we obtain that there exists $C'' > 0$ such that:

$$\mathbb{P}_{f_0}\left((P_n - P)(2\Pi_{S_m}(f_0) - 2f_0) - \|\Pi_{S_m}(f_0) - f_0\|_2^2 > \frac{C''\|f_0\|_\infty \lambda_n \log n}{n}\right) \leq \frac{\alpha_n}{2}. \quad (23)$$

We derive from (20), (21), and (23) that

$$\mathbb{P}_{f_0}\left(\hat{T}_m > \frac{C}{n}\left(\sqrt{D\tilde{\lambda}_n}\,(\|f_0\|_\infty + \sqrt{\|f_0\|_\infty}) + \|f_0\|_\infty \tilde{\lambda}_n + \frac{D\tilde{\lambda}_n^2}{n}\right) + \frac{C''\|f_0\|_\infty \lambda_n \log n}{n}\right) \leq \alpha_n.$$

Finally, we notice that there exist some positive constants $c(\alpha)$ and $c'(\alpha)$ such that for $n \geq 3$, $\lambda_n \leq c(\alpha) \log\log n$ and $\tilde{\lambda}_n \leq c'(\alpha) \log\log n$, which completes the proof of Proposition 1.

5.2.2. *Uniform separation rates.* Let us fix β in $]0,1[$ and l in $\{1,2,3\}$ such that $\mathcal{D}_l = \{2^J, 0 \leq J \leq \log_2(n^2/(\log\log n)^3)\}$. From Theorem 1 and Proposition 1, we deduce that if f satisfies :

$$\|f - f_0\|_2^2 > (1+\varepsilon) \inf_{D \in \mathcal{D}_l} \left\{\|f - \Pi_{S_{(l,D)}}(f)\|_2^2 + W_{(l,D)}(\alpha) + V_{(l,D)}(\beta)\right\},$$

then

$$\mathbb{P}_f\left(T_\alpha \leq 0\right) \leq \beta.$$

It is thus a matter of giving an upper bound for

$$\inf_{D \in \mathcal{D}_l} \left\{\|f - \Pi_{S_{(l,D)}}(f)\|_2^2 + W_{(l,D)}(\alpha) + V_{(l,D)}(\beta)\right\},$$

when f belongs to some specified classes of functions. Recall that

$$\mathcal{B}_s^{(l)}(R,M) = \left\{f \in \mathbb{L}_2(\mathbb{R}),\ \forall D \in \mathbb{D}_l\,,\ \|f - \Pi_{S_{(l,D)}}(f)\|_2^2 \leq R^2 D^{-2s},\ \|f\|_\infty \leq M\right\}.$$

We now suppose that f belongs to $\mathcal{B}_s^{(l)}(R,M)$. Since $\|f - \Pi_{S_{(l,D)}}(f)\|_2^2 \leq R^2 D^{-2s}$ and since the constant $C_2(\beta, \varepsilon, \|f\|_\infty, \|f_0\|_\infty)$ in Theorem 1 can be taken such that $C_2(\beta, \varepsilon, \|f\|_\infty, \|f_0\|_\infty) \leq C_2(\beta, \varepsilon, M, \|f_0\|_\infty)$, we only need an upper bound for

$$\inf_{D \in \mathcal{D}_l} \left\{R^2 D^{-2s} + C_1(\beta)(\sqrt{M} + M)\frac{\sqrt{D}}{n} + C(\alpha)(\|f_0\|_\infty + \sqrt{\|f_0\|_\infty})\frac{\sqrt{D \log\log n}}{n}\right.$$
$$\left. + C_1(\beta)\frac{D}{n^2} + C(\alpha)\frac{D(\log\log n)^2}{n^2} + C(\alpha)\|f_0\|_\infty \frac{(\log\log n)\log n}{n} + \frac{C_2(\beta, \varepsilon, M, \|f_0\|_\infty)}{n}\right\}.$$

Assuming that $n \geq 16$, this quantity is bounded from above by

$$\inf_{D \in \mathcal{D}_l} \left\{R^2 D^{-2s} + \left(C_1(\beta)(\sqrt{M} + M) + C(\alpha)(\|f_0\|_\infty + \sqrt{\|f_0\|_\infty})\right)\frac{\sqrt{D \log\log n}}{n}\right.$$
$$\left. + (C_1(\beta) + C(\alpha))\frac{D(\log\log n)^2}{n^2}\right\} + (C(\alpha)\|f_0\|_\infty + C_2(\beta, \varepsilon, M, \|f_0\|_\infty)\,\frac{(\log\log n)\log n}{n}.$$

Since every D in $\mathcal{D}_l$ is smaller than $n^2/(\log\log n)^3$,

$$\inf_{D\in\mathcal{D}_l}\left\{R^2D^{-2s}+\frac{\sqrt{D\log\log n}}{n}+\frac{D(\log\log n)^2}{n^2}\right\}\le 2\inf_{D\in\mathcal{D}_l}\left\{R^2D^{-2s}+\frac{\sqrt{D\log\log n}}{n}\right\}.$$

We have $R^2D^{-2s}<\sqrt{D\log\log n}/n$ if and only if $D>\left(R^4n^2/\log\log n\right)^{1/(1+4s)}$, so we define D_* by

$$\log_2(D_*)=\left[\log_2\left(\left(\frac{R^4n^2}{\log\log n}\right)^{\frac{1}{1+4s}}\right)\right]+1,$$

and we consider three cases.
The first one is the case where $1\le D_*\le 2^{[\log_2(n^2/(\log\log n)^3)]}$, which means that $D_*\in\mathcal{D}_l$. In this case, we have that

$$\inf_{D\in\mathcal{D}_l}\left\{R^2D^{-2s}+\frac{\sqrt{D\log\log n}}{n}\right\}\le R^2D_*^{-2s}+\frac{\sqrt{D_*\log\log n}}{n}.$$

Since

$$R^2D_*^{-2s}\le R^{\frac{2}{4s+1}}\left(\frac{\sqrt{\log\log n}}{n}\right)^{\frac{4s}{4s+1}},$$

and

$$\frac{\sqrt{D_*\log\log n}}{n}\le\sqrt{2}\left(\frac{nR^2}{\sqrt{\log\log n}}\right)^{\frac{1}{4s+1}}\frac{\sqrt{\log\log n}}{n}\le\sqrt{2}R^{\frac{2}{4s+1}}\left(\frac{\sqrt{\log\log n}}{n}\right)^{\frac{4s}{4s+1}},$$

we obtain that

$$\inf_{D\in\mathcal{D}_l}\left\{R^2D^{-2s}+\frac{\sqrt{D\log\log n}}{n}\right\}\le(1+\sqrt{2})R^{\frac{2}{4s+1}}\left(\frac{\sqrt{\log\log n}}{n}\right)^{\frac{4s}{4s+1}}.$$

The second one is the case where $D_*>2^{[\log_2(n^2/(\log\log n)^3)]}$. In this case, for all D in $\mathcal{D}_l$,

$$\frac{\sqrt{D\log\log n}}{n}\le R^2D^{-2s}.$$

By taking $D_0=2^{\left[\log_2\left(n^2/(\log\log n)^3\right)\right]}$, we obtain that

$$\inf_{D\in\mathcal{D}_l}\left\{R^2D^{-2s}+\frac{\sqrt{D\log\log n}}{n}\right\}\le 2R^2D_0^{-2s}\le 2^{2s+1}R^2\left(\frac{(\log\log n)^3}{n^2}\right)^{2s}.$$

The third one is the case where $D_*<1$. In this case, for all D in $\mathcal{D}_l$, $R^2D^{-2s}\le\sqrt{D\log\log n}/n$, so by taking $D=1$, we obtain that

$$\inf_{D\in\mathcal{D}_l}\left\{R^2D^{-2s}+\frac{\sqrt{D\log\log n}}{n}\right\}\le\frac{\sqrt{\log\log n}}{n}.$$

This ends the proof of the corollary.

