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

Béatrice Laurent

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 $\int f^2(x)dx$ 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.


1. Introduction

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

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 $\alpha$. They built an estimator $\hat{\theta}_a$ 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 build 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:

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

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

© EDP Sciences, SMAI 2005
after some projection onto an orthonormal basis of $L^2([0,1])$. 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 ellipsoid.

Efroimovich and Low [8] were the first to propose adaptive estimators of quadratic functionals in the framework of the Gaussian sequence model. The estimator $\hat{\theta}_n$ 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 $\alpha$, provided that the sequence $(\beta_\lambda)_{\lambda \geq 1}$ satisfies to the condition $\beta_\lambda^{2\alpha+1} \leq R^2$ for all $\lambda$ (which means that $(\beta_\lambda)_{\lambda \geq 1}$ belongs to some hyperrectangle) one has

\[ \mathbb{E} \left( (\hat{\theta}_n - \theta)^2 \right) \leq C(R, \alpha) \left( n^{-2} \log(n) \right)^{4\alpha/(1+4\alpha)}, \text{ if } \alpha \leq 1/4. \]

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 $\beta$ 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 $\theta$.

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 $\text{pen} : \mathcal{M} \rightarrow \mathbb{R}^+$. The penalized estimator of $\theta = \sum_{\lambda \in \mathbb{N}^*} \beta_\lambda^2$ is defined as

\[ \hat{\theta} = \sup_{m \in \mathcal{M}} \left( \sum_{\lambda \in m} Y_\lambda^2 - \text{pen}(m) \right). \]

For suitable choices of the set $\mathcal{M}$ and of the penalty function, this estimator is adaptive over a general class 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 $\int_\mathbb{R} f^2(x) \, dx$ in a density model which is also 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 $\text{pen} : \mathcal{M} \rightarrow \mathbb{R}^+$. The penalized estimator of $\theta = \sum_{\lambda \in \mathbb{N}^*} \beta_\lambda^2$ is defined as

\[ \hat{\theta} = \sup_{m \in \mathcal{M}} \left( \sum_{\lambda \in m} Y_\lambda^2 - \text{pen}(m) \right). \]

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 $\int_\mathbb{R} f^2(x) \, dx$ 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 $L^2(\mathbb{R})$. Our aim is to estimate $\int_\mathbb{R} f^2(x) \, dx$. To do this, we consider the orthogonal projection of $f$ onto the Haar basis. We first introduce some
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 \( L^2(\mathbb{R}) \). 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
\[ \hat{\theta}_J = \frac{1}{n(n-1)} \sum_{k \in \mathbb{Z}} \sum_{l \neq 1} \phi_{J,k}(X_l) \phi_{J,k}(X_l^c). \tag{1} \]
In order to evaluate the quadratic risk of this estimator, we use the decomposition
\[ \mathbb{E} \left[ (\hat{\theta}_J - \theta)^2 \right] = \text{Bias}^2(\hat{\theta}_J) + \text{Var}(\hat{\theta}_J). \]
Since the expectation of \( \hat{\theta}_J \) equals \( \int f_J^2 \), \( |\text{Bias}(\hat{\theta}_J)| = \int (f - f_J)^2 \). Assuming that \( \|f\|_\infty \) is finite, one can easily show that
\[ \text{Var}(\hat{\theta}_J) \leq C(\|f\|_\infty) \left( \frac{2J}{n^2} + \frac{1}{n} \right) \]
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 \( B_{\alpha,2,\infty}(R) \) defined by
\[ B_{\alpha,2,\infty}(R) = \left\{ \beta = (\beta_{j,k})_{j \geq 0, k \in \mathbb{Z}}, \forall j \geq 0 \sum_{k \in \mathbb{Z}} \beta_{j,k}^2 \leq R^2 2^{-2j\alpha} \right\}. \]
This implies that for all \( J \in \mathbb{N} \),
\[ \int_R (f - f_J)^2 = \sum_{j \geq J} \sum_{k \in \mathbb{Z}} \beta_{j,k}^2 \leq C(\alpha) R^2 2^{-2J\alpha}, \]
where \( C(\alpha) \) is a constant depending on \( \alpha \). Choosing \( J \) in such a way that
\[ \frac{2J}{n^2} \approx R^2 4^{-J\alpha}, \]
we obtain that
\[
E \left[ (\hat{\theta}_J - \theta)^2 \right] \leq C \left( \|f\|_{\infty} \right) \left( R^{4/1+4\alpha} n^{-8\alpha/1+4\alpha} + \frac{1}{n} \right).
\]
The rate obtained here corresponds to the minimax rate for estimating \( \theta \) over Hölderian balls with index \( \alpha \) 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 \( \alpha \) 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
\[
\int (f - f_J)^2 + \frac{2^{J/2}}{n} = \int f^2 - \int f_J^2 + \frac{2^{J/2}}{n}
\]
Since \( \int f^2 \) does not depend on \( J \), this is equivalent to maximize the quantity \( \int f_J^2 - \frac{2^{J/2}}{n} \). \( \int f_J^2 \) is estimated without bias by \( \hat{\theta}_J \), hence we set
\[
\hat{J} = \arg\max_{J \in \mathcal{J}} \left[ \hat{\theta}_J - \text{pen}(J) \right]
\]
and
\[
\hat{\theta} = \hat{\theta}_J - \text{pen}(\hat{J}) = \sup_{J \in \mathcal{J}} \left[ \hat{\theta}_J - \text{pen}(J) \right],
\]
where \( \mathcal{J} \) is some subset of \( \mathbb{N} \) and \( \text{pen}(J) \) 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 \( \text{pen}(J) \) appearing in the definition of the estimator \( \hat{\theta} \) given by (2), the estimator \( \hat{\theta} \) 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
\[
\theta = \int_{\mathbb{R}} f^2(x) \, dx.
\]
Let \( \mathcal{J} = \{ J \in \mathbb{N} : 2^J \leq n^2/\log^3(n) \} \). For all \( J \in \mathbb{N} \), let \( \hat{\theta}_J \) be defined by (1).

