ESAIM: PS
February 2005, Vol. 9, p. 1–18
DOI: 10.1051/ps:2005001

ESAIM: Probability and Statistics

ADAPTIVE ESTIMATION OF A QUADRATIC FUNCTIONAL OF A DENSITY BY MODEL SELECTION

Béatrice Laurent[1]

Abstract. We consider the problem of estimating the integral of the square of a density f from the observation of a n sample. Our method to estimate $\int_{\mathbb{R}} f^2(x)\mathrm{d}x$ is based on model selection *via* some penalized criterion. We prove that our estimator achieves the adaptive rates established by Efroimovich and Low on classes of smooth functions. A key point of the proof is an exponential inequality for U-statistics of order 2 due to Houdré and Reynaud.

Mathematics Subject Classification. 62G05, 62G20, 62J02.

Received October 23, 2003. Revised July 16, 2004.

1. Introduction

Let $X_1, \ldots, X_n$ be i.i.d. real random variables with common density $f \in \mathbb{L}^2(\mathbb{R})$. The aim of this paper is to propose an adaptive estimator $\int_{\mathbb{R}} f^2(x)\mathrm{d}x$.

Bickel and Ritov [1] and Laurent [16,17] have built estimators of $\int f^2$ in a density model but these estimators depend on some prior information on f. Bickel and Ritov [1] assumed that f belongs to some compact set included in the class of Hölderian functions of order α. They built an estimator $\widehat{\theta}_\alpha$ of $\int f^2$ that is efficient if $\alpha > 1/4$ and achieves the rate $n^{-4\alpha/(1+4\alpha)}$ if $\alpha \leq 1/4$. Moreover, they proved that this rate is optimal. Similar results are also obtained in Laurent [16] with a simpler method of estimation based on projection estimators, which allows to built efficient estimators of more general functionals of the form $\int \Phi(f)$ if $\alpha > 1/4$. Birgé and Massart [2] have established minimax lower bounds for the estimation of integral functionals of a density.

Several papers are devoted to the estimation of quadratic functionals in the Gaussian sequence model, that is when one observes

$$Y_\lambda = \beta_\lambda + \frac{1}{\sqrt{n}}\epsilon_\lambda, \quad \lambda \in \mathbb{N}^*,$$

where $(\epsilon_\lambda, \lambda \in \mathbb{N}^*)$ is a sequence of i.i.d. standard Gaussian variables. This model can be derived from the Gaussian white noise model:

$$Y(t) = \int_0^t f(u)\mathrm{d}u + \frac{1}{\sqrt{n}}W(t), \quad t \in [0,1]$$

Keywords and phrases. Adaptive estimation, quadratic functionals, model selection, Besov bodies, efficient estimation.

[1] INSA-LSP. Departement GMM, 135 avenue de Rangueil, 31077 Toulouse Cedex 4, France;
beatrice.laurent@insa-toulouse.fr

after some projection onto an orthonormal basis of $\mathbb{L}^2([0,1])$. The sequence $(\beta_\lambda, \lambda \in \mathbb{N}^*)$ corresponds to the sequence of coefficients of the signal f onto this orthonormal basis. The quantity to be estimated here is $\theta = \sum_{\lambda\in\mathbb{N}^*} \beta_\lambda^2 = \int_0^1 f^2$.

Ibragimov, Nemirovskii and Hasminskii [14] considered the problem of estimating a general functional of the signal f in the white noise model and established the conditions (in terms of regularity of the functional and of the signal) under which the functional can be estimated efficiently.

Donoho and Nussbaum [6] proposed an estimator of $\sum_{\lambda\in\mathbb{N}^*} \beta_\lambda^2$ in the Gaussian sequence model, with the prior information that the sequence $(\beta_\lambda, \lambda \in \mathbb{N}^*)$ belongs to some ellipsoïd.

Efroimovich and Low [8] were the first to propose adaptive estimators of quadratic functionals in the framework of the Gaussian sequence model. The estimator $\widehat{\theta}_n$ proposed by Efroimovich and Low is inspired by Lepskii's method of adaptation and it has the following adaptive properties: for any positive R and α, provided that the sequence $(\beta_\lambda)_{\lambda\geq 1}$ satisfies to the condition $\beta_\lambda^2\lambda^{2\alpha+1} \leq R^2$ for all λ (which means that $(\beta_\lambda)_{\lambda\geq 1}$ belongs to some hyperrectangle) one has

- $\mathbb{E}\Big[\Big(\widehat{\theta}_n - \theta\Big)^2\Big] \leq C\,(R,\alpha)\,\big(n^{-2}\log(n)\big)^{4\alpha/(1+4\alpha)}$, if $\alpha \leq 1/4$.
- $\widehat{\theta}_n$ is asymptotically efficient if $\alpha > 1/4$.

Efroimovich and Low also proved that this rate is optimal, which means that the logarithmic factor that appears in the rates of convergence cannot be avoided if we do not know *a priori* to what hyperrectangle the sequence β belongs.

Johnstone [15] proposed estimators of $\theta = \sum_{\lambda\in\mathbb{N}^*} \beta_\lambda^2$ in the Gaussian sequence model which are based on wavelet thresholding methods and proved that these estimators are adaptive on Hölder classes. Gayraud and Tribouley [10] also proposed estimators based on wavelet thresholding methods and proved the adaptivity on Besov balls. They also give asymptotic confidence intervals for θ.

Laurent and Massart [18] built adaptive estimators of quadratic functionals in a Gaussian framework covering both the Gaussian sequence model and the finite dimensional Gaussian regression. These estimators are based on model selection *via* some penalized criterion. In the framework of the Gaussian sequence model, the penalized estimator is defined in the following way: one considers a collection $\mathcal{M}$ of subset of $\mathbb{N}^*$ and a penalty function $\mathrm{pen} : \mathcal{M} \to \mathbb{R}^+$. The penalized estimator of $\theta = \sum_{\lambda\in\mathbb{N}^*} \beta_\lambda^2$ is defined as

$$\widehat{\theta} = \sup_{m\in\mathcal{M}} \left(\sum_{\lambda\in m} Y_\lambda^2 - \mathrm{pen}(m) \right).$$

For suitable choices of the set $\mathcal{M}$ and of the penalty function, this estimator is adaptive over more general classes of sequences of coefficients $(\beta_\lambda)_{\lambda\geq 1}$ than the estimators proposed by the previous authors.

In this paper, we propose an adaptive estimator of $\int_{\mathbb{R}} f^2(x)\mathrm{d}x$ in a density model which is also based on model selection *via* some penalized criterion. We show that this estimator achieves the minimax adaptive rate established by Efroimovich and Low [7] over Besov bodies $B_{\alpha,2,\infty}(R)$.

This paper is also motivated by applications to adaptive goodness-of-fit tests in a density model that are proposed in Fromont and Laurent [9].

A crucial point in the proof of our results is an exponential inequality for U-statistics of order 2 due to Houdré and Reynaud [13].

The paper is organized as follows: in Section 2, we recall some results concerning the estimation of $\int_{\mathbb{R}} f^2(x)\mathrm{d}x$ and we introduce the estimation *via* model selection. In Section 3, we give our main results. Section 4 contains the main tool for the proof of Theorem 1. Section 5 is devoted to the proofs.