5.3. *Proof of Theorem 2.* Let $\tilde{T}_\alpha$ be the test statistic defined by (12). We have that

$$\mathbb{P}_f(\tilde{T}_\alpha \leq 0) \leq \inf_{m\in\mathcal{M}} \mathbb{P}_f(\tilde{T}_m(X_1,\ldots,X_n) \leq \tilde{q}_{m,\alpha}),$$

and since $\mathcal{D}_2 \neq \emptyset$, setting $\mathcal{M}' = \{(2,D),\ D\in\mathcal{D}_2\}$,

$$\mathbb{P}_f(\tilde{T}_\alpha \leq 0) \leq \inf_{m\in\mathcal{M}'} \mathbb{P}_f(\tilde{T}_m(X_1,\ldots,X_n) \leq \tilde{q}_{m,\alpha}). \tag{24}$$

Let us introduce a few notations. Using the definitions given in (13), for $m\in\mathcal{M}'$, we recall that

$$\hat{T}_m\Big(\frac{X_1-\mu}{\sigma},\ldots,\frac{X_n-\mu}{\sigma}\Big) = \frac{1}{n(n-1)}\sum_{l\in\mathcal{L}_m}\sum_{i\neq j=1}^{n} p_l\Big(\frac{X_i-\mu}{\sigma}\Big)p_l\Big(\frac{X_j-\mu}{\sigma}\Big) + \|f_0\|_2^2 - \frac{2}{n}\sum_{i=1}^{n} f_0\Big(\frac{X_i-\mu}{\sigma}\Big),$$

and we set

$$a_l(\mu,\sigma) = \int p_l\Big(\frac{x-\mu}{\sigma}\Big)f(x)dx,$$

$$b(\mu,\sigma) = \int f_0\Big(\frac{x-\mu}{\sigma}\Big)f(x)dx,$$

$$Z_m(\mu,\sigma) = \frac{1}{n(n-1)}\sum_{l\in\mathcal{L}_m}\sum_{i\neq j=1}^{n}\left[p_l\Big(\frac{X_i-\mu}{\sigma}\Big) - a_l(\mu,\sigma)\right]\left[p_l\Big(\frac{X_j-\mu}{\sigma}\Big) - a_l(\mu,\sigma)\right],$$

$$L_m(\mu,\sigma) = (P_n - P)\left(2\sum_{l\in\mathcal{L}_m} a_l(\mu,\sigma)p_l\Big(\frac{.-\mu}{\sigma}\Big) - 2f_0\Big(\frac{.-\mu}{\sigma}\Big)\right).$$

The following identity holds:

$$\hat{T}_m\Big(\frac{X_1-\mu}{\sigma},\ldots,\frac{X_n-\mu}{\sigma}\Big) = Z_m(\mu,\sigma) + L_m(\mu,\sigma) + \|f_0\|_2^2 + \sum_{l\in\mathcal{L}_m} a_l^2(\mu,\sigma) - 2b(\mu,\sigma).$$

Recall that S_m is the linear subspace of $\mathbb{L}^2(\mathbb{R})$ spanned by the functions $\{p_l, l\in\mathcal{L}_m\}$ and that Π_{S_m} denotes the orthogonal projection onto S_m in $\mathbb{L}^2(\mathbb{R})$. Then,

$$\sum_{l\in\mathcal{L}_m} a_l^2(\mu,\sigma) = \|\Pi_{S_m}(\sigma f(\sigma.+\mu))\|_2^2. \tag{25}$$

After some easy computations, one can prove that

$$\|f_0\|_2^2 + \sum_{l\in\mathcal{L}_m} a_l^2\mu,\sigma) - 2b(\mu,\sigma) = \sigma\Big\|\frac{1}{\sigma}f_0\Big(\frac{.-\mu}{\sigma}\Big) - f\Big\|_2^2 - \|\sigma f(\sigma.+\mu) - \Pi_{S_m}(\sigma f(\sigma.+\mu))\|_2^2.$$

This implies that

$$\mathbb{P}_f(\tilde{T}_m(X_1,\ldots,X_n) \leq \tilde{q}_{m,\alpha}) = \mathbb{P}_f\left(\inf_{(\mu,\sigma)\in K}\left\{Z_m(\mu,\sigma) + L_m(\mu,\sigma) + \sigma\Big\|\frac{1}{\sigma}f_0\Big(\frac{.-\mu}{\sigma}\Big) - f\Big\|_2^2 - \|\sigma f(\sigma.+\mu) - \Pi_{S_m}(\sigma f(\sigma.+\mu))\|_2^2\right\} \leq \tilde{q}_{m,\alpha}\right),$$

or:

$$\mathbb{P}_f(\tilde{T}_m(X_1,\dots,X_n) \le \tilde{q}_{m,\alpha}) = \mathbb{P}_f\Bigg(\inf_{(\mu,\sigma)\in K} \Bigg\{ Z_m(\mu,\sigma) + L_m(\mu,\sigma) + \frac{\varepsilon}{4}\sigma \Big\| \frac{1}{\sigma} f_0\Big(\frac{.-\mu}{\sigma}\Big) - f \Big\|_2^2 + \frac{\varepsilon}{4} \|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2^2 + \Big(1 - \frac{\varepsilon}{4}\Big) \sigma \Big\| \frac{1}{\sigma} f_0\Big(\frac{.-\mu}{\sigma}\Big) - f \Big\|_2^2 - \Big(1 + \frac{\varepsilon}{4}\Big) \|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2^2 \Bigg\} \le \tilde{q}_{m,\alpha} \Bigg).$$

Therefore,

$$\mathbb{P}_f(\tilde{T}_m \le \tilde{q}_{m,\alpha}) \le \mathbb{P}_f\Bigg(\sup_{(\mu,\sigma)\in K} \Bigg\{ - Z_m(\mu,\sigma) - L_m(\mu,\sigma) - \frac{\varepsilon}{4}\sigma \Big\| \frac{1}{\sigma} f_0\Big(\frac{.-\mu}{\sigma}\Big) - f \Big\|_2^2 - \frac{\varepsilon}{4} \|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2^2 \Bigg\} \ge (1 - \frac{\varepsilon}{4}) \inf_{(\mu,\sigma)\in K} \sigma \Big\| \frac{1}{\sigma} f_0\Big(\frac{.-\mu}{\sigma}\Big) - f \Big\|_2^2 - (1 + \frac{\varepsilon}{4}) \sup_{(\mu,\sigma)\in K} \|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2^2 - \tilde{q}_{m,\alpha} \Bigg).$$

Let $\tau_m(\beta)$ denote the $(1-\beta)$ quantile of the variable $\sup_{(\mu,\sigma)\in K} \Gamma_m(\mu,\sigma)$, where

$$\Gamma_m(\mu,\sigma) = -Z_m(\mu,\sigma) - L_m(\mu,\sigma) - \frac{\varepsilon}{4}\sigma \Big\| \frac{1}{\sigma} f_0\Big(\frac{.-\mu}{\sigma}\Big) - f \Big\|_2^2 - \frac{\varepsilon}{4} \|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2^2.$$

It follows from the above inequality and from (24) that if

$$(1 - \frac{\varepsilon}{4}) \inf_{(\mu,\sigma)\in K} \sigma \Big\| \frac{1}{\sigma} f_0\Big(\frac{.-\mu}{\sigma}\Big) - f \Big\|_2^2 > \inf_{m\in\mathcal{M}'} \Big((1 + \frac{\varepsilon}{4}) \sup_{(\mu,\sigma)\in K} \|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2^2 + \tau_m(\beta) + \tilde{q}_{m,\alpha} \Big),$$

then

$$\mathbb{P}_f(\tilde{T}_\alpha \le 0) \le \beta.$$

Let $m = (2, 2^J) \in \mathcal{M}'$. We now need to compute an upper bound for $\tau_m(\beta)$.

To compute this upper bound, we introduce a finite grid on $K = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}, \overline{\sigma}]$.
Let $\mu_0 < \mu_1 < \ldots < \mu_N$ and $\sigma_0 < \sigma_1 < \ldots < \sigma_N$ such that: $\mu_0 = \underline{\mu}$, $\mu_N = \overline{\mu}$, $\sigma_0 = \underline{\sigma}$, $\sigma_N = \overline{\sigma}$, for all $\delta \in \{0, \ldots, N-2\}$, $|\mu_{\delta+1} - \mu_\delta| = \Delta_\mu$, $|\sigma_{\delta+1} - \sigma_\delta| = \Delta_\sigma$ and $|\mu_N - \mu_{N-1}| \le \Delta_\mu$, $|\sigma_N - \sigma_{N-1}| \le \Delta_\sigma$. Δ_μ and Δ_σ will be chosen later. Let for $(\delta, \delta') \in \{0, \ldots, N-1\}^2$,

$$A_{\delta,\delta'} = [\mu_\delta, \mu_{\delta+1}] \times [\sigma_{\delta'}, \sigma_{\delta'+1}].$$

The following inequality holds:

$$\sup_{(\mu,\sigma)\in K} \Gamma_m(\mu,\sigma) \le \sup_{(\delta,\delta')\in\{0,\ldots,N-1\}^2} \Gamma_m(\mu_\delta, \sigma_{\delta'}) + \sup_{(\delta,\delta')\in\{0,\ldots,N-1\}^2} \sup_{(\mu,\sigma)\in A_{\delta,\delta'}} \left(\Gamma_m(\mu,\sigma) - \Gamma_m(\mu_\delta, \sigma_{\delta'})\right).$$

5.3.1. *Control of sup* $_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\Gamma_m(\mu_\delta,\sigma_{\delta'})$.