There exists some absolute constant \( \kappa > 0 \) such that if we set for all \( J \in \mathcal{J} \)
\[
\text{pen}(J) = \frac{\kappa}{n} \left[ \sqrt{\hat{\theta}_J + 1} 2^J \log(2^J + 1) \right],
\]
then the estimator \( \hat{\theta} \) 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 \( \alpha \), \( R \) and \( M \) such that the following inequality holds for all \( n \geq n_0(\alpha, R, M) \):
\[
\sup_{f, \beta(f) \in \mathcal{B}_{\alpha, 2}(\mathbb{R}), \|f\|_{\infty} \leq M} E \left[ \left( \hat{\theta} - \theta - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) \right)^2 \right] \leq C(\alpha)(R(M + 1)^\alpha) \frac{n^{4\alpha}}{\log(nR^2)} \frac{\log(nR^2)}{n}^\frac{8\alpha}{1+4\alpha}.
\]

This leads to
QUADRATIC FUNCTIONALS OF A DENSITY

• for all α > 1/4, R > 0, and M > 0, there exists some integer \( n_1(\alpha, R, M) \) such that the following inequality holds for all \( n \geq n_1(\alpha, R, M) \):

\[
\sup_{f, \beta(f) \in B_{\alpha,2,\infty}(R), \|f\|_\infty \leq M} \mathbb{E} \left[ \left( \hat{\theta} - \theta \right)^2 \right] \leq C(\alpha) \frac{M^2}{n};
\]

(3)

• for all α ≤ 1/4, R > 0, and M > 0, there exists some integer \( n_2(\alpha, R, M) \) such that the following inequality holds for all \( n \geq n_2(\alpha, R, M) \):

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

(4)

If \( f \in L_\infty(\mathbb{R}) \) and \( \beta(f) \in B_{\alpha,2,\infty}(R) \) for some \( \alpha > 1/4 \) and \( R > 0 \), then

\[
\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{\mathbb{D}} N(0, \text{Var}(2f(X_1))) \quad \text{as} \quad n \to \infty
\]

(5)

\[
n\mathbb{E} \left[ (\hat{\theta} - \theta)^2 \right] \xrightarrow{\text{as}} \text{Var}(2f(X_1)).
\]

(6)

Comments:

1) We derive from (5) that if \( \beta(f) \in B_{\alpha,2,\infty}(R) \) for some \( \alpha > 1/4 \), then \( \hat{\theta} \) is an efficient estimator of \( \theta \) (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 H"olderian ball defined by

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

(7)

The minimax rate for estimating \( \int R \) over \( \mathcal{H}_\alpha(R) \) if \( \alpha \leq 1/4 \) is \( n^{-4\alpha/(1+4\alpha)} \) (see Birg"e 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 \( \hat{\theta}_n \) is an estimator of \( \theta \) based on the \( n \) sample \( X_1, \ldots, X_n \). If, for some \( \alpha > 1/4 \),

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

then for every \( \alpha < 1/4 \),

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

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

3) If the sequence \( \beta(f) \) belongs to \( B_{\alpha,2,\infty}(R) \) for some \( R > 0 \) and \( \alpha > 1/2 \), then \( f \) is uniformly bounded (see inequality (8.15) of Proposition (8.3) in Ha"rdle 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 \( \hat{\theta} \) defined by (2).