2. Estimation *via* model selection

Let $X_1, \ldots, X_n$ be i.i.d. random variables with common density f belonging to $\mathbb{L}^2(\mathbb{R})$. Our aim is to estimate $\int_{\mathbb{R}} f^2(x)\mathrm{d}x$. To do this, we consider the orthogonal projection of f onto the Haar basis. We first introduce some

Article published by EDP Sciences and available at http://www.edpsciences.org/ps or http://dx.doi.org/10.1051/ps:2005001

notations. Let

$$\phi(x) = \mathbb{1}_{[0,1]}(x), \quad \psi(x) = \mathbb{1}_{[0,1/2[}(x) - \mathbb{1}_{[1/2,1[}(x),$$

and for any $k \in \mathbb{Z}$ and $j \in \mathbb{N}$, let

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k), \quad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k).$$

The functions $(\phi_{0,k}, \psi_{j,k}, j \in \mathbb{N}, k \in \mathbb{Z})$ form the Haar basis of $\mathbb{L}^2(\mathbb{R})$. The decomposition of f onto this basis may be written as:

$$\sum_{k\in\mathbb{Z}} \alpha_{0,k}(f)\phi_{0,k} + \sum_{j\geq 0}\sum_{k\in\mathbb{Z}} \beta_{j,k}(f)\psi_{j,k}$$

where $\alpha_{j,k}(f) = \int f\phi_{j,k}$ and $\beta_{j,k}(f) = \int f\psi_{j,k}$. For any $J \in \mathbb{N}$, this decomposition is also equal to

$$\sum_{k\in\mathbb{Z}} \alpha_{J,k}(f)\phi_{J,k} + \sum_{j\geq J}\sum_{k\in\mathbb{Z}} \beta_{j,k}(f)\psi_{j,k}.$$

We set $\beta(f) = (\beta_{j,k}(f))_{j\geq 0, k\in\mathbb{Z}}$.

We define the function f_J by

$$f_J = \sum_{k\in\mathbb{Z}} \alpha_{J,k}(f)\phi_{J,k}.$$

For any $J \in \mathbb{N}$, we consider the unbiased estimator of $\int f_J^2 = \sum_{k\in\mathbb{Z}} \alpha_{J,k}^2$ defined by

$$\widehat{\theta}_J = \frac{1}{n(n-1)} \sum_{k\in\mathbb{Z}} \sum_{l\neq l'=1}^{n} \phi_{J,k}(X_l)\phi_{J,k}(X_{l'}). \tag{1}$$

In order to evaluate the quadratic risk of this estimator, we use the decomposition

$$\mathbb{E}\left[(\widehat{\theta_J} - \theta)^2\right] = \mathrm{Bias}^2(\widehat{\theta}_J) + \mathrm{Var}(\widehat{\theta}_J).$$

Since the expectation of $\widehat{\theta}_J$ equals $\int f_J^2$, $|\mathrm{Bias}(\widehat{\theta}_J)| = \int (f - f_J)^2$. Assuming that $\|f\|_\infty$ is finite, one can easily show that

$$\mathrm{Var}(\widehat{\theta}_J) \leq C(\|f\|_\infty)\left(\frac{2^J}{n^2} + \frac{1}{n}\right)$$

where $C(\|f\|_\infty)$ is a constant depending on $\|f\|_\infty$.

Let us assume that we have some prior information on f, for example that the sequence $\beta(f)$ belongs to the Besov body $\mathcal{B}_{\alpha,2,\infty}(R)$ defined by

$$\mathcal{B}_{\alpha,2,\infty}(R) = \left\{\beta = (\beta_{j,k})_{j\geq 0,k\in\mathbb{Z}}, \forall j \geq 0 \sum_{k\in\mathbb{Z}} \beta_{j,k}^2 \leq R^2 2^{-2j\alpha}\right\}.$$

This implies that for all $J \in \mathbb{N}$,

$$\int_{\mathbb{R}} (f - f_J)^2 = \sum_{j\geq J}\sum_{k\in\mathbb{Z}} \beta_{j,k}^2 \leq C(\alpha)R^2 2^{-2J\alpha},$$

where $C(\alpha)$ is a constant depending on α. Choosing J in such a way that

$$\frac{2^J}{n^2} \approx R^4 2^{-4J\alpha},$$

we obtain that
$$\mathbb{E}\left[(\widehat{\theta}_J-\theta)^2\right] \le C(\|f\|_\infty)\left(R^{4/1+4\alpha}n^{-8\alpha/1+4\alpha}+\frac{1}{n}\right)\cdot$$
The rate obtained here corresponds to the minimax rate for estimating θ over Hölderian balls with index α and radius R as it is proved by Bickel and Ritov [1] and Birgé and Massart [2].

Since our choice of J depends on the unknown parameters α and R, this is not satisfactory. We shall now present some heuristics of the adaptive procedure *via* model selection.

Adaptive estimation of $\int f^2$.

We have seen that the "ideal" choice of J minimizes the quantity
$$\int (f-f_J)^2+\frac{2^{J/2}}{n}=\int f^2-\int f_J^2+\frac{2^{J/2}}{n}\cdot$$
Since $\int f^2$ does not depend on J, this is equivalent to maximize the quantity $\int f_J^2 - 2^{J/2}/n$. $\int f_J^2$ is estimated without bias by $\widehat{\theta}_J$, hence we set
$$\widehat{J}=\mathrm{argmax}_{J\in\mathcal{J}}\left[\widehat{\theta}_J-\mathrm{pen}(J)\right]$$
and
$$\widehat{\theta}=\widehat{\theta}_{\widehat{J}}-\mathrm{pen}(\widehat{J})=\sup_{J\in\mathcal{J}}\left[\widehat{\theta}_J-\mathrm{pen}(J)\right], \tag{2}$$
where $\mathcal{J}$ is some subset of $\mathbb{N}$ and pen(J) is non-negative quantity that we shall specify in the following.

3. Main results

In Theorem 1, we consider the classes of functions f which are uniformly bounded and for which the sequence of coefficients onto the Haar basis belongs to some Besov body.

We show that for a suitable choice of the penalty term pen(J) appearing in the definition of the estimator $\widehat{\theta}$ given by (2), the estimator $\widehat{\theta}$ is adaptive over these classes.

Theorem 1. *Let $X_1,\ldots,X_n$ be i.i.d. real random variables with common density f belonging to $\mathbb{L}_\infty(\mathbb{R})$. Let $\theta=\int_{\mathbb{R}} f^2(x)\mathrm{d}x$.*

Let $\mathcal{J}=\left\{J\in\mathbb{N}, 2^J\le n^2/\log^3(n)\right\}$. For all $J\in\mathbb{N}$, let $\widehat{\theta}_J$ be defined by (1).

There exists some absolute constant $\kappa>0$ such that if we set for all $J\in\mathcal{J}$
$$\mathrm{pen}(J)=\frac{\kappa}{n}\left[\sqrt{(\widehat{\theta}_J+1)2^J\log(2^J+1)}\right],$$
then the estimator $\widehat{\theta}$ defined by (2) has the following properties:

For any $\alpha>0$, $R>0$ and $M>0$, there exists some integer $n_0(\alpha,R,M)$ depending on α, R and M such that the following inequality holds for all $n\ge n_0(\alpha,R,M)$:
$$\sup_{f,\beta(f)\in\mathcal{B}_{\alpha,2,\infty}(R),\|f\|_\infty\le M}\mathbb{E}\left[\left(\widehat{\theta}-\theta-\frac{2}{n}\sum_{i=1}^n(f(X_i)-\theta)\right)^2\right]\le C(\alpha)(R(M+1)^\alpha)^{\frac{4}{1+4\alpha}}\left(\frac{\sqrt{\log(nR^2)}}{n}\right)^{\frac{8\alpha}{1+4\alpha}}.$$

This leads to

- *for all $\alpha > \frac{1}{4}$, $R > 0$, and $M > 0$, there exists some integer $n_1(\alpha, R, M)$ such that the following inequality holds for all $n \geq n_1(\alpha, R, M)$:*

$$\sup_{f,\beta(f)\in\mathcal{B}_{\alpha,2,\infty}(R),\|f\|_\infty\leq M} \mathbb{E}\left[\left(\widehat{\theta}-\theta\right)^2\right] \leq C(\alpha)\frac{M^2}{n}; \tag{3}$$

- *for all $\alpha \leq \frac{1}{4}$, $R > 0$, and $M > 0$, there exists some integer $n_2(\alpha, R, M)$ such that the following inequality holds for all $n \geq n_2(\alpha, R, M)$:*

$$\sup_{f,\beta(f)\in\mathcal{B}_{\alpha,2,\infty}(R),\|f\|_\infty\leq M} \mathbb{E}\left[\left(\widehat{\theta}-\theta\right)^2\right] \leq C(\alpha)(R(M+1)^\alpha)^{\frac{4}{1+4\alpha}}\left(\frac{\sqrt{\log(nR^2)}}{n}\right)^{\frac{8\alpha}{1+4\alpha}}. \tag{4}$$