Since

$$-L_m(\mu,\sigma) = (P_n-P)\Big(2\sigma f - 2\Big(\Pi_{S_m}(\sigma f(\sigma.+\mu))\Big)\Big(\frac{.-\mu}{\sigma}\Big)\Big) + (P_n-P)\Big(2f_0\Big(\frac{.-\mu}{\sigma}\Big) - 2\sigma f\Big),$$

we have that

$$\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\Gamma_m(\mu_\delta,\sigma_{\delta'}) \le \sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}|Z_m(\mu_\delta,\sigma_{\delta'})| + \sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}L_m^{(1)}(\mu_\delta,\sigma_{\delta'}) + \sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}L_m^{(2)}(\mu_\delta,\sigma_{\delta'}),$$

with

$$L_m^{(1)}(\mu,\sigma) = (P_n-P)\left(2\sigma f - 2\Big(\Pi_{S_m}(\sigma f(\sigma.+\mu))\Big)\Big(\frac{.-\mu}{\sigma}\Big)\right) - \frac{\varepsilon}{4}\|\sigma f(\sigma.+\mu) - \Pi_{S_m}(\sigma f(\sigma.+\mu))\|_2^2$$

and

$$L_m^{(2)}(\mu,\sigma) = (P_n-P)\left(2f_0\Big(\frac{.-\mu}{\sigma}\Big) - 2\sigma f\right) - \frac{\varepsilon}{4}\sigma\left\|\frac{1}{\sigma}f_0\Big(\frac{.-\mu}{\sigma}\Big) - f\right\|_2^2.$$

- Control of sup $_{(\delta,\delta')\in\{0,\dots,N-1\}^2}|Z_m(\mu_\delta,\sigma_{\delta'})|$.

We apply Lemma 1 by replacing the variables X_i's by the variables $(X_i-\mu)/\sigma$'s and the density f by the density of the variables $(X_i-\mu)/\sigma$ that is $\sigma f(\sigma.+\mu)$. With $m=(2,2^J)$ and $\{p_l, l\in\mathcal{L}_m\} = \{\varphi_{J,k}, k\in\mathbb{Z}\}$, we obtain that there exists some constant $C>0$ such that for any $(\mu,\sigma)\in K$, for any $x>0$

$$\mathbb{P}\left(|Z_m(\mu,\sigma)| > \frac{C}{n}\left(2^{J/2}\sqrt{x}\left(\overline{\sigma}\|f\|_\infty + \sqrt{\overline{\sigma}\|f\|_\infty}\right) + \overline{\sigma}\|f\|_\infty x + \frac{2^J}{n}x^2\right)\right) \le 5.6e^{-x}.$$

Hence,

$$\mathbb{P}\left(\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}|Z_m(\mu_\delta,\sigma_{\delta'})| > \frac{C}{n}\left(2^{J/2}\sqrt{x}\left(\overline{\sigma}\|f\|_\infty + \sqrt{\overline{\sigma}\|f\|_\infty}\right) + \overline{\sigma}\|f\|_\infty x + \frac{2^J}{n}x^2\right)\right) \le 5.6N^2e^{-x}. \quad (26)$$

- Control of sup $_{(\delta,\delta')\in\{0,\dots,N-1\}^2}L_m^{(1)}(\mu_\delta,\sigma_{\delta'})$.

To get an upper bound for this supremum, we use Lemma 2.
We recall that for $m=(2,2^J)$, $\{p_l, l\in\mathcal{L}_m\} = \{\varphi_{J,k}, k\in\mathbb{Z}\}$, where $\varphi_{J,k} = 2^{J/2}\varphi(2^J.-k)$ and φ is a compactly supported scaling function. Then, for all $l\in\mathcal{L}_m$,

$$\|p_l\|_\infty \le 2^{J/2}\|\varphi\|_\infty.$$

Assume that the support of the function φ is included in $[-\Phi,\Phi]$, with $\Phi>0$. This implies that for every x in $\mathbb{R}$, the cardinality of the set $\{l\in\mathcal{L}_m, p_l(x)\neq 0\}$ is smaller than 2Φ and that for all $(\mu,\sigma)\in K$, for all $l\in\mathcal{L}_m$,

$$|a_l(\mu,\sigma)| \le 2\Phi\sigma 2^{-J/2}\|f\|_\infty\|\varphi\|_\infty.$$

Hence, for all $x \in \mathbb{R}$,

$$\left|\Big(\Pi_{S_m}(\sigma f(\sigma . + \mu))\Big)\Big(\frac{x-\mu}{\sigma}\Big)\right| = \left|\sum_{l \in \mathcal{L}_m} a_l(\mu,\sigma) p_l\Big(\frac{x-\mu}{\sigma}\Big)\right|$$
$$\leq 2\Phi\sigma 2^{-J/2}\|f\|_\infty\|\varphi\|_\infty \sum_{l\in\mathcal{L}_m}\left|p_l\Big(\frac{x-\mu}{\sigma}\Big)\right|$$
$$\leq 4\Phi^2\overline{\sigma}\|f\|_\infty\|\varphi\|_\infty^2.$$

Therefore, there exists some constant $C' > 0$ such that

$$\|2\sigma f - 2\Big(\Pi_{S_m}(\sigma f(\sigma . + \mu))\Big)\Big(\frac{.-\mu}{\sigma}\Big)\|_\infty \leq C'\overline{\sigma}\|f\|_\infty.$$

Furthermore, we have

$$\int\Big(2\sigma f - 2\Big(\Pi_{S_m}(\sigma f(\sigma . + \mu))\Big)\Big(\frac{.-\mu}{\sigma}\Big)\Big)^2 f \leq 4\overline{\sigma}\|f\|_\infty\|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2^2.$$

Then, for all $(\mu,\sigma) \in K$, Lemma 2 gives the following exponential inequality: for all $x > 0$,

$$\mathbb{P}_f\Bigg((P_n - P)\Big(2\sigma f - 2\Big(\Pi_{S_m}(\sigma f(\sigma . + \mu))\Big)\Big(\frac{.-\mu}{\sigma}\Big)\Big) > \frac{C'\overline{\sigma}\|f\|_\infty}{3}\frac{x}{n}$$
$$+ 2\sqrt{2\overline{\sigma}\|f\|_\infty}\|\sigma f(\sigma . + \mu) - \Pi_{S_m}(\sigma f(\sigma . + \mu))\|_2\sqrt{\frac{x}{n}}\Bigg) \leq e^{-x}.$$

By using the elementary inequality $2ab \leq 4a^2/\varepsilon + \varepsilon b^2/4$, we obtain that

$$\mathbb{P}_f\left(\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2} L_m^{(1)}(\mu_\delta,\sigma_{\delta'}) > \left(\frac{8}{\varepsilon}+\frac{C'}{3}\right)\overline{\sigma}\|f\|_\infty\frac{x}{n}\right) \leq N^2 e^{-x}. \tag{27}$$

- Control of $\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2} L_m^{(2)}(\mu_\delta,\sigma_{\delta'})$.