**Proposition 2.** Let \( X_1, \ldots, X_n \) be i.i.d. real random variables with common density \( f \) belonging to \( L_\infty(\mathbb{R}) \). Let \( \theta = \int_R f^2(x)dx \). Let \( J_n \) be a subset of \( \{ J \in \mathbb{N}, 2^J \leq \frac{n^2}{\log(n)} \} \). There exists some absolute constant \( \kappa_0 \) such that if we set for all \( J \in J_n \)

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

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

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

\( C \) 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 \( \hat{\theta} \), indeed

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

and

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

2) One can also deduce from Proposition 2 that if

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

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

\[
\sqrt{n} (\hat{\theta} - \theta) \overset{D}{\rightarrow} \mathcal{N}(0, \text{Var}(2f(X_1))).
\]

In this situation \( \hat{\theta} \) is an efficient estimator of \( \theta \) (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 \( \kappa_0 \), but this upper bound would be very large. A simulation study would be necessary to know how to calibrate \( \kappa_0 \) 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 $\alpha$ 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 $X_1, \ldots, X_n$ be i.i.d. real random variables with common density $f$ belonging to $L_\infty(\mathbb{R})$. We assume that $\|f\|_\infty \leq M$ where $M$ is known. Let $\theta = \int_\mathbb{R} f^2(x)dx$. Let $\mathcal{J}$ be a subset of $\mathbb{N}$. For all $J \in \mathcal{J}$, let $\hat{\theta}_J$ be defined by (1). There exists some constant $\kappa_0 > 0$ such that if we set for all $J \in \mathcal{J}$

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

with $\kappa \geq \kappa_0$ and

$$\hat{\theta} = \sup_{J \in \mathcal{J}} (\hat{\theta}_J - \text{pen}(J)),$$

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

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

where $C$ is some absolute constant.

**5.1. Proof of Proposition 3**

We use the canonical decomposition of the $U$-statistics $\hat{\theta}_J$. We denote by $U_n$ the process defined by $U_n(H) = (1/n(n-1)) \sum_{i \neq j=1}^n H(X_i, X_j)$ 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:

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

Let us denote by $V_J$ the variable

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

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

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

Moreover, since

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

$$\hat{\theta} - \theta - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) = \sup_{J \in \mathcal{J}} (V_J).$$
we obtain that
\[ \mathbb{E}\left[ \left( \sup_{J \in J} V_J \right)^2 \right] \leq \sum_{J \in J} \mathbb{E}\left[ (V_J)^2 \right] + \inf_{J \in J} \mathbb{E}\left[ (V_J)^2 \right]. \]

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
\[
U_n(H_J) = \frac{1}{n(n-1)} \sum_{i \neq j=1}^n H_J(X_i, X_j).
\]

There exists some absolute constant \( C_0 > 0 \) such that for all \( J \in \mathbb{N} \), for all \( t > 0 \),
\[
\mathbb{P}\left( |U_n(H_J)| > \frac{C_0}{n-1} \left( \sqrt{2j\theta_J} + \|f\|_\infty t + \frac{2j^2 t^2}{n} \right) \right) \leq 5.6 \exp(-t).
\]

The proof of the lemma is postponed to the Appendix.

We set for all \( t > 0 \)
\[
u_J(t) = \frac{C_0}{n-1} \left( \sqrt{2j\theta_J} + \|f\|_\infty t + \frac{2j^2 t^2}{n} \right).
\]

We derive from Lemma 1 that for all \( t \geq 0 \),
\[
\mathbb{P}\left( |U_n(H_J)| > \nu_J(t) \right) \leq 5.6 \exp(-t).
\]

Noticing that for all \( t_1 > 0 \) and all \( t_2 > 0 \)
\[
u_J \left( \frac{t_1 + t_2}{\sqrt{2}} \right) \leq \nu_J(t_1) + \nu_J(t_2),
\]
we derive from (10) that for all \( t \geq 0 \) and \( y_J \geq 0 \),
\[
\mathbb{P}\left( |U_n(H_J)| > \nu_J(\sqrt{2}y_J) + \nu_J(\sqrt{2}t) \right) \leq 5.6 \exp(-t + y_J).
\]