If $f \in \mathbb{L}_\infty(\mathbb{R})$ and $\beta(f) \in \mathcal{B}_{\alpha,2,\infty}(R)$ for some $\alpha > 1/4$ and $R > 0$, then

$$\sqrt{n}(\widehat{\theta}-\theta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\ \mathrm{Var}(2f(X_1))\right) \ \text{ as } n\to\infty \tag{5}$$

$$n\mathbb{E}\left[(\widehat{\theta}-\theta)^2\right] \overset{n\to\infty}{\to} \mathrm{Var}(2f(X_1)). \tag{6}$$

Comments:

1) We derive from (5) that if $\beta(f) \in \mathcal{B}_{\alpha,2,\infty}(R)$ for some $\alpha > 1/4$, then $\widehat{\theta}$ is an efficient estimator of θ (see Laurent [16]).
2) If $\alpha \leq 1/4$, we obtain the rate $(\sqrt{\log(n)}/n)^{4\alpha/(1+4\alpha)}$.

Let $\mathcal{H}_\alpha(R)$ denote the Hölderian ball defined by

$$\mathcal{H}_\alpha(R) = \{f : [0,1] \to \mathbb{R}, \forall x, y \in [0,1], |f(x) - f(y)| \leq R|x-y|^\alpha\}. \tag{7}$$

The minimax rate for estimating $\int_{\mathbb{R}} f^2$ over $\mathcal{H}_\alpha(R)$ if $\alpha \leq 1/4$ is $n^{-4\alpha/(1+4\alpha)}$ (see Birgé and Massart [2]).

Efroimovich and Low [7] proved that the logarithmic loss with respect to the minimax rate that appears in the adaptive lower bounds for estimating $\int f^2$ is unavoidable. This is the purpose of the following proposition:

Proposition 1 (Efroimovich and Low [7])**.** *Suppose that $\widehat{\theta}_n$ is an estimator of θ based on the n sample $X_1, \ldots, X_n$. If, for some $\alpha > 1/4$,*

$$\lim_{n\to\infty} \sup_{f\in\mathcal{H}_\alpha(R)} n\mathbb{E}\left[\left(\widehat{\theta}_n-\theta\right)^2\right] < \infty,$$

then for every $\alpha < 1/4$,

$$\lim_{n\to\infty} \sup_{f\in\mathcal{H}_\alpha(R)} \left(\frac{n^2}{\log(n)}\right)^{\frac{4\alpha}{1+4\alpha}} \mathbb{E}\left[\left(\widehat{\theta}_n-\theta\right)^2\right] > 0.$$

Since $\mathcal{H}_\alpha(R) \subset \{f, \beta(f) \in \mathcal{B}_{\alpha,2,\infty}(R)\}$ this proves that the rate of convergence obtained in Theorem 1 corresponds to the minimax adaptive rate for estimating θ over the set $\{f, \beta(f) \in \mathcal{B}_{\alpha,2,\infty}(R)\}$.

3) If the sequence $\beta(f)$ belongs to $\mathcal{B}_{\alpha,2,\infty}(R)$ for some $R > 0$ et $\alpha > 1/2$, then f is uniformly bounded (see inequality (8.15) of Proposition (8.3) in Haŕdle *et al.* [12]). Therefore, the assumption of boundedness on f is only a restriction if $\alpha \leq 1/2$. In all cases, assuming that $\|f\|_\infty \leq M$, we provide an upper bound for the quadratic risk, where the dependency with respect to M is given.

4. An oracle inequality

The result stated in this section is the main tool for the proof of Theorem 1. It provides a non asymptotic bound for the risk of the penalized estimator $\widehat{\theta}$ defined by (2).

Proposition 2. *Let $X_1, \ldots, X_n$ be i.i.d. real random variables with common density f belonging to $\mathbb{L}_\infty(\mathbb{R})$. Let $\theta = \int_{\mathbb{R}} f^2(x)\mathrm{d}x$. Let $\mathcal{J}_n$ be a subset of $\left\{J \in \mathbb{N}, 2^J \leq \frac{n^2}{\log^3(n)}\right\}$. There exists some absolute constant κ_0 such that if we set for all $J \in \mathcal{J}_n$*

$$\mathrm{pen}(J) = \frac{\kappa}{n}\left[\sqrt{(\widehat{\theta}_J + 1)2^J \log(2^J + 1)}\right]$$

with $\kappa \geq \kappa_0$, then the estimator $\widehat{\theta}$ defined by (2) satisfies the following inequality for all $n \geq 2$ provided that $\|f\|_\infty \leq M$:

$$\mathbb{E}\left[\left(\widehat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^n (f(X_i) - \theta)\right)^2\right] \leq C \inf_{J \in \mathcal{J}_n}\left[\|f_J - f\|_2^4 + \frac{2^J(M+1)\log(2^J+1)}{n^2}\right] + \frac{C(M)}{n^2}.$$

C is an absolute constant and $C(M)$ is some constant depending on M only.

Comments.

1) One can derive from Proposition 2 an upper bound for the quadratic risk of $\widehat{\theta}$, indeed

$$\mathbb{E}\left[\left(\widehat{\theta} - \theta\right)^2\right] \leq 2\mathbb{E}\left[\left(\widehat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^n (f(X_i) - \theta)\right)^2\right] + 2\mathbb{E}\left[\left(\frac{2}{n}\sum_{i=1}^n (f(X_i) - \theta)\right)^2\right]$$

and

$$\mathbb{E}\left[\left(\frac{2}{n}\sum_{i=1}^n (f(X_i) - \theta)\right)^2\right] \leq \frac{4}{n}\int f^3.$$

2) One can also deduce from Proposition 2 that if

$$\inf_{J \in \mathcal{J}_n}\left[\|f_J - f\|_2^4 + \frac{2^J(M+1)\log(2^J+1)}{n^2}\right] = o(1/n),$$

then $\sqrt{n}\left(\widehat{\theta} - \theta - 2\sum_{i=1}^n (f(X_i) - \theta)/n\right)$ tends to zero in probability, which implies that

$$\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\ \mathrm{Var}(2f(X_1))\right).$$

In this situation $\widehat{\theta}$ is an efficient estimator of θ (see Laurent [16]).

3) In order to prove Proposition 2, we use an exponential inequality with explicit constants for U-statistics of order 2 due to Houdré and Reynaud [13]. It is worth mentioning the paper by Giné, Latala and Zinn [11] where an exponential inequality for general U-statistics is given, and the paper by Bretagnolle [5] where an exponential inequality for U-statistics of order 2 is also established.

4) We could derive from the explicit constants given in Houdré and Reynaud's inequality an upper bound for κ_0, but this upper bound would be very large. A simulation study would be necessary to know how to calibrate κ_0 in practice. Such a simulation study was carried out by Birgé and Rozenholc [4] in the case of density estimation with histograms.

5. Proofs

In the sequel, we denote by C some absolute constant whose value may vary from one line to another. We always mention the dependency of a constant with respect to some parameters: for example $C(\alpha, R)$ stands for a constant depending on α and R.

Before proving Proposition 2, we prove the following Proposition, where an upper bound for $\|f\|_\infty$ is assumed to be known.