We control $\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2} L_m^{(2)}(\mu_\delta,\sigma_{\delta'})$ in the same way, noticing that for any $(\mu,\sigma) \in K$,

$$\left\|2f_0\Big(\frac{.-\mu}{\sigma}\Big) - 2\sigma f\right\|_\infty \leq 2(\|f_0\|_\infty + \overline{\sigma}\|f\|_\infty),$$

and

$$\int\Big(2f_0\Big(\frac{.-\mu}{\sigma}\Big) - 2\sigma f\Big)^2 f \leq 4\sigma^2\|f\|_\infty\left\|\frac{1}{\sigma}f_0\Big(\frac{.-\mu}{\sigma}\Big) - f\right\|_2^2.$$

We get:

$$\mathbb{P}_f\left(\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2} L_m^{(2)}(\mu_\delta,\sigma_{\delta'}) > 2\left(\frac{4}{\varepsilon}\overline{\sigma}\|f\|_\infty + \frac{\|f_0\|_\infty + \overline{\sigma}\|f\|_\infty}{3}\right)\frac{x}{n}\right) \leq N^2 e^{-x}. \tag{28}$$

Assume that $N \geq 2$. Collecting the inequalities (26), (27) and (28) and choosing x such that $7.6N^2e^{-x} = \beta/2$, we obtain that there exists some positive constant $C = C(\overline{\sigma},\varepsilon,\|f\|_\infty,\|f_0\|_\infty,\beta)$ such that

$$\mathbb{P}_f\left(\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\Gamma_m(\mu_\delta,\sigma_{\delta'}) > C\Big(2^{J/2}\frac{\sqrt{\log N}}{n} + \frac{\log N}{n} + 2^J\frac{\log^2 N}{n^2}\Big)\right) \leq \frac{\beta}{2}. \tag{29}$$

5.3.2. *Control of sup* $_{(\delta,\delta')\in\{0,\ldots,N-1\}^2}$ *sup* $_{(\mu,\sigma)\in A_{\delta,\delta'}}(\Gamma_m(\mu,\sigma)-\Gamma_m(\mu_\delta,\sigma_{\delta'}))$.

We have

$$\Gamma_m(\mu,\sigma) = -\frac{1}{n(n-1)}\sum_{l\in\mathcal{L}_m}\sum_{i\neq j=1}^{n} p_l\Big(\frac{X_i-\mu}{\sigma}\Big)p_l\Big(\frac{X_j-\mu}{\sigma}\Big) + \sum_{l\in\mathcal{L}_m} a_l^2(\mu,\sigma)$$
$$-\frac{\varepsilon}{4}\|\sigma f(\sigma.+\mu)-\Pi_{S_m}(\sigma f(\sigma.+\mu))\|_2^2-\frac{\varepsilon}{4}\sigma\Big\|\frac{1}{\sigma}f_0\Big(\frac{.-\mu}{\sigma}\Big)-f\Big\|_2^2+(P_n-P)\Big(2f_0\Big(\frac{.-\mu}{\sigma}\Big)\Big).$$

Using (25), this implies that

$$\Gamma_m(\mu,\sigma) = -\frac{1}{n(n-1)}\sum_{l\in\mathcal{L}_m}\sum_{i\neq j=1}^{n} p_l\Big(\frac{X_i-\mu}{\sigma}\Big)p_l\Big(\frac{X_j-\mu}{\sigma}\Big) - \frac{\varepsilon}{2}\sigma\|f\|_2^2-\frac{\varepsilon}{4}\|f_0\|_2^2$$
$$+(1+\frac{\varepsilon}{4})\sum_{l\in\mathcal{L}_m} a_l^2(\mu,\sigma)+\frac{\varepsilon}{2}b(\mu,\sigma)+(P_n-P)\Big(2f_0\Big(\frac{.-\mu}{\sigma}\Big)\Big).$$

Let

$$\Gamma_m^{(1)}(\mu,\sigma,\delta,\delta')=\frac{-1}{n(n-1)}\sum_{l\in\mathcal{L}_m}\sum_{i\neq j=1}^{n}\left(p_l\Big(\frac{X_i-\mu}{\sigma}\Big)p_l\Big(\frac{X_j-\mu}{\sigma}\Big)-p_l\Big(\frac{X_i-\mu_\delta}{\sigma_{\delta'}}\Big)p_l\Big(\frac{X_j-\mu_\delta}{\sigma_{\delta'}}\Big)\right),$$
$$\Gamma_m^{(2)}(\mu,\sigma,\delta,\delta') = \sum_{l\in\mathcal{L}_m} a_l^2(\mu,\sigma)-\sum_{l\in\mathcal{L}_m} a_l^2(\mu_\delta,\sigma_{\delta'}),$$
$$\Gamma_m^{(3)}(\mu,\sigma,\delta,\delta') = b(\mu,\sigma)-b(\mu_\delta,\sigma_{\delta'}),$$
$$\Gamma_m^{(4)}(\mu,\sigma,\delta,\delta') = (P_n-P)\Big(2f_0\Big(\frac{.-\mu}{\sigma}\Big)-2f_0\Big(\frac{.-\mu_\delta}{\sigma_{\delta'}}\Big)\Big).$$

Then,

$$\sup_{(\delta,\delta')\in\{0,\ldots,N-1\}^2}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}}(\Gamma_m(\mu,\sigma)-\Gamma_m(\mu_\delta,\sigma_{\delta'}))\le\sup_{(\delta,\delta')}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}}|\Gamma_m^{(1)}(\mu,\sigma,\delta,\delta')|$$
$$+(1+\frac{\varepsilon}{4})\sup_{(\delta,\delta')}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}}|\Gamma_m^{(2)}(\mu,\sigma,\delta,\delta')|+\frac{\varepsilon}{2}\sup_{(\delta,\delta')}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}}|\Gamma_m^{(3)}(\mu,\sigma,\delta,\delta')|$$
$$+\sup_{(\delta,\delta')}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}}|\Gamma_m^{(4)}(\mu,\sigma,\delta,\delta')|+\frac{\varepsilon}{2}\Delta_\sigma\|f\|_\infty.$$

- Control of sup $_{(\delta,\delta')\in\{0,\ldots,N-1\}^2}$sup $_{(\mu,\sigma)\in A_{\delta,\delta'}}|\Gamma_m^{(1)}(\mu,\sigma,\delta,\delta')|$.

One can easily see that

$$\Gamma_m^{(1)}(\mu,\sigma,\delta,\delta') = -\frac{1}{n(n-1)}\sum_{l\in\mathcal{L}_m}\sum_{i\neq j=1}^{n}\left[\left(p_l\Big(\frac{X_i-\mu}{\sigma}\Big)-p_l\Big(\frac{X_i-\mu_\delta}{\sigma_{\delta'}}\Big)\right)p_l\Big(\frac{X_j-\mu}{\sigma}\Big)\right.$$
$$\left.+\left(p_l\Big(\frac{X_j-\mu}{\sigma}\Big)-p_l\Big(\frac{X_j-\mu_\delta}{\sigma_{\delta'}}\Big)\right)p_l\Big(\frac{X_i-\mu_\delta}{\sigma_{\delta'}}\Big)\right].$$

We recall that for all $x,y\in\mathbb{R}$, the cardinality of the set $\{l\in\mathcal{L}_m, p_l(x)-p_l(y)\neq 0\}$ is not larger than 4Φ and that for $l\in\mathcal{L}_m$,

$$\|p_l\|_\infty\le 2^{J/2}\|\varphi\|_\infty.$$

Since φ is a Lipschitz function with Lipschitz constant C_φ, we also have that for all $x,y\in\mathbb{R}$,

$$\sum_{l\in\mathcal{L}_m}|p_l(x)-p_l(y)|\le 4\Phi 2^{3J/2}C_\varphi|x-y|. \tag{30}$$

This implies that

$$|\Gamma_m^{(1)}(\mu,\sigma,\delta,\delta')| \leq 8\Phi C_\varphi \|\varphi\|_\infty \frac{2^{2J}}{n}\sum_{i=1}^n \left|\frac{X_i-\mu}{\sigma} - \frac{X_i-\mu_\delta}{\sigma_{\delta'}}\right|,$$

and

$$\sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(1)}(\mu,\sigma,\delta,\delta')| \leq 8\Phi C_\varphi \|\varphi\|_\infty 2^{2J}\left\{\frac{1}{n}\sum_{i=1}^n |X_i|\left(\frac{\Delta_\sigma}{\underline{\sigma}^2}\right)+\frac{1}{\underline{\sigma}^2}\left(\overline{\sigma}\Delta_\mu + (|\overline{\mu}|\vee|\underline{\mu}|)\Delta_\sigma\right)\right\}.$$

Assuming that (h_2) holds, we derive from Lemma 2 that for all $x>0$,

$$\mathbb{P}_f\left(\sum_{i=1}^n (|X_i| - \mathbb{E}(|X_i|)) \geq \sqrt{2\nu n x} + cx\right) \leq e^{-x}. \tag{31}$$

Since $\mathbb{E}(|X_i|) \leq \sqrt{\nu}$ and since N is assumed to be larger than 2, by taking $x = \log(4N^2/\beta)$, we obtain that there exists some constant $C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\nu,c,\beta) > 0$ such that:

$$\mathbb{P}_f\left(\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(1)}(\mu,\sigma,\delta,\delta')| \geq C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\nu,c,\beta)2^{2J}(\Delta_\mu\vee\Delta_\sigma)\log N\right) \leq \frac{\beta}{4}. \tag{32}$$

- Control of $\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(2)}(\mu,\sigma,\delta,\delta')|$.

