



**HAL**  
open science

# Wavelet linear estimation of a density and its derivatives from observations of mixtures under quadrant dependence

Christophe Chesneau

► **To cite this version:**

Christophe Chesneau. Wavelet linear estimation of a density and its derivatives from observations of mixtures under quadrant dependence. 2011. hal-00510690v3

**HAL Id: hal-00510690**

**<https://hal.science/hal-00510690v3>**

Preprint submitted on 3 Jul 2011 (v3), last revised 3 Oct 2011 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# WAVELET LINEAR ESTIMATION OF A DENSITY AND ITS DERIVATIVES FROM OBSERVATIONS OF MIXTURES UNDER QUADRANT DEPENDENCE

Christophe Chesneau

Laboratoire de Mathématiques Nicolas Oresme  
Université de Caen Basse-Normandie  
Campus II, Science 3, 14032 Caen, France.

**Abstract.** The estimation of a density and its derivatives from a finite mixture under the pairwise positive quadrant dependence assumption is considered. A new wavelet based linear estimator is constructed. We evaluate its asymptotic performance by determining an upper bound of the mean integrated squared error.

**Mathematics Subject Classification.** 62G07, 62G20.

**Keywords.** density estimation, mixtures, quadrant dependence, wavelets.

## 1 Introduction

The following mixture density model is considered: we observe  $n$  random variables  $X_1, \dots, X_n$  such that, for any  $i \in \{1, \dots, n\}$ , the density of  $X_i$  is the finite mixture:

$$h_i(x) = \sum_{d=1}^m w_d(i) f_d(x), \quad x \in [0, 1],$$

where

- $(w_d(i))_{(i,d) \in \{1, \dots, n\} \times \{1, \dots, m\}}$  are known positive weights such that, for any  $i \in \{1, \dots, n\}$ ,

$$\sum_{d=1}^m w_d(i) = 1,$$

- $f_1, \dots, f_m$  are unknown densities.

For a fixed  $\nu \in \{1, \dots, m\}$ , we aim to estimate  $f_\nu$  and, more generally, the  $r$ -th derivative  $f_\nu^{(r)}$ , from Pairwise Positive Quadrant Dependent (PPQD)  $X_1, \dots, X_n$ .

Let us present a brief survey related to this problem under various configurations. When  $X_1, \dots, X_n$  are independent, the estimation of  $f_\nu$  has been considered in e.g. Maiboroda (1996), Hall and Zhou (2003) and Pokhyl'ko (2005), and the estimation of  $f_\nu^{(r)}$  by Prakasa Rao (2010). On the other hand,

when  $X_1, \dots, X_n$  are identically distributed i.e.  $h = h_1 = \dots = h_n$ , the estimation of  $h$  for associated  $X_1, \dots, X_n$  (including PPQD) has been investigated in e.g. Cai and Roussas (1997), Dewan and Prakasa Rao (1999), Masry (2001), Prakasa Rao (2003) and Chaubey *et al.* (2006). However, to the best of our knowledge, the combination of these two statistical frameworks i.e. the estimation of  $f_\nu^{(r)}$  (including  $f_\nu$ ) under associated dependence conditions is a new challenge.

This estimation problem occurs in the study of medical, biological and other types of data. The most common situation is the following: for any  $i \in \{1, \dots, n\}$ ,  $X_i$  depends on an unobserved random indicator  $I_i$  taking its values in  $\{1, \dots, m\}$ . Applying the Bayes theorem, the density of  $X_i$  is the finite mixture  $h_i$  defined with  $w_d(i) = \mathbb{P}(I_i = d)$  and  $f_d$  is the conditional density of  $X_i$  given  $\{I_i = d\}$ . We refer to Maiboroda (1996) and the references there in. Further details and applications on the concept of associated random variables can be found in Roussas (1999) and Prakasa Rao and Dewan (1999).

To estimate  $f_\nu^{(r)}$ , several methods are possible as kernel, spline, wavelet, ... (see e.g. Prakasa Rao (1983, 1999), Härdle *et al.* (1998) and Tsybakov (2004)). In this study, we focus our attention on the multiresolution analysis techniques by adopting the wavelet methodology of Pokhyl'ko (2005) and Prakasa Rao (2010). We construct a linear wavelet estimator and explore its asymptotic performance by taking the mean integrated squared error (MISE) and assuming that  $f_\nu^{(r)}$  belongs to a Besov ball.