**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 $i = 1, \ldots, n$. 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 $t > 0$,

$$P\left(P_n(h_J) > 2\sqrt{2\|f\|_\infty \|f_J - f\|_2^2} \cdot \frac{\sqrt{7}}{\sqrt{n}} + \frac{4\|f\|_\infty t}{3n}\right) \leq \exp(-t).$$

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

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

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

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

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

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 $J \in \mathcal{J}$,

$$\text{pen}(J) \geq \frac{\kappa}{n} \left(\theta_J 2^J \log(2^J + 1) + M \log(2^J + 1) + \frac{2^J \log^2(2^J + 1)}{n}\right),$$

(14) implies that

$$\text{pen}(J) \geq u_J(\sqrt{2y_J}) + \frac{10\|f\|_\infty y_J}{3n}.$$ \hspace{1cm} (15)

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

$$P\left(V_J > u_J(t\sqrt{2}) + \frac{10M}{3n}\right) \leq 6.6 \exp(-(t + y_J)).$$ \hspace{1cm} (16)

We now use the identity

$$\mathbb{E}[(V_J)^2] = 2 \int_0^{\infty} t \mathbb{P}(V_J > t) \, dt.$$  

This identity, together with (16) leads to

$$\mathbb{E}[(V_J)^2] \leq C \left\{ \frac{2^J M^2}{n^2} + \frac{M^2}{n^2} + \frac{2^J}{n^4} \right\} \exp(-y_J).$$
Recalling that $J$ is a subset of $N$, $\sum_{J \in J} 2^{2J} \exp(-y_J) \leq \sum_{J \geq 0} 2^{-J}$ which implies that

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

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

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

We use the identity

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

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

We derive from (10) that

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

Since

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

this implies that

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

In order to give an upper bound for $E(P_n^2(h_J))$, 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 $y > 0$,

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

Using (17), this leads to

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

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

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

Collecting these evaluations, we have

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

This concludes the proof of Proposition 3.

5.2. Proof of Proposition 2

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

$$E \left[ \left( \hat{\theta} - \theta - \frac{2}{n} \sum_{i=1}^n (f(X_i) - \theta) \right)^2 \mathbb{I}_A \right].$$
Using the same notations as in the proof of Proposition 3, 
\[
\mathbb{E} \left[ \left( \hat{\theta} - \theta - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) \right)^2 \mathbb{1}_A \right] \leq \sum_{J \in \mathcal{J}_n} \mathbb{E} \left[ (V_J)^2 \mathbb{1}_A \right] + \inf_{J \in \mathcal{J}_n} \mathbb{E} \left[ (V_J)^2 \mathbb{1}_A \right].
\]
Let \( C_0(M) = \inf \left\{ J \in \mathbb{N}, \frac{y_J}{\log(2^J + 1)} \geq M^2 \right\} \). 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
\[
\mathbb{E} \left[ (V_J)^2 \mathbb{1}_A \right] = 2 \int_0^\infty t \mathbb{P} (V_J > t \cap A) \, dt.
\]
We set \( y_J = 3 \log(2^J + 1) \). We derive from (11) and (13) that if, on the event \( A \),
\[
\text{pen}(J) \geq u_J (\sqrt{2} y_J + \frac{10 M y_J}{3n}),
\]
(where \( u_J \) is defined by (9)), then we have that
\[
\mathbb{P} \left( \left( V_J > u_J (t \sqrt{2} + \frac{10 M t}{3n}) \right) \cap A \right) \leq 6.6 \exp(-t + y_J).
\]
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 \( J \geq C_0(M) \),
\[
n \text{pen}(J) \geq \frac{\sqrt{2} x_J}{\sqrt{2} x_J} \frac{\log(2^J + 1)}{x_J} \frac{1}{\log^3(n)} \leq C_1'
\]
where \( C_1' \) is some positive constant.
This implies that, on the event \( A \), for all \( J \in \mathcal{J}_n \) such that \( 2^J \geq C_0(M) \),
\[
n \text{pen}(J) \geq \frac{\sqrt{2} x_J}{\sqrt{2} x_J} \frac{\log(2^J + 1)}{x_J} \frac{1}{\log^3(n)} \leq C_1'
\]
By definition of \( u_J \) given in (9), we obtain that if \( \kappa_0' \geq \max(32\sqrt{2} C_0 + 54.256 C_1' C_0) \), then, on the event \( A \), for all \( J \in \mathcal{J}_n \) such that \( 2^J \geq C_0(M) \), (20) holds.
If $2^J \leq C_0(M)$, by using the inequalities (12) and (18), one obtains
\[
E \left( (V_J)^2 \mathbb{1}_A \right) \leq 2 \left\{ E \left[ U_n^2(H_J) \right] + E \left[ (P_n(h_J) - \|f - f_J\|_2^2) \right] \right\} \\
\leq C(M) \left( \frac{2^J}{n^2} + \frac{2^{2J}}{n^4} \right) \\
\leq C(M) \frac{1}{n^2}
\]
possibly enlarging $C(M)$ since $2^J \leq C_0(M)$. Collecting these evaluations, one obtains that
\[
\sum_{J \in \mathcal{J}_n} E \left( (V_J)^2 \mathbb{1}_A \right) \leq C(M) \frac{1}{n^2}
\] (21)
since $\sum_{J \in \mathcal{J}_n} 2^{2J} \exp(-y_J) \leq \sum_{J \geq 0} 2^{-J}$.