Proposition 3. *Let $X_1, \dots, X_n$ be i.i.d. real random variables with common density f belonging to $\mathbb{L}_\infty(\mathbb{R})$. We assume that $\|f\|_\infty \leq M$ where M is known. Let $\theta = \int_{\mathbb{R}} f^2(x)\mathrm{d}x$. Let $\mathcal{J}$ be a subset of $\mathbb{N}$. For all $J \in \mathcal{J}$, let $\widehat{\theta}_J$ be defined by (1). There exists some constant $\kappa_0 > 0$ such that if we set for all $J \in \mathcal{J}$*

$$n\mathrm{pen}(J) = \kappa \left[\sqrt{M 2^J \log(2^J + 1)} + M \log(2^J + 1) + \frac{2^J \log^2(2^J + 1)}{n} \right]$$

with $\kappa \geq \kappa_0$ and

$$\widehat{\theta} = \sup_{J \in \mathcal{J}} \left(\widehat{\theta}_J - \mathrm{pen}(J) \right),$$

then the following inequality holds for all $n \geq 2$:

$$\mathbb{E}\left[\left(\widehat{\theta} - \theta - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) \right)^2 \right] \leq C \inf_{J \in \mathcal{J}} \left[\|f_J - f\|_2^4 + \mathrm{pen}^2(J) \right]$$

where C is some absolute constant.

5.1. Proof of Proposition 3

We use the canonical decomposition of the U-statistics $\widehat{\theta}_J$. We denote by U_n the process defined by $U_n(H) = (1/n(n-1)) \sum_{l \neq l'=1}^{n} H(X_l, X_{l'})$ and we denote by P_n the empirical measure $P_n(h) = (1/n) \sum_{i=1}^{n} h(X_i) - \int hf$. We set $\alpha_{J,k} = \int f \phi_{J,k}$,

$$H_J(x, y) = \sum_{k \in \mathbb{Z}} (\phi_{J,k}(x) - \alpha_{J,k}) (\phi_{J,k}(y) - \alpha_{J,k})$$

and

$$h_J = 2(f_J - f).$$

The following decomposition holds:

$$\widehat{\theta}_J - \theta - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) = U_n(H_J) + P_n(h_J) - \|f - f_J\|_2^2. \tag{8}$$

Let us denote by V_J the variable

$$V_J = U_n(H_J) + P_n(h_J) - \|f - f_J\|_2^2 - \mathrm{pen}(J).$$

By definition of $\widehat{\theta}$, and by (8),

$$\widehat{\theta} - \theta - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) = \sup_{J \in \mathcal{J}} (V_J).$$

Moreover, since

$$|\sup_{J \in \mathcal{J}} V_J| = \left[\sup_{J \in \mathcal{J}} (V_J)_+ \right] \vee \left[\inf_{J \in \mathcal{J}} (V_J)_- \right],$$

we obtain that

$$\mathbb{E}\left[\left(\sup_{J\in\mathcal{J}} V_J\right)^2\right] \leq \sum_{J\in\mathcal{J}} \mathbb{E}\left[(V_J)_+^2\right] + \inf_{J\in\mathcal{J}} \mathbb{E}\left[(V_J)_-^2\right].$$

We first control $\mathbb{E}\left[(V_J)_+^2\right]$. To do this, we use an exponential inequality for U-statistics of order 2 due to Houdré and Reynaud [13] in order to control the term $U_n(H_J)$. In order to control the term $P_n(h_J) - \|f - f_J\|_2^2$, we use Bernstein's inequality.

Control of $U_n(H_J)$.

We shall use the following Lemma that is a consequence of Houdré and Reynaud's exponential inequality for U-statistics of order 2:

Lemma 1. *Let $X_1, \ldots, X_n$ be i.i.d. with common density $f \in \mathbb{L}_\infty(\mathbb{R})$. Let for all $J \in \mathbb{N}$ and $k \in \mathbb{Z}$, let $\alpha_{J,k} = \int f\phi_{J,k}$, and $\theta_J = \sum_{k\in\mathbb{Z}} \alpha_{J,k}^2$. Let*

$$H_J(x,y) = \sum_{k\in\mathbb{Z}} (\phi_{J,k}(x) - \alpha_{J,k})(\phi_{J,k}(y) - \alpha_{J,k})$$

and

$$U_n(H_J) = \frac{1}{n(n-1)} \sum_{l\neq l'=1}^{n} H_J(X_l, X_{l'}).$$

There exists some absolute constant $C_0 > 0$ such that for all $J \in \mathbb{N}$, for all $t > 0$,

$$\mathbb{P}\left(|U_n(H_J)| > \frac{C_0}{n-1}\left(\sqrt{2^J\theta_J t} + \|f\|_\infty t + \frac{2^J t^2}{n}\right)\right) \leq 5.6\exp(-t).$$

The proof of the lemma is postponed to the Appendix.

We set for all $t > 0$

$$u_J(t) = \frac{C_0}{n-1}\left(\sqrt{2^J\theta_J t} + \|f\|_\infty t + \frac{2^J t^2}{n}\right). \tag{9}$$

We derive from Lemma 1 that for all $t \geq 0$,

$$\mathbb{P}\left(|U_n(H_J)| > u_J(t)\right) \leq 5.6\exp(-t). \tag{10}$$

Noticing that for all $t_1 > 0$ and all $t_2 > 0$

$$u_J\left(\frac{t_1 + t_2}{\sqrt{2}}\right) \leq u_J(t_1) + u_J(t_2),$$

we derive from (10) that for all $t \geq 0$ and $y_J \geq 0$,

$$\mathbb{P}\left(|U_n(H_J)| > u_J(\sqrt{2}y_J) + u_J(\sqrt{2}t)\right) \leq 5.6\exp - (t + y_J). \tag{11}$$

Control of $P_n(h_J) - \|f - f_J\|_2^2$.

We use the following lemma due to Birgé and Massart [3] which provides a special version of Bernstein's inequality.

Lemma 2. *Let $U_1, \ldots, U_n$ be independent random variables such that for all $i \in \{1, \ldots, n\}$, $|U_i| \leq b$ and $\mathbb{E}(U_i^2) \leq \delta^2$. Then for all $t > 0$*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(U_i - \mathbb{E}(U_i)) > \frac{\delta\sqrt{2t}}{\sqrt{n}} + \frac{bt}{3n}\right) \leq \exp(-t).$$

We apply Lemma 2 with $U_i = 2(f_J - f)(X_i)$ for all $i = 1, \dots, n$. Note that

$$|U_i| \leq 2\|f_J - f\|_\infty \leq 4\|f\|_\infty$$

since $\|f_J\|_\infty \leq \|f\|_\infty$. Moreover,

$$\mathbb{E}(U_i^2) = 4\int (f_J - f)^2 f \leq 4\|f\|_\infty \|f_J - f\|_2^2.$$

We deduce from Lemma 2 that for all $t > 0$,

$$\mathbb{P}\left(P_n(h_J) > 2\sqrt{2\|f\|_\infty}\|f_J - f\|_2 \frac{\sqrt{t}}{\sqrt{n}} + \frac{4\|f\|_\infty t}{3n}\right) \leq \exp(-t).$$

Using the inequality $2ab \leq a^2 + b^2$, one obtains for all $t > 0$,

$$\mathbb{P}\left(P_n(h_J) - \|f - f_J\|_2^2 > \frac{10\|f\|_\infty t}{3n}\right) \leq \exp(-t), \tag{12}$$

which implies that for all $t \geq 0$ and $y_J \geq 0$

$$\mathbb{P}\left(P_n(h_J) - \|f - f_J\|_2^2 > \frac{10\|f\|_\infty y_J}{3n} + \frac{10\|f\|_\infty t}{3n}\right) \leq \exp -(t + y_J). \tag{13}$$

We now turn to the control of $\mathbb{E}\left[(V_J)_+^2\right]$.