Since

$$\Gamma_m^{(2)}(\mu,\sigma,\delta,\delta') = \sum_{l\in\mathcal{L}_m} \left(a_l(\mu,\sigma) - a_l(\mu_\delta,\sigma_{\delta'})\right)\left(a_l(\mu,\sigma) + a_l(\mu_\delta,\sigma_{\delta'})\right),$$

and $|a_l(\mu,\sigma) + a_l(\mu_\delta,\sigma_{\delta'})| \leq 2.2^{J/2}\|\varphi\|_\infty$,

$$\left|\Gamma_m^{(2)}(\mu,\sigma,\delta,\delta')\right| \leq 2.2^{J/2}\|\varphi\|_\infty \int \sum_{l\in\mathcal{L}_m}\left|p_l\left(\frac{x-\mu}{\sigma}\right) - p_l\left(\frac{x-\mu_\delta}{\sigma_{\delta'}}\right)\right| f(x)dx.$$

By (30), we obtain:

$$\begin{aligned}\left|\Gamma_m^{(2)}(\mu,\sigma,\delta,\delta')\right| &\leq 8\Phi\|\varphi\|_\infty C_\varphi 2^{2J}\int \left|\frac{x-\mu}{\sigma} - \frac{x-\mu_\delta}{\sigma_{\delta'}}\right| f(x)dx,\\ &\leq 8\Phi\|\varphi\|_\infty C_\varphi 2^{2J}\int \left(|x|\frac{\Delta_\sigma}{\underline{\sigma}^2} + \frac{1}{\underline{\sigma}^2}\left(\overline{\sigma}\Delta_\mu + (|\overline{\mu}|\vee|\underline{\mu}|)\Delta_\sigma\right)\right) f(x)dx.\end{aligned}$$

We deduce from (h_2) that there exists some constant $C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\nu) > 0$ such that

$$\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(2)}(\mu,\sigma,\delta,\delta')| \leq C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\nu)2^{2J}(\Delta_\mu\vee\Delta_\sigma). \tag{33}$$

- Control of $\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(3)}(\mu,\sigma,\delta,\delta')|$.

Assuming that (h_1) holds, we have that for any (μ,σ) in $A_{\delta,\delta'}$, x in the support of f,

$$\left|f_0\Big(\frac{x-\mu}{\sigma}\Big) - f_0\Big(\frac{x-\mu_\delta}{\sigma_{\delta'}}\Big)\right| \leq C_{f_0}\left[|x|\left(\frac{\Delta_\sigma}{\underline{\sigma}^2}\right) + \frac{1}{\underline{\sigma}^2}\left(\overline{\sigma}\Delta_\mu + (|\overline{\mu}|\vee|\underline{\mu}|)\Delta_\sigma\right)\right]. \tag{34}$$

Hence, we derive from (h_2) that there exists some constant $C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},C_{f_0},\nu) > 0$ such that

$$\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(3)}(\mu,\sigma,\delta,\delta')| \leq C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},C_{f_0},\nu)(\Delta_\mu\vee\Delta_\sigma). \tag{35}$$

- Control of $\sup_{(\delta,\delta')\in\{0,\dots,N-1\}^2}\sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(4)}(\mu,\sigma,\delta,\delta')|$.

It follows from (34) that there exists $C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},C_{f_0})>0$ such that

$$\sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(4)}(\mu,\sigma,\delta,\delta')| \leq C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},C_{f_0}) \left[\frac{1}{n}\sum_{i=1}^n (|X_i|+\mathbb{E}(|X_i|)+1)(\Delta_\mu \vee \Delta_\sigma)\right].$$

Using again (31), we prove that there exists some constant $C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\nu,c,\beta,C_{f_0})>0$ such that

$$\mathbb{P}_f\left(\sup_{(\delta,\delta')\in\{0,\ldots,N-1\}^2}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}} |\Gamma_m^{(4)}(\mu,\sigma,\delta,\delta')| > C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\nu,c,\beta,C_{f_0})(\Delta_\mu\vee\Delta_\sigma)\log N\right) \leq \frac{\beta}{4}. \quad (36)$$

Collecting (32), (33), (35) and (36), we get

$$\mathbb{P}_f\left(\sup_{(\delta,\delta')\in\{0,\ldots,N-1\}^2}\ \sup_{(\mu,\sigma)\in A_{\delta,\delta'}} (\Gamma_m(\mu,\sigma)-\Gamma_m(\mu_\delta,\sigma_{\delta'})) > C2^{2J}(\Delta_\mu\vee\Delta_\sigma)\log N\right) \leq \frac{\beta}{2}, \quad (37)$$

for some constant $C=C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},C_{f_0},\|f\|_\infty,\nu,c,\beta,\varepsilon)>0$. Finally, by setting $\Delta_\mu=\Delta_\sigma=n^{-2}2^{-J}$, we have that

$$n^2 2^J(\overline{\mu}-\underline{\mu})\wedge(\overline{\sigma}-\underline{\sigma}) \leq N \leq n^2 2^J(\overline{\mu}-\underline{\mu})\vee(\overline{\sigma}-\underline{\sigma})+1.$$

Hence, we deduce from (29) and (37) that if $n^2(\overline{\mu}-\underline{\mu})\wedge(\overline{\sigma}-\underline{\sigma})\geq 2$ and $n\geq 2$, there exists some constant $C=C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},\|f_0\|_\infty,C_{f_0},\|f\|_\infty,\nu,c,\beta,\varepsilon)>0$ such that

$$\tau_m(\beta) \leq C\Big(\frac{2^{J/2}}{n}\sqrt{\log(n^2 2^J)}+\frac{2^J}{n^2}\log^2(n^2 2^J)+\frac{\log(n^2 2^J)}{n}\Big).$$

This concludes the proof of Theorem 2.

5.4. *Proof of Corollary 2.* Since $\tilde{q}_{(2,D),\alpha}=t_{(2,D)}(u_\alpha)$ with the same notations as in Section 2, we have that

$$\tilde{q}_{(2,D),\alpha} \leq \tilde{W}_D(\alpha),$$

where

$$\tilde{W}_D(\alpha)=\frac{C(\alpha)}{n}\Big((\|f_0\|_\infty+\sqrt{\|f_0\|_\infty})\sqrt{D\log\log n}+\frac{D(\log\log n)^2}{n}+\|f_0\|_\infty(\log\log n)\Big),$$

$C(\alpha)$ being some positive constant. This upper bound is obtained by replacing α_n by $\alpha/2(1+\log_2 n^2)$ in the proof of Proposition 1.

Let f belong to $\tilde{B}_s(R,M)$ and satisfy (h_1) and (h_2).
From Theorem 2 and the above upper bound, we deduce that there exists some positive constant $C=C(\underline{\mu},\overline{\mu},\underline{\sigma},\overline{\sigma},C_{f_0},\|f_0\|_\infty,M,\nu,c,\alpha,\beta,s)$ such that if f satisfies

$$\inf_{(\mu,\sigma)\in K}\Big\|f-\frac{1}{\sigma}f_0\Big(\frac{\cdot-\mu}{\sigma}\Big)\Big\|_2^2 \geq C\inf_{D\in\mathcal{D}_2}\Big\{R^2D^{-2s}+\frac{\sqrt{D\log(n^2D)}}{n}+\frac{D\log^2(n^2D)}{n^2}+\frac{\log(n^2D)}{n}\Big\},$$

then

$$\mathbb{P}_f(\tilde{T}_\alpha\leq 0)\leq\beta.$$

Since $\mathcal{D}_2 = \{2^J, 0 \leq J \leq 2^{[\log_2(n^2/\log^3 n)]}\}$ and $n \geq 3$, there exists $c > 0$ such that:

$$\inf_{D\in\mathcal{D}_2} \left\{ R^2 D^{-2s} + \frac{\sqrt{D\log(n^2 D)}}{n} + \frac{D\log^2(n^2D)}{n^2} + \frac{\log(n^2D)}{n} \right\}$$
$$\leq c\left(\inf_{D\in\mathcal{D}_2} \left\{ R^2D^{-2s} + \frac{\sqrt{D\log n}}{n} \right\} + \frac{\log n}{n} \right),$$

and Corollary 2 can be proved in the same way as Corollary 1.