We now give an upper bound for $E \left[ (V_J)^2 \right]$.
\[
E \left[ (V_J)^2 \right] \leq 4 \left\{ E \left[ U_n^2(H_J) \right] + E \left[ P_n^2(h_J) \right] + \|f - f_J\|_2^2 + E \left[ \text{pen}^2(J) \right] \right\}.
\]

We use inequalities (18) and (19) to control $E \left[ U_n^2(H_J) \right]$ and $E \left[ P_n^2(h_J) \right]$. Moreover, for all $J \in \mathcal{J}_n$
\[
\text{pen}^2(J) = \frac{\kappa^2}{n^2} (\hat{\theta}_J + 1) 2^J f_J,
\]
and since $E(\hat{\theta}_J) = \theta_J \leq \|f\|_\infty \leq M$,
\[
E \left[ \text{pen}^2(J) \right] \leq C(M + 1) \frac{2^J f_J}{n^2}.
\]

It follows that for all $J \in \mathcal{J}_n$,
\[
E \left[ (V_J)^2 \right] \leq C \left( \|f - f_J\|_2^2 + (M + 1) \frac{2^J f_J}{n^2} + M^2 \right).
\] (22)

It remains to evaluate
\[
E \left[ \left( \hat{\theta} - \theta - \frac{2}{n} \sum_{i=1}^n (f(X_i) - \theta) \right) \mathbb{1}_A \right].
\]

We first give an upper bound for $\mathbb{P}(A_C)$. Note that
\[
\mathbb{P}(A_C) \leq \sum_{J \in \mathcal{J}_n} \mathbb{P} \left( \hat{\theta}_J - \theta_J \leq -1/2 \right) \leq \sum_{J \in \mathcal{J}_n} \mathbb{P} \left( |\hat{\theta}_J - \theta_J| > 1/2 \right).
\]

Since
\[
\hat{\theta}_J - \theta_J = U_n(H_J) + P_n(2f_J),
\]
\[
\mathbb{P} \left( |\hat{\theta}_J - \theta_J| > \frac{1}{2} \right) \leq \mathbb{P} \left( |U_n(H_J)| > \frac{1}{4} \right) + \mathbb{P} \left( |P_n(2f_J)| > \frac{1}{4} \right).
\]