Noticing that

$$\theta_J = \int f_J^2 \leq \int f^2 \leq \|f\|_\infty \leq M$$

since $\int f = 1$, we obtain that for all $J \in \mathcal{J}$,

$$\mathrm{pen}(J) \geq \frac{\kappa}{n}\left(\sqrt{\theta_J 2^J \log(2^J + 1)} + M\log(2^J + 1) + \frac{2^J \log^2(2^J + 1)}{n}\right). \tag{14}$$

Let for all $J \in \mathbb{N}$, $y_J = 3\log(2^J + 1)$. Using the inequality $1/(n-1) \leq 2/n$ and setting

$$\kappa_0 = \max\left(36C_0, 6\sqrt{2}C_0 + 10\right),$$

(14) implies that

$$\mathrm{pen}(J) \geq u_J(\sqrt{2}y_J) + \frac{10\|f\|_\infty y_J}{3n}. \tag{15}$$

It follows from the definition of V_J and from (11), (13) and (15) that for all $J \in \mathcal{J}$,

$$\mathbb{P}\left(V_J > u_J(t\sqrt{2}) + \frac{10Mt}{3n}\right) \leq 6.6\exp -(t + y_J). \tag{16}$$

We now use the identity

$$E\left[(V_J)_+^2\right] = 2\int_0^{+\infty} t\mathbb{P}\left(V_J > t\right) \mathrm{d}t.$$

This identity, together with (16) leads to

$$\mathbb{E}\left[(V_J)_+^2\right] \leq C\left\{\frac{2^J M}{n^2} + \frac{M^2}{n^2} + \frac{2^{2J}}{n^4}\right\}\exp(-y_J).$$

Article published by EDP Sciences and available at http://www.edpsciences.org/ps or http://dx.doi.org/10.1051/ps:2005001

Recalling that $\mathcal{J}$ is a subset of $\mathbb{N}$, $\sum_{J\in\mathcal{J}} 2^{2J}\exp(-y_J) \leq \sum_{J\geq 0} 2^{-J}$ which implies that

$$\sum_{J\in\mathcal{J}} E\left((V_J)_+^2\right) \leq C\left(\frac{M^2+M}{n^2}+\frac{1}{n^4}\right).$$

Let us now give an upper bound for $E\left[(V_J)_-^2\right]$. It follows from the definition of V_J that

$$\mathbb{E}\left[(V_J)_-^2\right] \leq 4\left\{E\left[U_n^2(H_J)\right]+E\left[P_n^2(h_J)\right]+\mathrm{pen}^2(J)+\|f_J-f\|_2^4\right\}.$$

We use the identity

$$\mathbb{E}\left(X^2\right) = 2\int_0^{\infty} t\mathbb{P}\left(|X|>t\right)\mathrm{d}t, \tag{17}$$

which holds for any random variable X such that $E(X^2) < +\infty$.

We derive from (10) that

$$\mathbb{E}\left[U_n^2(H_J)\right] \leq C\left(\frac{2^J M}{n^2}+\frac{M^2}{n^2}+\frac{2^{2J}}{n^4}\right). \tag{18}$$

Since

$$\mathrm{pen}^2(J) \geq \frac{2^J M}{n^2}\log(2)+\frac{M^2}{n^2}\log^2(2)+\frac{2^{2J}}{n^4}\log^4(2),$$

this implies that

$$\mathbb{E}\left[U_n^2(H_J)\right] \leq C\mathrm{pen}^2(J).$$

In order to give an upper bound for $\mathbb{E}(P_n^2(h_J))$, we set

$$u(y) = 2\sqrt{2y\|f\|_\infty}\frac{\|f_J-f\|_2}{\sqrt{n}}+\frac{4}{3}\|f\|_\infty\frac{y}{n}.$$

We deduce from Lemma 2 that for any $y>0$,

$$\mathbb{P}\left(|P_n(h_J)|>u(y)\right) \leq 2\exp(-y).$$

Using (17), this leads to

$$\mathbb{E}\left[P_n^2(h_J)\right] \leq C\left(\frac{M\|f_J-f\|_2^2}{n}+\frac{M^2}{n^2}\right).$$

Using the inequality $2ab \leq a^2+b^2$, one obtains that

$$\mathbb{E}\left[P_n^2(h_J)\right] \leq C\left(\|f_J-f\|_2^4+\frac{M^2}{n^2}\right). \tag{19}$$

Collecting these evaluations,

$$\mathbb{E}\left[(V_J)_-^2\right] \leq C\left(\|f_J-f\|_2^4+\mathrm{pen}^2(J)\right).$$

This concludes the proof of Proposition 3.

5.2. **Proof of Proposition 2**

Let A denote the event $\left\{\forall J\in\mathcal{J}_n, \hat{\theta}_J+\frac{1}{2}\geq\theta_J\right\}$. We first give an upper bound for

$$\mathbb{E}\left[\left(\widehat{\theta}-\theta-\frac{2}{n}\sum_{i=1}^{n}(f(X_i)-\theta)\right)^2 \mathbb{1}_A\right].$$

Using the same notations as in the proof of Proposition 3,

$$\mathbb{E}\left[\left(\widehat{\theta}-\theta-\frac{2}{n}\sum_{i=1}^{n}(f(X_i)-\theta)\right)^2 \mathbb{1}_A\right] \leq \sum_{J\in\mathcal{J}_n}\mathbb{E}\left[(V_J)_+^2 \ \mathbb{1}_A\right]+\inf_{J\in\mathcal{J}_n}\mathbb{E}\left[(V_J)_-^2\right].$$

Let $C_0(M)=\inf\left\{J\in\mathbb{N}, \frac{2^J}{\log(2^J+1)}\geq M^2\right\}$. We first show that for all $J\in\mathcal{J}_n$ such that $2^J\geq C_0(M)$,

$$\mathbb{E}\left[(V_J)_+^2 \ \mathbb{1}_A\right]\leq\frac{C(M)}{n^2}2^{2J}\exp(-y_J).$$

To prove this result, we use the identity

$$\mathbb{E}\left[(V_J)_+^2 \ \mathbb{1}_A\right]=2\int_0^\infty t\mathbb{P}\left(V_J>t\cap A\right)\mathrm{d}t.$$

We set $y_J=3\log(2^J+1)$. We derive from (11) and (13) that if, on the event A,

$$\mathrm{pen}(J)\geq u_J(\sqrt{2}y_J)+\frac{10My_J}{3n}, \tag{20}$$

(where u_J is defined by (9)), then we have that

$$\mathbb{P}\left(\left(V_J>u_J(t\sqrt{2})+\frac{10}{3}M\frac{t}{n}\right)\cap A\right)\leq 6.6\exp-(t+y_J).$$

This implies that

$$\mathbb{E}\left[(V_J)_+^2 \ \mathbb{1}_A\right]\leq\frac{C(M)}{n^2}2^{2J}\exp(-y_J).$$

Let us show that (20) holds for all $J\in\mathcal{J}_n$ such that $J\geq C_0(M)$. Setting $x_J=\log(2^J+1)$, on the event A, for all $J\in\mathcal{J}_n$ such that $J\geq C_0(M)$,

$$\begin{aligned} n\mathrm{pen}(J) &\geq \kappa\sqrt{(\theta_J+1/2)2^Jx_J}\\ &\geq \frac{\kappa}{\sqrt{2}}\sqrt{\theta_J2^Jx_J}+\frac{\kappa}{2}\sqrt{2^Jx_J}\\ &\geq \frac{\kappa}{\sqrt{2}}\sqrt{\theta_J2^Jx_J}+\frac{\kappa}{4}Mx_J+\frac{\kappa}{4}\sqrt{2^Jx_J}. \end{aligned}$$

Since for all $J\in\mathcal{J}_n$, $2^J\leq n^2/\log^3(n)$,

$$\frac{1}{\sqrt{2^Jx_J}}\frac{2^Jx_J^2}{n}\leq\frac{\log^{3/2}(n^2+1)}{\log^{3/2}(n)}\leq C_1'$$

where C_1' is some positive constant.

This implies that, on the event A, for all $J\in\mathcal{J}_n$ such that $2^J\geq C_0(M)$,

$$n\mathrm{pen}(J)\geq\frac{\kappa}{\sqrt{2}}\sqrt{\theta_J2^Jx_J}+\frac{\kappa}{4}Mx_J+\frac{\kappa}{4C_1'}\frac{2^Jx_J^2}{n}.$$

By definition of u_J given in (9), we obtain that if $\kappa_0'\geq\max(32\sqrt{2}C_0+54.256C_1'C_0)$, then, on the event A, for all $J\in\mathcal{J}_n$ such that $2^J\geq C_0(M)$, (20) holds.