6. Appendix: Proof of Lemma 1. We deduce from Theorem 3.4 in the paper by Houdré and Reynaud-Bouret (2002) that there exists some absolute constant $\kappa > 0$ such that for all $m = (l, D)$, $l \in \{1,2,3\}$, $D \in \mathbb{D}_l$, for all $x > 0$,

$$\mathbb{P}\left(|Z_m| > \frac{\kappa}{n^2}\left\{ A_1\sqrt{x} + A_2 x + A_3 x^{3/2} + A_4 x^2 \right\} \right) \leq 5.6\, e^{-x},$$

where

$$A_1^2 = n^2 \mathbb{E}(H_m^2(X_1,X_2)),$$
$$A_2 = \sup\left\{ \left| \mathbb{E}\left(\sum_{i=1}^n \sum_{j=1}^{i-1} H_m(X_1,X_2)\alpha_i(X_1)\beta_j(X_2) \right) \right|, \mathbb{E}\left(\sum_{i=1}^n \alpha_i^2(X_1) \right) \leq 1, \mathbb{E}\left(\sum_{j=1}^n \beta_j^2(X_1) \right) \leq 1 \right\},$$
$$A_3^2 = n \sup_x \left\{ \mathbb{E}_{X_2}(H_m^2(x, X_2)) \right\},$$
$$A_4 = \sup_{x,y} |H_m(x,y)|.$$

We therefore need to evaluate A_1, A_2, A_3 and A_4 for every $m = (l, D)$ with $l \in \{1,2,3\}$ and $D \in \mathbb{D}_l$.

6.1. *Evaluation in the case iii) (Fourier basis).* Let $m = (3, D)$ with $D \in \mathbb{D}_3$.

- Evaluation of A_1.

Setting for all $k \in \mathbb{N}$, $a_k = \langle f, g_k \rangle$, one has

$$H_m^2(X_1,X_2) = \sum_{k,k'=0}^{D} (g_k(X_1) - a_k)(g_{k'}(X_1) - a_{k'})(g_k(X_2) - a_k)(g_{k'}(X_2) - a_{k'}).$$

This implies that

$$\mathbb{E}\left(H_m^2(X_1,X_2)\right) = \sum_{k,k'=0}^{D} \left[\int (g_k(x) - a_k)(g_{k'}(x) - a_{k'}) f(x)dx \right]^2$$
$$= \sum_{k,k'=0}^{D} \left[\int g_k g_{k'} f - a_k a_{k'} \right]^2.$$

Using the inequality $(a-b)^2 \leq 2(a^2+b^2)$, one gets:

$$\begin{aligned}
\mathbb{E}\left(H_m^2(X_1,X_2)\right) &\leq 2\sum_{k,k'=0}^{D}\left(\int g_k g_{k'} f\right)^2 + 2\left(\sum_{k=0}^{D} a_k^2\right)^2 \\
&\leq 2\int\int\left(\sum_{k=0}^{D} g_k(x)g_k(y)\right)^2 f(x)f(y)dxdy + 2\left(\sum_{k=0}^{D} a_k^2\right)^2 \\
&\leq 2\|f\|_\infty^2 \int\int \sum_{k,k'=0}^{D} g_k(x)g_k(y)g_{k'}(x)g_{k'}(y)dxdy + 2\left(\sum_{k=0}^{D} a_k^2\right)^2 \\
&\leq 2\|f\|_\infty^2 (D+1) + 2\left(\sum_{k=0}^{D} a_k^2\right)^2 \text{ by orthogonality.}
\end{aligned}$$

Moreover, $\sum_{k=0}^{D} a_k^2 \leq \int f^2 \leq \|f\|_\infty$. This implies that

$$\mathbb{E}\left(H_m^2(X_1,X_2)\right) \leq 2(D+2)\|f\|_\infty^2 \leq 6D\|f\|_\infty^2,$$

and therefore

$$A_1 \leq n\sqrt{6D}\|f\|_\infty.$$

- Evaluation of A_2.

Let $(\alpha_i)_{1\leq i\leq n}$ and $(\beta_j)_{1\leq j\leq n}$ satisfy $\mathbb{E}(\sum_{i=1}^n \alpha_i^2(X_1)) \leq 1$ and $\mathbb{E}(\sum_{j=1}^n \beta_j^2(X_1)) \leq 1$. Setting $f_D = \sum_{k=0}^{D} a_k g_k$, one has

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{i-1}\mathbb{E}\left(H_m(X_1,X_2)\alpha_i(X_1)\beta_j(X_2)\right) &= \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=0}^{D}\int (g_k - a_k)\,\alpha_i f \int (g_k - a_k)\,\beta_j f \\
&= \sum_{i=1}^{n}\sum_{j=1}^{i-1}\left(\sum_{k=0}^{D}\int g_k\alpha_i f\int g_k\beta_j f + \sum_{k=0}^{D} a_k^2 \int\alpha_i f\int\beta_j f\right. \\
&\qquad \left. - \int\alpha_i f\int f_D\beta_j f - \int\beta_j f\int f_D\alpha_i f\right).
\end{aligned}$$

Using repeatedly the Cauchy-Schwarz inequality and the fact that for all $g \in \mathbb{L}^2([0,1])$, $\sum_{k=0}^{D}(\int g_k g)^2 \leq \int g^2$, one obtains that

$$\begin{aligned}
\left|\sum_{k=0}^{D}\int g_k\alpha_i f\int g_k\beta_j f\right| &\leq \left[\sum_{k=0}^{D}\left(\int g_k\alpha_i f\right)^2\right]^{1/2}\left[\sum_{k=0}^{D}\left(\int g_k\beta_j f\right)^2\right]^{1/2} \\
&\leq \left(\int \alpha_i^2 f^2\right)^{1/2}\left(\int \beta_j^2 f^2\right)^{1/2} \\
&\leq \|f\|_\infty \left(\int \alpha_i^2 f\right)^{1/2}\left(\int \beta_j^2 f\right)^{1/2}.
\end{aligned}$$

Therefore,

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\left|\sum_{k=0}^{D}\int g_k\alpha_i f\int g_k\beta_j f\right| \le \|f\|_\infty \sum_{i=1}^{n}\left(\int \alpha_i^2 f\right)^{1/2}\sum_{j=1}^{n}\left(\int \beta_j^2 f\right)^{1/2}$$
$$\le n\|f\|_\infty\left(\sum_{i=1}^{n}\int \alpha_i^2 f\right)^{1/2}\left(\sum_{j=1}^{n}\int \beta_j^2 f\right)^{1/2}$$
$$\le n\|f\|_\infty .$$

Moreover, still using the Cauchy-Schwarz inequality, the assumption on the α_i's and β_j's, together with the fact that $\sum_{k=0}^{D} a_k^2 \le \int f^2 \le \|f\|_\infty$, one gets

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=0}^{D} a_k^2\left|\int \alpha_i f\int \beta_j f\right| \le \|f\|_\infty \sum_{i=1}^{n}\left(\int \alpha_i^2 f\int f\right)^{1/2}\sum_{j=1}^{n}\left(\int \beta_j^2 f\int f\right)^{1/2}$$
$$\le \|f\|_\infty n.$$

Finally, since $\int f_D^2 \le \int f^2 \le \|f\|_\infty$,

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\int \alpha_i f\int f_D\beta_j f\right| \le \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\int \alpha_i^2 f\right)^{1/2}\left(\int f_D^2 f\right)^{1/2}\left(\int \beta_j^2 f\right)^{1/2}$$
$$\le \sum_{i=1}^{n}\sum_{j=1}^{n}\|f\|_\infty\left(\int \alpha_i^2 f\right)^{1/2}\left(\int \beta_j^2 f\right)^{1/2}$$
$$\le \|f\|_\infty n.$$

In the same way,

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\int \beta_j f\int f_D\alpha_i f\right| \le \|f\|_\infty n,$$

which gives:

$$A_2 \le 4n\|f\|_\infty .$$

- Evaluation of A_3.

$$\mathbb{E}_{X_2}(H_m^2(x,X_2)) = \sum_{k,k'=0}^{D}\left(\int g_k g_{k'} f\right)(g_k(x)-a_k)(g_{k'}(x)-a_{k'}) - \left(\sum_{k=0}^{D} a_k(g_k(x)-a_k)\right)^2$$
$$\le \sum_{k,k'=0}^{D}\left(\int g_k g_{k'} f\right)(g_k(x)-a_k)(g_{k'}(x)-a_{k'})$$
$$\le \|f\|_\infty \int \sum_{k,k'=0}^{D} g_k(y)g_{k'}(y)(g_k(x)-a_k)(g_{k'}(x)-a_{k'})dy$$
$$\le \|f\|_\infty \sum_{k=0}^{D}(g_k(x)-a_k)^2 \quad \text{by orthogonality.}$$

Hence,

$$\sup_x \mathbb{E}_{X_2}(H_m^2(x,X_2)) \le 2\|f\|_\infty \sum_{k=0}^{D}\left(\|g_k\|_\infty^2 + a_k^2\right).$$

Since $||g_k||_\infty = \sqrt{2}$ and

$$\sum_{k=0}^{D} a_k^2 = \sum_{k=0}^{D} (\int f g_k)^2 \leq \sum_{k=0}^{D} ||g_k||_\infty^2 \leq 2(D+1),$$

one gets:

$$A_3 \leq 4\sqrt{Dn||f||_\infty}.$$

- Evaluation of A_4.