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
\[
P \left( \left| U_n(H_J) \right| > \frac{2C_0}{n} \left( \sqrt{2^J M t_0(n, M) + M t_0(n, M) + \frac{2^J t_0^2(n, M)}{n}} \right) \right) \leq 5.6 \exp(-t).
\]
Let
\[
t_0(n, M) = \inf \left\{ \frac{\log^3(n)}{(24C_0)^2M}, \frac{n}{24C_0}, \frac{\log^{3/2}(n)}{\sqrt{24C_0}} \right\}.
\]
Note that, since $2^J \leq n^2/\log^3(n)$,
\[
\frac{2C_0}{n} \left( \sqrt{2^J M t_0(n, M) + M t_0(n, M) + \frac{2^J t_0^2(n, M)}{n}} \right) \leq \frac{1}{4},
\]
which implies that
\[
P \left( \left| U_n(H_J) \right| > \frac{1}{4} \right) \leq 5.6 \exp(-t_0(n, M)) \leq \frac{C(M)}{n^8}.
\]
Let us now give an upper bound for $P \left( \left| P_n(2f_J) \right| > \frac{1}{4} \right)$. By Lemma 2,
\[
P \left( \left| P_n(2f_J) \right| > 2M \frac{\sqrt{2y}}{\sqrt{n}} + \frac{2My}{3n} \right) \leq 2 \exp(-y).
\]
Let
\[
y_0(n, M) = \inf \left\{ \frac{n}{2^9 M^2}, \frac{3}{16}, \frac{n}{16 M} \right\}.
\]
Since
\[
2M \frac{\sqrt{2y_0(n, M)}}{\sqrt{n}} + \frac{2M y_0(n, M)}{3n} \leq \frac{1}{4},
\]
we obtain
\[
P \left( \left| P_n(2f_J) \right| > \frac{1}{4} \right) \leq 2 \exp(-y_0(n, M)) \leq \frac{C(M)}{n^8}.
\]
Since the cardinality of $\mathcal{J}_n$ is not larger than $n^2$, we finally obtain that
\[
P (A^C) \leq \frac{C(M)}{n^6}.
\]
It follows from the definition of $\hat{\theta}_J$ that $0 \leq \hat{\theta}_J \leq 2^J$ for all $J \in \mathbb{N}$. This implies that for all $J \in \mathcal{J}_n$
\[
0 \leq \text{pen}(J) \leq Cn.
\]
Hence,
\[
\left| \hat{\theta} \right| = \sup_{J \in \mathcal{J}_n} \left( \hat{\theta}_J - \text{pen}(J) \right) \leq \sup_{J \in \mathcal{J}_n} \left( 2^J + Cn \right) \leq Cn^2.
\]
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

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

and therefore,

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

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

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

where $\lfloor x \rfloor$ denotes the integer part of $x$.

Since $\alpha > 0$, $J_n \in J_n$ 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 $f \in B_{\alpha,2,\infty}(R)$, for all $J \in \mathbb{N}$

$$\|f - f_J\|_2^2 = \sum_{j \geq J} \sum_{k \in \mathbb{Z}} \beta_{j,k}^2(f) \leq C(\alpha)R^22^{-2J_n},$$

we derive from Proposition 2 that

$$\sup_{f,\beta(f) \in B_{\alpha,2,\infty}(R), \|f\|_\infty \leq M} \mathbb{E}\left[\left(\hat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^{n}(f(X_i) - \theta)\right)^2\right] \leq C(\alpha) \left[ R^{4\alpha - 4J_n}\alpha + \frac{(M+1)2^{J_n}}{n^2} \log(2^{J_n} + 1) \right] + \frac{C(M)}{n^2} $$

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

$$\sup_{f,\beta(f) \in B_{\alpha,2,\infty}(R), \|f\|_\infty \leq M} \mathbb{E}\left[\left(\hat{\theta} - \theta - \frac{2}{n}\sum_{i=1}^{n}(f(X_i) - \theta)\right)^2\right] \leq C(\alpha)(R(M+1)^\alpha)^\frac{\sqrt{\log(nR^2)}}{n^\gamma}.$$
In order to prove (3) and (4), we use the inequality
\[
\mathbb{E} \left( \hat{\theta} - \theta \right)^2 \leq 2\mathbb{E} \left( \hat{\theta} - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) \right)^2 + 2\mathbb{E} \left( \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) \right)^2
\]

together with the fact that
\[
\mathbb{E} \left( \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) \right)^2 \leq C \frac{M^2}{n}.
\]

Finally (5) and (6) derive from the fact that if \( \alpha > 1/4 \), then
\[
n\mathbb{E} \left( \hat{\theta} - \frac{2}{n} \sum_{i=1}^{n} (f(X_i) - \theta) \right)^2 \xrightarrow{n \to \infty} 0,
\]

and from the fact that
\[
\frac{2}{\sqrt{n}} \sum_{i=1}^{n} (f(X_i) - \theta) \xrightarrow{D} \mathcal{N}(0, \text{Var}(2f(X_1))).
\]