If $2^J \le C_0(M)$, by using the inequalities (12) and (18), one obtains

$$\begin{aligned}\mathbb{E}\left((V_J)_+^2 \ \mathbb{1}_A\right) &\le \mathbb{E}\left((V_J)_+^2\right)\\ &\le 2\left\{\mathbb{E}\left[U_n^2(H_J)\right] + \mathbb{E}\left[\left(P_n(h_J) - \|f - f_J\|_2^2\right)_+^2\right]\right\}\\ &\le C(M)\left(\frac{2^J}{n^2} + \frac{2^{2J}}{n^4}\right)\\ &\le C(M)\frac{1}{n^2}\end{aligned}$$

possibly enlarging $C(M)$ since $2^J \le C_0(M)$. Collecting these evaluations, one obtains that

$$\sum_{J\in\mathcal{J}_n} \mathbb{E}\left((V_J)_+^2 \ \mathbb{1}_A\right) \le C(M)\frac{1}{n^2} \tag{21}$$

since $\sum_{J\in\mathcal{J}_n} 2^{2J}\exp(-y_J) \le \sum_{J\ge 0} 2^{-J}$.

We now give an upper bound for $\mathbb{E}\left[(V_J)_-^2\right]$.

$$\mathbb{E}\left[(V_J)_-^2\right] \le 4\left\{\mathbb{E}\left[U_n^2(H_J)\right] + \mathbb{E}\left[P_n^2(h_J)\right] + \|f_J - f\|_2^4 + \mathbb{E}\left[\mathrm{pen}^2(J)\right]\right\}.$$

We use inequalities (18) and (19) to control $\mathbb{E}\left[U_n^2(H_J)\right]$ and $\mathbb{E}\left[P_n^2(h_J)\right]$. Moreover, for all $J \in \mathcal{J}_n$

$$\mathrm{pen}^2(J) = \frac{\kappa^2}{n^2}(\widehat{\theta}_J + 1)2^J x_J,$$

and since $\mathbb{E}(\widehat{\theta}_J) = \theta_J \le \|f\|_\infty \le M$,

$$\mathbb{E}\left[\mathrm{pen}^2(J)\right] \le C(M+1)\frac{2^J x_J}{n^2}\cdot$$

It follows that for all $J \in \mathcal{J}_n$,

$$\mathbb{E}\left[(V_J)_-^2\right] \le C\left(\|f_J - f\|_2^4 + (M+1)\frac{2^J x_J}{n^2} + \frac{M^2}{n^2}\right)\cdot \tag{22}$$

It remains to evaluate

$$\mathbb{E}\left[\left(\widehat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^n (f(X_i) - \theta)\right)^2 \ \mathbb{1}_{A^C}\right].$$

We first give an upper bound for $\mathbb{P}(A^C)$. Note that

$$\mathbb{P}(A^C) \le \sum_{J\in\mathcal{J}_n} \mathbb{P}\left(\widehat{\theta}_J - \theta_J \le -1/2\right) \le \sum_{J\in\mathcal{J}_n} \mathbb{P}\left(|\widehat{\theta}_J - \theta_J| > 1/2\right).$$

Since

$$\widehat{\theta}_J - \theta_J = U_n(H_J) + P_n(2f_J),$$

$$\mathbb{P}\left(|\widehat{\theta}_J - \theta_J| > \frac{1}{2}\right) \le \mathbb{P}\left(|U_n(H_J)| > \frac{1}{4}\right) + \mathbb{P}\left(|P_n(2f_J)| > \frac{1}{4}\right)\cdot$$

Let us first give an upper bound for $\mathbb{P}\left(|U_n(H_J)| > \frac{1}{4}\right)$.

Since $\forall n \geq 2$, $1/(n-1) \leq 2/n$, we deduce from (10) that

$$\mathbb{P}\left(|U_n(H_J)| > \frac{2C_0}{n}\left(\sqrt{2^J M t} + Mt + \frac{2^J t^2}{n}\right)\right) \leq 5.6\exp(-t).$$

Let

$$t_0(n, M) = \inf\left\{\frac{\log^3(n)}{(24C_0)^2 M}, \frac{n}{24C_0 M}, \frac{\log^{3/2}(n)}{\sqrt{24C_0}}\right\}.$$

Note that, since $2^J \leq n^2/\log^3(n)$,

$$\frac{2C_0}{n}\left(\sqrt{2^J M t_0(n,M)} + Mt_0(n,M) + \frac{2^J t_0^2(n,M)}{n}\right) \leq \frac{1}{4},$$

which implies that

$$\mathbb{P}\left(|U_n(H_J)| > \frac{1}{4}\right) \leq 5.6\exp(-t_0(n,M)) \leq \frac{C(M)}{n^8}.$$

Let us now give an upper bound for $\mathbb{P}\left(|P_n(2f_J)| > \frac{1}{4}\right)$. By Lemma 2,

$$\mathbb{P}\left(|P_n(2f_J)| > 2M\frac{\sqrt{2y}}{\sqrt{n}} + \frac{2My}{3n}\right) \leq 2\exp(-y).$$

Let

$$y_0(n, M) = \inf\left\{\frac{n}{2^9 M^2}, \frac{3}{16}\frac{n}{M}\right\}.$$

Since

$$2M\frac{\sqrt{2y_0(n,M)}}{\sqrt{n}} + \frac{2My_0(n,M)}{3n} \leq \frac{1}{4},$$

we obtain

$$\mathbb{P}\left(|P_n(2f_J)| > \frac{1}{4}\right) \leq 2\exp\left(-y_0(n,M)\right) \leq \frac{C(M)}{n^8}.$$

Since the cardinality of $\mathcal{J}_n$ is not larger that n^2, we finally obtain that

$$\mathbb{P}\left(A^C\right) \leq \frac{C(M)}{n^6}.$$

It follows from the definition of $\widehat{\theta}_J$ that $0 \leq \widehat{\theta}_J \leq 2^J$ for all $J \in \mathbb{N}$. This implies that for all $J \in \mathcal{J}_n$

$$0 \leq \mathrm{pen}(J) \leq Cn.$$

Hence,

$$\begin{aligned} |\widehat{\theta}| &= \left|\sup_{J\in\mathcal{J}_n}\left(\widehat{\theta}_J - \mathrm{pen}(J)\right)\right| \\ &\leq \sup_{J\in\mathcal{J}_n}\left(2^J + Cn\right) \\ &\leq Cn^2. \end{aligned}$$

Moreover, $0 \leq \theta = \int f^2 \leq M$ and $|\frac{2}{n}\sum_{i=1}^n (f(X_i)-\theta)| \leq 4M$. These evaluations imply that

$$\left(\widehat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^n (f(X_i)-\theta)\right)^2 \leq C(M)n^4$$

and therefore,

$$\mathbb{E}\left[\left(\widehat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^n (f(X_i)-\theta)\right)^2 \mathbb{1}_A^C\right] \leq C(M)n^4\mathbb{P}(A^C) \leq \frac{C(M)}{n^2}\cdot \tag{23}$$

Collecting (21), (22) and (23), we conclude the proof of Proposition 2.

5.3. **Proof of Theorem 1**

We now apply Proposition 2. We set

$$J_n = \left[\log_2\left(\left(\frac{n^2R^4}{(M+1)\log(nR^2)}\right)^{1/(1+4\alpha)}\right)\right] + 1$$

where $[x]$ denotes the integer part of x.

Since $\alpha > 0$, $J_n \in \mathcal{J}_n$ for $n \geq n_0(\alpha, R, M)$.