Since for all $k \in \mathbb{N}$, $||g_k||_\infty \leq \sqrt{2}$, it is easy to see that $A_4 \leq 8(D+1)$.

Collecting the above evaluations, one obtains that there exists some positive, absolute constant C such that for all $D \in \mathbb{N}^*$,

$$\mathbb{P}\left(|Z_{(3,D)}| > \frac{C}{n}\left\{\sqrt{Dx}||f||_\infty + ||f||_\infty x + \sqrt{D||f||_\infty}\frac{x^{3/2}}{\sqrt{n}} + \frac{Dx^2}{n}\right\}\right) \leq 5.6e^{-x}.$$

It follows from the elementary inequality $2ab \leq a^2 + b^2$ that

$$2\sqrt{D||f||_\infty}\frac{x^{3/2}}{\sqrt{n}} \leq ||f||_\infty x + \frac{Dx^2}{n},$$

which allows to conclude.

6.2. *Proof in the cases i) and ii) (Constant piecewise functions and scaling functions).* We can write the same proof for the two cases by setting

$$g_{k,D} = \sqrt{D}\phi(Dx-k), \quad a_{k,D} = \int g_{k,D} f,$$

with $\phi = \mathbb{I}_{[0,1]}$ for the case $i)$, $\phi = \varphi$ and $D = 2^J$ for the case $ii)$. Let $m = (1, D)$ with $D \in \mathbb{D}_1$ or $m = (2, D)$ with $D \in \mathbb{D}_2$. All along the proof, $C(\phi)$ denotes some constant depending on ϕ that may vary from line to line.

- Evaluation of A_1.

By the same arguments as for the case $iii)$,

$$\mathbb{E}\left(H_m^2(X_1, X_2)\right) \leq 2 \sum_{k,k' \in \mathbb{Z}} \left(\int g_{k,D} g_{k',D} f\right)^2 + 2\left(\sum_{k \in \mathbb{Z}} a_{k,D}^2\right)^2.$$

Since ϕ is compactly supported, we can assume that the support of ϕ is included in $[a_\phi, b_\phi]$, where a_ϕ and b_ϕ are integers. Let

$$\mathcal{I}_{k,D} = [(a_\phi + k)/D, (b_\phi + k)/D].$$

We shall use the fact that

$$|k - k'| > b_\phi - a_\phi \text{ implies } \mathcal{I}_{k,D} \cap \mathcal{I}_{k',D} = \emptyset, \tag{38}$$

and that the support of the function $g_{k,D}$ is included in $\mathcal{I}_{k,D}$. This implies that

$$\sum_{k,k' \in \mathbb{Z}} \left(\int g_{k,D} g_{k',D} f\right)^2 \leq D^2 ||\phi||_\infty^4 \sum_{k \in \mathbb{Z}} \left(\int_{\mathcal{I}_{k,D}} f\right) \sum_{k' \in \mathbb{Z}, |k'-k| \leq b_\phi - a_\phi} \left(\int_{\mathcal{I}_{k',D}} f\right).$$

Moreover, we derive from the definition of the intervals $\mathcal{I}_{k,D}$ that for any function g in $\mathbb{L}_1(\mathbb{R})$,

$$\sum_{k \in \mathbb{Z}} \left(\int_{\mathcal{I}_{k,D}} g\right) \leq (b_\phi - a_\phi) \int g. \tag{39}$$

This leads to

$$\sum_{k,k'\in\mathbb{Z}} \left(\int g_{k,D} g_{k',D} f \right)^2 \leq \|\phi\|_\infty^4 D^2 \left(2(b_\phi - a_\phi) + 1\right) \sup_{k'\in\mathbb{Z}} \left| \int_{\mathcal{I}_{k',D}} f \right| \leq C(\phi) D \|f\|_\infty .$$

Let us now evaluate $\left(\sum_{k\in\mathbb{Z}} a_{k,D}^2\right)^2$. On the one hand, using the fact that the functions $\{g_{k,D}, k \in \mathbb{Z}\}$ are orthonormal in $\mathbb{L}^2(\mathbb{R})$, one obtains that

$$\sum_{k\in\mathbb{Z}} a_{k,D}^2 \leq \int f^2 \leq \|f\|_\infty .$$

On the other hand, using (39) and the inequality $\int_{\mathcal{I}_{k,D}} f \leq \int f \leq 1$,

$$\sum_{k\in\mathbb{Z}} a_{k,D}^2 \leq D\|\phi\|_\infty^2 \sum_{k\in\mathbb{Z}} \int_{\mathcal{I}_{k,D}} f \leq C(\phi) D.$$

The above evaluations imply that $\mathbb{E}\left(H_m^2(X_1, X_2)\right) \leq C(\phi) D \|f\|_\infty$ and therefore,

$$A_1 \leq C(\phi) n \sqrt{D\|f\|_\infty} .$$

- Evaluation of A_2.

Let $(\alpha_i)_{1\leq i\leq n}$ and $(\beta_j)_{1\leq j\leq n}$ satisfy $\mathbb{E}(\sum_{i=1}^n \alpha_i^2(X_1)) \leq 1$ and $\mathbb{E}(\sum_{j=1}^n \beta_j^2(X_1)) \leq 1$. Setting $f_D = \sum_{k\in\mathbb{Z}} a_{k,D} g_{k,D}$, one has:

$$\begin{aligned}
&\sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{E}\Big(H_m(X_1, X_2) \alpha_i(X_1) \beta_j(X_2) \Big) \\
&\qquad = \sum_{i=1}^n \sum_{j=1}^{i-1} \sum_{k\in\mathbb{Z}} \int \left(g_{k,D} - a_{k,D}\right) \alpha_i f \int \left(g_{k,D} - a_{k,D}\right) \beta_j f \\
&\qquad = \sum_{i=1}^n \sum_{j=1}^{i-1} \left(\sum_{k\in\mathbb{Z}} \int g_{k,D} \alpha_i f \int g_{k,D} \beta_j f + \sum_{k=0}^{D} a_{k,D}^2 \int \alpha_i f \int \beta_j f \right. \\
&\qquad \qquad \left. - \int \alpha_i f \int f_D \beta_j f - \int \beta_j f \int f_D \alpha_i f \right).
\end{aligned}$$

Using repeatedly the Cauchy-Schwarz inequality, one obtains that

$$\begin{aligned}
\left| \sum_{k\in\mathbb{Z}} \int g_{k,D} \alpha_i f \int g_{k,D} \beta_j f \right| &\leq \left[\sum_{k\in\mathbb{Z}} \left(\int g_{k,D} \alpha_i f \right)^2 \right]^{1/2} \left[\sum_{k\in\mathbb{Z}} \left(\int g_{k,D} \beta_j f \right)^2 \right]^{1/2} \\
&\leq \left[\sum_{k\in\mathbb{Z}} \int_{\mathcal{I}_{k,D}} \alpha_i^2 f \int (g_{k,D})^2 f \right]^{1/2} \left[\sum_{k\in\mathbb{Z}} \int_{\mathcal{I}_{k,D}} \beta_j^2 f \int (g_{k,D})^2 f \right]^{1/2} .
\end{aligned}$$

For all $k \in \mathbb{Z}$, by the definition of $\mathcal{I}_{k,D}$,

$$\int (g_{k,D})^2 f \leq D\|\phi\|_\infty^2 \int_{\mathcal{I}_{k,D}} f \leq C(\phi) \|f\|_\infty .$$

Moreover, using (39),

$$\sum_{k\in\mathbb{Z}} \int_{\mathcal{I}_{k,D}} \alpha_i^2 f \leq (b_\phi - a_\phi) \int \alpha_i^2 f \quad \text{and} \quad \sum_{k\in\mathbb{Z}} \int_{\mathcal{I}_{k,D}} \beta_j^2 f \leq (b_\phi - a_\phi) \int \beta_j^2 f.$$