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 \( t > 0 \),
\[
\mathbb{P} \left( \left| \sum_{j \neq i=1}^{n} H_j(X_i, X_j) \right| > C \left( A_1 \sqrt{t} + A_2 t + A_3 t^{3/2} + A_4 t^2 \right) \right) \leq 5.6 \exp(-t)
\]

where

\[
A_1^2 = n(n-1)\mathbb{E}(H^2_2(X_1, X_2)),
\]

\[
A_2 = \sup \left\{ \mathbb{E} \left( \sum_{i=1}^{n-1} \sum_{j=1}^{i-1} H_j(X_1, X_2) a_i(X_1) b_j(X_2) \right), \mathbb{E} \left( \sum_{i=1}^{n} a_i^2(X_1) \right) \leq 1, \mathbb{E} \left( \sum_{j=1}^{n} b_j^2(X_1) \right) \leq 1 \right\},
\]

\[
A_3^2 = n \sup_x \{ \mathbb{E}_{X_2}(H^2_2(x, X_2)) \},
\]

\[
A_4 = \sup_{x,y} \mathbb{E}_{H_j(x,y)}.
\]

Let us now evaluate \( A_1, A_2, A_3 \) and \( A_4 \).

Evaluation of \( A_1 \).

\[
H^2_2(X_1, X_2) = \sum_{k \in \mathbb{Z}} (\phi_{j,k}(X_1) - \alpha_{j,k})^2 (\phi_{j,k}(X_2) - \alpha_{j,k})^2
\]

\[
+ \sum_{k \neq k' \in \mathbb{Z}} (\phi_{j,k}(X_1) - \alpha_{j,k}) (\phi_{j,k'}(X_1) - \alpha_{j,k'}) (\phi_{j,k}(X_2) - \alpha_{j,k}) (\phi_{j,k'}(X_2) - \alpha_{j,k'}).
\]
This implies that
\[
\mathbb{E} \left( H^2_j(X_1, X_2) \right) = \sum_{k \in \mathbb{Z}} \left[ \int (\phi_{J,k}(x) - \alpha_{J,k})^2 f(x) dx \right]^2 \\
+ \sum_{k \neq k' \in \mathbb{Z}} \left[ \int (\phi_{J,k}(x) - \alpha_{J,k})(\phi_{J,k'}(x) - \alpha_{J,k'}) f(x) dx \right]^2 \\
= \sum_{k \in \mathbb{Z}} \left[ \int \phi_{J,k}^2 f - \alpha_{J,k}^2 \right]^2 + \sum_{k \neq k' \in \mathbb{Z}} \alpha_{J,k} \alpha_{J,k'} \\
\leq \sum_{k \in \mathbb{Z}} \left( \int \phi_{J,k}^2 f \right)^2 + \left( \sum_{k \in \mathbb{Z}} \alpha_{J,k}^2 \right)^2 \\
\leq 2^J \sum_{k \in \mathbb{Z}} \alpha_{J,k}^2 + \left( \sum_{k \in \mathbb{Z}} \alpha_{J,k}^2 \right)^2 \quad \text{since } \phi_{J,k}^2 = 2^{J/2} \phi_{J,k} \\
\leq 2^{J+1} \theta_j
\]

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
\[
A_1 \leq \sqrt{2} \theta_j.
\]

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

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

Using repeatedly the Cauchy-Schwarz inequality,
\[
\left| \sum_{k \in \mathbb{Z}} \phi_{J,k} a_i f \int \phi_{J,k} b_j f \right| \leq \left[ \sum_{k \in \mathbb{Z}} \left( \int \phi_{J,k} a_i f \right)^2 \right]^{1/2} \left[ \sum_{k \in \mathbb{Z}} \left( \int \phi_{J,k} b_j f \right)^2 \right]^{1/2} \\
\leq 2^J \left[ \sum_{k \in \mathbb{Z}} \int_{I_{J,k}} a_i^2 f \int_{I_{J,k}} f \right]^{1/2} \left[ \sum_{k \in \mathbb{Z}} \int_{I_{J,k}} b_j^2 f \int_{I_{J,k}} f \right]^{1/2} \\
\leq 2^J \left( \|f\|_{\infty} \int a_i^2 f \right)^{1/2} \left( \|f\|_{\infty} \int b_j^2 f \right)^{1/2} \\
\leq \|f\|_{\infty} \left( \int a_i^2 f \right)^{1/2} \left( \int b_j^2 f \right)^{1/2}.
\]
QUADRATIC FUNCTIONALS OF A DENSITY

Therefore,
\[
\sum_{i=1}^{n} \sum_{j=1}^{n} \left| \sum_{k \in \mathbb{Z}} \phi_{j,k} a_i f \int \phi_{j,k} b_j f \right| \leq \|f\|_{\infty} \sum_{i=1}^{n} \left( \int a_i^2 f \right)^{1/2} \sum_{j=1}^{n} \left( \int b_j f \right)^{1/2}
\]
\[
\leq \|f\|_{\infty} n \left( \sum_{i=1}^{n} \int a_i^2 f \right) \sum_{j=1}^{n} \left( \int b_j f \right)^{1/2}
\]
\[
\leq n \|f\|_{\infty}
\]
since by assumption \(\sum_{i=1}^{n} a_i^2 f \leq 1\) and \(\sum_{j=1}^{n} b_j^2 f \leq 1\).