Note that

$$\left(\frac{n^2R^4}{(M+1)\log(nR^2)}\right)^{1/(1+4\alpha)} \leq 2^{J_n} \leq 2\left(\frac{n^2R^4}{(M+1)\log(nR^2)}\right)^{1/(1+4\alpha)}$$

and that there exists $n_0(\alpha, R, M)$ such that for all $n \geq n_0(\alpha, R, M)$, $J_n \geq 0$ and

$$\log(2^{J_n}+1) \leq C(\alpha)\log(nR^2).$$

Noting that if $f \in \mathcal{B}_{\alpha,2,\infty}(R)$, for all $J \in \mathbb{N}$

$$\|f - f_J\|_2^2 = \sum_{j\geq J}\sum_{k\in\mathbb{Z}} \beta_{j,k}^2(f) \leq C(\alpha)R^2 2^{-2J\alpha},$$

we derive from Proposition 2 that

$$\sup_{f,\beta(f)\in\mathcal{B}_{\alpha,2,\infty}(R), \|f\|_\infty\leq M} \mathbb{E}\left[\left(\widehat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^n (f(X_i)-\theta)\right)^2\right] \leq$$

$$C(\alpha)\left[R^4 2^{-4J_n\alpha} + \frac{(M+1)2^{J_n}}{n^2}\log(2^{J_n}+1)\right] + \frac{C(M)}{n^2}\cdot$$

By definition of J_n, possibly enlarging $n_0(\alpha, R, M)$, for all $n \geq n_0(\alpha, R, M)$,

$$\sup_{f,\beta(f)\in\mathcal{B}_{\alpha,2,\infty}(R), \|f\|_\infty\leq M} \mathbb{E}\left[\left(\widehat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^n (f(X_i)-\theta)\right)^2\right] \leq C(\alpha)(R(M+1)^\alpha)^{\frac{4}{1+4\alpha}}\left(\frac{\sqrt{\log(nR^2)}}{n}\right)^{\frac{8\alpha}{1+4\alpha}}.$$

In order to prove (3) and (4), we use the inequality

$$\mathbb{E}\left[\left(\widehat{\theta}-\theta\right)^2\right] \leq 2\mathbb{E}\left[\left(\widehat{\theta}-\theta-\frac{2}{n}\sum_{i=1}^{n}(f(X_i)-\theta)\right)^2\right] + 2\mathbb{E}\left[\left(\frac{2}{n}\sum_{i=1}^{n}(f(X_i)-\theta)\right)^2\right]$$

together with the fact that

$$\mathbb{E}\left[\left(\frac{2}{n}\sum_{i=1}^{n}(f(X_i)-\theta)\right)^2\right] \leq C\frac{M^2}{n}\cdot$$

Finally (5) and (6) derive from the fact that if $\alpha > 1/4$, then

$$n\mathbb{E}\left[\left(\widehat{\theta}-\theta-\frac{2}{n}\sum_{i=1}^{n}(f(X_i)-\theta)\right)^2\right] \overset{n\to\infty}{\to} 0,$$

and from the fact that

$$\frac{2}{\sqrt{n}}\sum_{i=1}^{n}(f(X_i)-\theta) \overset{\mathcal{D}}{\to} \mathcal{N}\left(0,\ \mathrm{Var}(2f(X_1)\right).$$

6. Appendix

6.1. Proof of Lemma 1

We deduce from Theorem 3.4 in Houdré and Reynaud [13] that there exists some absolute constant $C > 0$ such that for all $t > 0$,

$$\mathbb{P}\left(\left|\sum_{l\neq l'=1}^{n} H_J(X_l, X_{l'})\right| > C\left(A_1\sqrt{t} + A_2 t + A_3 t^{3/2} + A_4 t^2\right)\right) \leq 5.6\exp(-t)$$

where

$$\begin{aligned}
A_1^2 &= n(n-1)\mathbb{E}(H_J^2(X_1,X_2)),\\
A_2 &= \sup\left\{\left|\mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{i-1} H_J(X_1,X_2)a_i(X_1)b_j(X_2)\right)\right|, \mathbb{E}\left(\sum_{i=1}^{n} a_i^2(X_1)\right) \leq 1, \mathbb{E}\left(\sum_{j=1}^{n} b_j^2(X_1)\right) \leq 1\right\},\\
A_3^2 &= n\sup_x\left\{\mathbb{E}_{X_2}(H_J^2(x,X_2))\right\},\\
A_4 &= \sup_{x,y}|H_J(x,y)|.
\end{aligned}$$

Let us now evaluate A_1, A_2, A_3 and A_4.

Evaluation of A_1.

$$\begin{aligned}
H_J^2(X_1,X_2) = \sum_{k\in\mathbb{Z}} &\left(\phi_{J,k}(X_1)-\alpha_{J,k}\right)^2\left(\phi_{J,k}(X_2)-\alpha_{J,k}\right)^2\\
&+ \sum_{k\neq k'\in\mathbb{Z}} \left(\phi_{J,k}(X_1)-\alpha_{J,k}\right)\left(\phi_{J,k'}(X_1)-\alpha_{J,k'}\right)\left(\phi_{J,k}(X_2)-\alpha_{J,k}\right)\left(\phi_{J,k'}(X_2)-\alpha_{J,k'}\right).
\end{aligned}$$

This implies that

$$\begin{aligned}
\mathbb{E}\left(H_J^2(X_1,X_2)\right) &= \sum_{k\in\mathbb{Z}}\left[\int \left(\phi_{J,k}(x)-\alpha_{J,k}\right)^2 f(x)\mathrm{d}x\right]^2 \\
&\quad + \sum_{k\neq k'\in\mathbb{Z}}\left[\int \left(\phi_{J,k}(x)-\alpha_{J,k}\right)\left(\phi_{J,k'}(x)-\alpha_{J,k'}\right) f(x)\mathrm{d}x\right]^2 \\
&= \sum_{k\in\mathbb{Z}}\left[\int \phi_{J,k}^2 f-\alpha_{J,k}^2\right]^2 + \sum_{k\neq k'\in\mathbb{Z}}\alpha_{J,k}^2\alpha_{J,k'}^2 \\
&\leq \sum_{k\in\mathbb{Z}}\left(\int \phi_{J,k}^2 f\right)^2 + \left(\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2\right)^2 \\
&\leq 2^J\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2 + \left(\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2\right)^2 \quad \text{since } \phi_{J,k}^2 = 2^{J/2}\phi_{J,k} \\
&\leq 2^{J+1}\theta_J
\end{aligned}$$

since $\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2=\theta_J$ and since, setting $I_{J,k}=]k/2^J,(k+1)/2^J]$,

$$\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2 = 2^J\sum_{k\in\mathbb{Z}}\int_{I_{J,k}} f\int_{I_{J,k}} f \leq 2^J\sum_{k\in\mathbb{Z}}\int_{I_{J,k}} f\leq 2^J.$$

It follows that

$$A_1\leq\sqrt{2n(n-1)2^J\theta_J}.$$

Evaluation of A_2.
Let $(a_i)_{1\leq i\leq n}$ and $(b_j)_{1\leq j\leq n}$ satisfy $\mathbb{E}(\sum_{i=1}^n a_i^2(X_1))\leq 1$ and $\mathbb{E}(\sum_{j=1}^n b_j^2(X_1))\leq 1$.

We recall that $f_J=\sum_{k\in\mathbb{Z}}\alpha_{J,k}\phi_{J,k}$.

$$\sum_{i=1}^n\sum_{j=1}^{i-1}\mathbb{E}\left(H_J(X_1,X_2)a_i(X_1)b_j(X_2)\right) = \sum_{i=1}^n\sum_{j=1}^{i-1}\sum_{k\in\mathbb{Z}}\int\left(\phi_{J,k}-\alpha_{J,k}\right)a_i f\int\left(\phi_{J,k}-\alpha_{J,k}\right)b_j f$$

$$=\sum_{i=1}^n\sum_{j=1}^{i-1}\sum_{k\in\mathbb{Z}}\int\phi_{J,k}a_i f\int\phi_{J,k}b_j f+\sum_{i=1}^n\sum_{j=1}^{i-1}\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2\int a_i f\int b_j f-\sum_{i=1}^n\sum_{j=1}^{i-1}\left(\int a_i f\int f_J b_j f+\int b_j f\int f_J a_i f\right).$$