This implies that

$$\left|\sum_{k\in\mathbb{Z}}\int g_{k,D}\alpha_i f\int g_{k,D}\beta_j f\right|\le C(\phi)\|f\|_\infty\left(\int\alpha_i^2 f\right)^{1/2}\left(\int\beta_j^2 f\right)^{1/2},$$

so, by using the same computations as for the case $iii)$,

$$\sum_{i=1}^n\sum_{j=1}^{i-1}\left|\sum_{k\in\mathbb{Z}}\int g_{k,D}\alpha_i f\int g_{k,D}\beta_j f\right|\le C(\phi)\|f\|_\infty n.$$

Since $\sum_{k\in\mathbb{Z}}a_{k,D}^2\le\|f\|_\infty$, one obtains

$$\sum_{i=1}^n\sum_{j=1}^{i-1}\sum_{k\in\mathbb{Z}}a_{k,D}^2\left|\int\alpha_i f\int\beta_j f\right|\le\|f\|_\infty n.$$

Finally, from the fact that $\sup_{k\in\mathbb{Z}}|a_{k,D}|\le\|\phi\|_\infty(b_\phi-a_\phi)\|f\|_\infty/\sqrt{D}$ and

$$\Big\|\sum_{k\in\mathbb{Z}}g_{k,D}\Big\|_\infty\le\sqrt{D}\|\phi\|_\infty\left(2(b_\phi-a_\phi)+1\right),\tag{40}$$

one derives that $\|f_D\|_\infty\le C(\phi)\|f\|_\infty$ and

$$\sum_{i=1}^n\sum_{j=1}^{i-1}\left|\int\alpha_i f\int f_D\beta_j f\right|\le\sum_{i=1}^n\sum_{j=1}^n\int|f_D\beta_j|f\int|\alpha_i|f\le C(\phi)\|f\|_\infty n.$$

$\sum_{i=1}^n\sum_{j=1}^{i-1}\left|\int\beta_j f\int f_D\alpha_i f\right|$ is also bounded by $\|f\|_\infty n$ so,

$$A_2\le C(\phi)\|f\|_\infty n.$$

- Evaluation of A_3.

$$\begin{aligned}
&\mathbb{E}_{X_2}(H_m^2(x,X_2))\\
&=\sum_{k,k'\in\mathbb{Z}}\left(\int g_{k,D}g_{k',D}f\right)(g_{k,D}(x)-a_{k,D})(g_{k',D}(x)-a_{k',D})-\left(\sum_{k\in\mathbb{Z}}a_{k,D}(g_{k,D}(x)-a_{k,D})\right)^2\\
&\le\sum_{k,k'\in\mathbb{Z},|k-k'|\le b_\phi-a_\phi}\left(\int g_{k,D}g_{k',D}f\right)(g_{k,D}(x)-a_{k,D})(g_{k',D}(x)-a_{k',D}).
\end{aligned}$$

On the one hand, for all $k,k'\in\mathbb{Z}$,

$$\left|\int g_{k,D}g_{k',D}f\right|\le\|\phi\|_\infty^2 D\int_{\mathcal{I}_{k,D}\cap\mathcal{I}_{k',D}}f\le C(\phi)\|f\|_\infty.$$

On the other hand, using (40) and the inequality

$$\Big|\sum_{k\in\mathbb{Z}}a_{k,D}\Big|=\Big|\int f\sum_{k\in\mathbb{Z}}g_{k,D}\Big|\le\Big\|\sum_{k\in\mathbb{Z}}g_{k,D}\Big\|_\infty,$$

one gets

$$\Big\|\sum_{k\in\mathbb{Z}}(g_{k,D}(x)-a_{k,D})\Big\|_\infty\le C(\phi)\sqrt{D}.\tag{41}$$

Hence,

$$\Big\|\sum_{k,k'\in\mathbb{Z}}(g_{k,D}(x)-a_{k,D})(g_{k',D}(x)-a_{k',D})\Big\|_\infty\le C(\phi)D,$$

which leads to the upper bound:

$$A_3 \leq C(\phi)\sqrt{Dn\|f\|_\infty}.$$

- Evaluation of A_4.

Using (41) and the fact that $\|g_{k,D}\|_\infty \leq \sqrt{D}\|\phi\|_\infty$, it is easy to see that

$$A_4 \leq C(\phi)D.$$

This completes the proof of Lemma 1.

REFERENCES

Baraud Y., Huet S., Laurent B. (2001). Test for convex hypotheses. *Technical report- 2001-66 Université Paris-Sud.*

Baraud Y., Huet S., Laurent B. (2003a). Adaptive tests of linear hypotheses by model selection. *Ann. Statist.*, Vol. 31, No. 1, pp. 225–251.

Baraud Y., Huet S., Laurent B. (2003b). Adaptive tests of qualitative hypotheses. *ESAIM Probab. Stat.*, Vol. 7, pp. 147–159.

Bickel P., Ritov Y. (1992). Testing for goodness-of-fit: a new approach. *Nonparametric statistics and related topics- North-Holland, Amsterdam*, pp. 51–57.

Birgé L., Massart P. (1998). Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli*, Vol. 4, No. 3, pp. 329–375.

DeVore R.A. and Lorentz G.G. (1993). *Constructive Approximation.* Springer, Berlin.

Devore R.A., Jawerth, B., Popov, V. (1992). Compression of wavelet decompositions. *Amer. J. Math.*, Vol. 114, pp. 737–785.

Fan, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *J. Amer. Statist. Assoc.*, Vol. 91, No. 434, pp. 674–688.

Fromont, M. 2003. *Quelques problèmes de sélection de modèles : construction de tests adaptatifs, ajustement de pénalités par des méthodes de bootstrap.* Ph.D. thesis, Université Paris-Sud.

Giné E., Latala R., Zinn J. (2000). Exponential and moment inequalities for U-statistics. *High Dimensional Probability 2, Progress in Probability*, Vol. 47, pp. 13–38.

Houdré C., Reynaud-Bouret P. (2002). Exponential inequalities for U-statistics of order 2 with constants. *To appear in Stochastic Inequalities and Applications. Birkhauser. Euroconference Barcelona - June 2002.*

Inglot T., Ledwina T. (1996). Asymptotic optimality of data-driven Neyman's tests for uniformity. *Ann. Statist.*, Vol. 24, No. 5, pp. 1982–2019.

Inglot T., Kallenberg W., Ledwina T. (1997). Data driven smooth tests for composite hypotheses. *Ann. Statist.*, Vol. 25, No. 3, pp. 1222–1250.

Ingster, Yu.I. (1993). Asymptotically minimax testing for nonparametric alternatives I-II-III. *Math. Methods Statist.*, Vol. 2, pp. 85–114, 171–189, 249–268.

Ingster, Yu.I. (2000). Adaptive chi-square tests. *J. Math. Sci*, Vol. 99, No. 2, pp. 1110–1119.

Kallenberg, W. (2002). The penalty in data driven Neyman's tests. *Math. Methods Statist.*, Vol. 11, pp. 323–340.

Kallenberg W., Ledwina T. (1995). Consistency and Monte Carlo simulation of a data driven version of smooth goodness-of-fit tests. *Ann. Statist.*, Vol. 23, No. 5, pp. 1594–1608.

Laurent, B. (2003). Adaptive estimation of a quadratic functional of a density by model selection. *Technical Report- Université Paris-Sud.*

Ledwina, T. (1994). Data-driven version of Neyman's smooth test of fit. *J. Amer. Statist. Assoc.*, Vol. 89, No. 427, pp. 1000–1005.

Neyman, J. (1937). Smooth test for goodness of fit. *Skand. Aktuarietidskr.*, pp. 150–199.

Pouet, C. (2002). Test asymptotiquement minimax pour une hypothèse nulle composite dans le modèle de densité. *C.R. Math. Acad. Sci. Paris*, Vol. 334, No. 10, pp. 913–916.

Magalie FROMONT
Laboratoire de Mathématiques
U.M.R. C 8628
UUNIVERSITÉ PARIS-SUD
Bât. 425
91405 ORSAY CEDEX
FRANCE
e-mail : magalie.fromont@math.u-psud.fr

Béatrice LAURENT
Département de Génie Mathématique
I.N.S.A.
135, Avenue de Rangueil
31077 TOULOUSE CEDEX 4
FRANCE
e-mail : beatrice.laurent@gmm.insa-tlse.fr