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 \(f = 1\), one obtains
\[
\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k \in \mathbb{Z}} \alpha_{j,k}^2 \left| \int a_i f \int b_j f \right| \leq \int f^2 \sum_{i=1}^{n} \left| a_i f \right| \sum_{j=1}^{n} \left| b_j f \right|
\]
\[
\leq \|f\|_{\infty} \sum_{i=1}^{n} \left( \int a_i^2 f \right)^{1/2} \sum_{j=1}^{n} \left( \int b_j^2 f \right)^{1/2}
\]
\[
\leq \|f\|_{\infty} n.
\]

Finally, using that \(\|f\|_{\infty} \leq \|f\|_{\infty}\),
\[
\sum_{i=1}^{n} \sum_{j=1}^{n} \left| \int a_i f \int f_j b_j f \right| \leq \sum_{i=1}^{n} \sum_{j=1}^{n} \left| f_j b_j f \int a_i f \right| \leq \|f\|_{\infty} n.
\]

\(\sum_{i=1}^{n} \sum_{j=1}^{n} | f_j b_j f \int f_j a_i f |\) is also controlled by \(\|f\|_{\infty} n\). We finally obtain that
\[
A_2 \leq 4n \|f\|_{\infty}.
\]

**Evaluation of \(A_3\).**

\[
\mathbb{E}_{X_2}(H^2_2(x, X_2)) = \sum_{k \in \mathbb{Z}} \left( \int (\phi_{j,k} - \alpha_{j,k})^2 f \right) (\phi_{j,k}(x) - \alpha_{j,k})^2
\]
\[
+ \sum_{k \neq k' \in \mathbb{Z}} \left( \int (\phi_{j,k} - \alpha_{j,k}) (\phi_{j,k'} - \alpha_{j,k'}) f \right) (\phi_{j,k}(x) - \alpha_{j,k}) (\phi_{j,k'}(x) - \alpha_{j,k'})
\]
\[
= \sum_{k \in \mathbb{Z}} \phi_{j,k}^2 f (\phi_{j,k}(x) - \alpha_{j,k})^2 - \left( \sum_{k \in \mathbb{Z}} \alpha_{j,k} (\phi_{j,k}(x) - \alpha_{j,k}) \right)^2
\]
\[
\leq \|f\|_{\infty} \sum_{k \in \mathbb{Z}} (\phi_{j,k}^2(x) + \alpha_{j,k}^2)
\]

since \(\phi_{j,k}^2 = 1\). Noticing that \(\sum_{k \in \mathbb{Z}} \phi_{j,k}^2(x) = 2^j\) and that \(\sum_{k \in \mathbb{Z}} \alpha_{j,k}^2 \leq 2^j\),
\[
\sup_x \mathbb{E}_{X_2}(H^2_2(x, X_2)) \leq 2^{j+1} \|f\|_{\infty}.
\]
Therefore
\[ A_3 \leq \sqrt{2^{J+1}n}\|f\|_\infty. \]

**Evaluation of \( A_4 \).**

It is easy to see that
\[ A_4 \leq 2^{J+2}. \]

We now derive from Houdré and Reynaud's inequality that there exists \( C > 0 \) such that for all \( t > 0 \),
\[
\mathbb{P}
\left( \frac{1}{n(n-1)} \sum_{l \neq l'}^{n} H_J(X_l, X_{l'}) \geq C (\frac{2^J \theta_J t + \|f\|_\infty t + \sqrt{2^J \|f\|_\infty t^3/2 \sqrt{n} + 2^{J+2} t^2}}{n}) \right) \leq 5.6 \text{exp}(-t).
\]

Since \( 2ab \leq a^2 + b^2 \),
\[
2\sqrt{2^J \|f\|_\infty t^{3/2} \sqrt{n}} \leq \|f\|_\infty t + \frac{2^J t^2}{n}.
\]

This concludes the proof of the lemma.

**References**