This paper is organized as follows. Assumptions on the model and some notations are introduced in Section 2. Section 3 briefly describes the wavelet basis on  $[0, 1]$  and the Besov balls. The linear wavelet estimator and the results are presented in Section 4. Section 5 is devoted to the proofs.

## 2 Assumptions

Additional assumptions on the model are presented below. The integers  $r$  and  $\nu$  refer to those in  $f_\nu^{(r)}$ .

**Assumption on  $f_1, \dots, f_m$ .** Without loss of generality, for any  $d \in \{1, \dots, m\}$ , we assume that the support of  $f_d$  is  $[0, 1]$  (our study can be extended to another compact support).

We suppose that there exists a constant  $C_* > 0$  such that

$$\sup_{d \in \{1, \dots, m\}} f_d^{(r)}(x) \leq C_*. \quad (2.1)$$

We suppose that, for any  $d \in \{1, \dots, m\}$ ,

$$f_d^{(r)}(0) = f_d^{(r)}(1). \quad (2.2)$$

**Assumptions on  $X_1, \dots, X_n$ .** We suppose that  $X_1, \dots, X_n$  are PPQD i.e. for any  $(i, \ell) \in \{1, \dots, n\}^2$  with  $i \neq \ell$  and any  $(x, y) \in [0, 1]^2$ ,

$$\mathbb{P}(X_i > x, X_\ell > y) \geq \mathbb{P}(X_i > x)\mathbb{P}(X_\ell > y).$$

This weak kind of dependence has been introduced by Lehmann (1966). Examples of PPQD variables can be found in Sancetta (2009).

Moreover, we assume that there exist positive real numbers  $b_0, \dots, b_{n-1}$  such that

- for any  $(i, \ell) \in \{1, \dots, n\}^2$ ,

$$\mathbb{C}_{ov}(X_i, X_\ell) = b_{|i-\ell|}, \quad (2.3)$$

- there exist two constants,  $C \geq 0$  and  $\theta \in [0, 1)$ , satisfying

$$\lim_{n \rightarrow \infty} \frac{1}{n^\theta} \sum_{u=0}^{n-1} b_u = C. \quad (2.4)$$

These two assumptions are not very restrictive. For instance, if there exists  $\gamma \geq 0$  such that, for any  $(i, \ell) \in \{1, \dots, n\}^2$ ,

$$\mathbb{C}_{ov}(X_i, X_\ell) = \frac{1}{1 + |i - \ell|^\gamma},$$

then (2.3) and (2.4) are satisfied with  $b_u = 1/(1 + |u|^\gamma)$  and

$$\theta = \begin{cases} 0 & \text{if } \gamma > 1, \\ 1 - \gamma & \text{if } \gamma \in [0, 1). \end{cases}$$

**Assumption on the weights of the mixture.** We suppose that the matrix

$$\Gamma_n = \left( \frac{1}{n} \sum_{i=1}^n w_k(i) w_\ell(i) \right)_{(k, \ell) \in \{1, \dots, m\}^2}$$

satisfies  $\det(\Gamma_n) > 0$ . For the considered  $\nu$  (the one which refers to the estimation of  $f_\nu^{(r)}$ ) and any  $i \in \{1, \dots, n\}$ , we set

$$a_\nu(i) = \frac{1}{\det(\Gamma_n)} \sum_{k=1}^m (-1)^{k+\nu} \gamma_{\nu, k}^n w_k(i), \quad (2.5)$$

where  $\gamma_{\nu, k}^n$  denotes the determinant of the minor  $(\nu, k)$  of the matrix  $\Gamma_n$ .

Then  $a_\nu(1), \dots, a_\nu(n)$  satisfy

$$(a_\nu(1), \dots, a_\nu(n)) = \underset{(u_1, \dots, u_n) \in \cap_{d=1}^m \mathcal{U}_{\nu, d}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^n u_i^2, \quad (2.6)$$

where

$$\mathcal{U}_{\nu, d} = \left\{ (u_1, \dots, u_n) \in \mathbb{R}^n; \frac{1}{n} \sum_{i=1}^n u_i w_d(i) = \delta_{\nu, d} \right\}$$

and  $\delta_{\nu, d}$  is the Kronecker delta.