Using repeatedly the Cauchy-Schwarz inequality,

$$\begin{aligned}
\left|\sum_{k\in\mathbb{Z}}\int\phi_{J,k}a_i f\int\phi_{J,k}b_j f\right| &\leq \left[\sum_{k\in\mathbb{Z}}\left(\int\phi_{J,k}a_i f\right)^2\right]^{1/2}\left[\sum_{k\in\mathbb{Z}}\left(\int\phi_{J,k}b_j f\right)^2\right]^{1/2} \\
&\leq 2^J\left[\sum_{k\in\mathbb{Z}}\int_{I_{J,k}} a_i^2 f\int_{I_{J,k}} f\right]^{1/2}\left[\sum_{k\in\mathbb{Z}}\int_{I_{J,k}} b_j^2 f\int_{I_{J,k}} f\right]^{1/2} \\
&\leq 2^J\left(\frac{\|f\|_\infty}{2^J}\int a_i^2 f\right)^{1/2}\left(\frac{\|f\|_\infty}{2^J}\int b_j^2 f\right)^{1/2} \\
&\leq \|f\|_\infty\left(\int a_i^2 f\right)^{1/2}\left(\int b_j^2 f\right)^{1/2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\left|\sum_{k\in\mathbb{Z}}\int\phi_{J,k}a_i f\int\phi_{J,k}b_j f\right| &\leq \|f\|_\infty\sum_{i=1}^{n}\left(\int a_i^2 f\right)^{1/2}\sum_{j=1}^{n}\left(\int b_j^2 f\right)^{1/2}\\ &\leq \|f\|_\infty n\left(\sum_{i=1}^{n}\int a_i^2 f\sum_{j=1}^{n}\int b_j^2 f\right)^{1/2}\\ &\leq n\|f\|_\infty\end{aligned}$$

since by assumption $\sum_{i=1}^n\int a_i^2 f\leq 1$ and $\sum_{j=1}^n\int b_j^2 f\leq 1$.

Moreover, still using repeatedly the Cauchy-Schwarz inequality, the assumption on the a_i's and b_j's, together with the fact that $\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2\leq\int f^2\leq\|f\|_\infty$ since $\int f=1$, one obtains

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2\left|\int a_i f\int b_j f\right| &\leq \int f^2\sum_{i=1}^{n}\int|a_i f|\sum_{j=1}^{n}\int|b_j f|\\ &\leq \|f\|_\infty\sum_{i=1}^{n}\left(\int a_i^2 f\int f\right)^{1/2}\sum_{j=1}^{n}\left(\int b_j^2 f\int f\right)^{1/2}\\ &\leq \|f\|_\infty n.\end{aligned}$$

Finally, using that $\|f_J\|_\infty\leq\|f\|_\infty$,

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\int a_i f\int f_J b_j f\right|\leq\sum_{i=1}^{n}\sum_{j=1}^{n}\int|f_J b_j|f\int|a_i|f\leq\|f\|_\infty n.$$

$\sum_{i=1}^n\sum_{j=1}^n|\int b_j f\int f_J a_i f|$ is also controlled by $\|f\|_\infty n$. We finally obtain that

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

Evaluation of A_3.

$$\begin{aligned}\mathbb{E}_{X_2}(H_J^2(x,X_2)) &= \sum_{k\in\mathbb{Z}}\left(\int(\phi_{J,k}-\alpha_{J,k})^2 f\right)(\phi_{J,k}(x)-\alpha_{J,k})^2\\ &\quad+\sum_{k\neq k'\in\mathbb{Z}}\left(\int(\phi_{J,k}-\alpha_{J,k})(\phi_{J,k'}-\alpha_{J,k'})f\right)(\phi_{J,k}(x)-\alpha_{J,k})(\phi_{J,k'}(x)-\alpha_{J,k'})\\ &=\sum_{k\in\mathbb{Z}}\int\phi_{J,k}^2 f(\phi_{J,k}(x)-\alpha_{J,k})^2-\left(\sum_{k\in\mathbb{Z}}\alpha_{J,k}(\phi_{J,k}(x)-\alpha_{J,k})\right)^2\\ &\leq\|f\|_\infty\sum_{k\in\mathbb{Z}}\left(\phi_{J,k}^2(x)+\alpha_{J,k}^2\right)\end{aligned}$$

since $\int\phi_{J,k}^2=1$. Noticing that $\sum_{k\in\mathbb{Z}}\phi_{J,k}^2(x)=2^J$ and that $\sum_{k\in\mathbb{Z}}\alpha_{J,k}^2\leq 2^J$,

$$\sup_x\mathbb{E}_{X_2}(H_J^2(x,X_2))\|_\infty\leq 2^{J+1}\|f\|_\infty.$$

Therefore

$$A_3 \le \sqrt{2^{J+1} n \|f\|_\infty}.$$

Evaluation of A_4.
It is easy to see that

$$A_4 \le 2^{J+2}.$$

We now derive from Houdré and Reynaud's inequality that there exists $C > 0$ such that for all $t > 0$,

$$\mathbb{P}\left(\frac{1}{n(n-1)} \left| \sum_{l \ne l'=1}^{n} H_J(X_l, X_{l'}) \right| > \frac{C}{n-1} \left\{ \sqrt{2^J \theta_J t} + \|f\|_\infty t + \sqrt{2^J \|f\|_\infty} \frac{t^{3/2}}{\sqrt{n}} + \frac{2^J t^2}{n} \right\} \right) \le 5.6 \exp(-t).$$

Since $2ab \le a^2 + b^2$,

$$2\sqrt{2^J \|f\|_\infty} \frac{t^{3/2}}{\sqrt{n}} \le \|f\|_\infty t + \frac{2^J t^2}{n}.$$

This concludes the proof of the lemma.

References

[1] P. Bickel and Y. Ritov, Estimating integrated squared density derivatives: sharp best order of convergence estimates. *Sankhya Ser. A.* **50** (1989) 381–393.

[2] L. Birgé and P. Massart, Estimation of integral functionals of a density. *Ann. Statist.* **23** (1995) 11–29.

[3] L. Birgé and P. Massart, Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli* **4** (1998) 329–375.

[4] L. Birgé and Y. Rozenholc, How many bins should be put in a regular histogram. *Technical Report Université Paris 6 et 7* (2002).

[5] J. Bretagnolle, A new large deviation inequality for U-statistics of order 2. *ESAIM: PS* **3** (1999) 151–162.

[6] D. Donoho and M. Nussbaum, Minimax quadratic estimation of a quadratic functional. *J. Complexity* **6** (1990) 290–323.

[7] S. Efroïmovich and M. Low, On Bickel and Ritov's conjecture about adaptive estimation of the integral of the square of density derivatives. *Ann. Statist.* **24** (1996) 682–686.

[8] S. Efroïmovich and M. Low, On optimal adaptive estimation of a quadratic functional. *Ann. Statist.* **24** (1996) 1106–1125.

[9] M. Fromont and B. Laurent, Adaptive goodness-of-fit tests in a density model. *Technical report.* Université Paris 11 (2003).

[10] G. Gayraud and K. Tribouley, Wavelet methods to estimate an integrated quadratic functional: Adaptivity and asymptotic law. *Statist. Probab. Lett.* **44** (1999) 109–122.

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

[12] W. Haïdle, G. Kerkyacharian, D. Picard, A. Tsybakov, Wavelets, Approximations and statistical applications. *Lect. Notes Stat.* **129** (1998).

[13] C. Houdré and P. Reynaud-Bouret, Exponential inequalities for U-statistics of order two with constants, in *Euroconference on Stochastic inequalities and applications.* Barcelona. Birkhauser (2002).

[14] I.A. Ibragimov, A. Nemirovski and R.Z. Hasminskii, Some problems on nonparametric estimation in Gaussian white noise. *Theory Probab. Appl.* **31** (1986) 391–406.

[15] I. Johnstone, Chi-square oracle inequalities. *State of the art in probability and statistics (Leiden 1999) - IMS Lecture Notes Monogr. Ser., 36. Inst. Math. Statist., Beachwood, OH* (1999) 399–418.

[16] B. Laurent, Efficient estimation of integral functionals of a density. *Ann. Statist.* **24** (1996) 659–681.

[17] B. Laurent, Estimation of integral functionals of a density and its derivatives. *Bernoulli* **3** (1997) 181–211.

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