Technical details can be found in Maiboroda (1996).

We set

$$z_n = \frac{1}{n} \sum_{i=1}^n a_\nu^2(i). \quad (2.7)$$

For technical reasons, we suppose that  $z_n < n^{1-\theta}$  where  $\theta$  refers to (2.4).

### 3 Wavelets and Besov balls

Let  $N \in \mathbb{N}^*$ , and  $\phi$  and  $\psi$  be the initial wavelet functions of the Daubechies wavelets  $dbN$ . In particular, these functions satisfy  $\operatorname{supp}(\phi) = \operatorname{supp}(\psi) = [1 - N, N]$  and belong to  $\mathcal{C}^{r+1}([1 - N, N])$ . Set

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

Then there exists an integer  $\eta$  satisfying  $2^\eta \geq 2N$  such that, for any  $\ell \geq \eta$ , the collection

$$\mathcal{B} = \{\phi_{\ell,k}(\cdot), k \in \{0, \dots, 2^\ell - 1\}; \psi_{j,k}(\cdot); j \in \mathbb{N} - \{0, \dots, \ell - 1\}, k \in \{0, \dots, 2^j - 1\}\},$$

with an appropriate treatment at the boundaries, is an orthonormal basis of  $\mathbb{L}^2([0, 1])$  (the set of square-integrable functions on  $[0, 1]$ ) and  $(\phi_{j,k})^{(r)}(0) = (\phi_{j,k})^{(r)}(1)$ . Details can be found in Cohen *et al.* (1993).

For any integer  $\ell \geq \eta$ , any  $h \in \mathbb{L}^2([0, 1])$  can be expanded on  $\mathcal{B}$  as

$$h(x) = \sum_{k=0}^{2^\ell - 1} \alpha_{\ell,k} \phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty} \sum_{k=0}^{2^j - 1} \beta_{j,k} \psi_{j,k}(x), \quad x \in [0, 1],$$

where  $\alpha_{j,k}$  and  $\beta_{j,k}$  are the wavelet coefficients of  $h$  defined by

$$\alpha_{j,k} = \int_0^1 h(x) \phi_{j,k}(x) dx, \quad \beta_{j,k} = \int_0^1 h(x) \psi_{j,k}(x) dx. \quad (3.1)$$

A function  $h$  belongs to  $B_{2,\infty}^s(M)$  if and only if there exists a constant  $M^* > 0$  (depending on  $M$ ) such that the associated wavelet coefficients (3.1) satisfy

$$\sup_{j \geq \eta} 2^{2js} \sum_{k \in \Lambda_j} \beta_{j,k}^2 \leq M^*.$$

We refer to Meyer (1990).

## 4 Estimator and results

Assuming that  $f_\nu^{(r)} \in B_{2,\infty}^s(M)$ , we define the linear estimator  $\hat{f}^{(r)}$  by

$$\hat{f}^{(r)}(x) = \sum_{k=0}^{2^{j_0}-1} \hat{\alpha}_{j_0,k}^{(r)} \phi_{j_0,k}(x), \quad x \in [0, 1], \quad (4.1)$$

where

$$\hat{\alpha}_{j_0,k}^{(r)} = \frac{(-1)^r}{n} \sum_{i=1}^n a_\nu(i) (\phi_{j_0,k})^{(r)}(X_i), \quad (4.2)$$

$a_\nu(1), \dots, a_\nu(n)$  are defined by (2.5),  $j_0$  is the integer satisfying

$$\frac{1}{2} \left( \frac{n^{1-\theta}}{z_n} \right)^{1/(2s+2r+4)} < 2^{j_0} \leq \left( \frac{n^{1-\theta}}{z_n} \right)^{1/(2s+2r+4)},$$

$z_n$  is defined by (2.7) and  $\theta$  is the one in (2.4). The definition of  $j_0$ , which takes into account the PPQD case, is chosen to minimize the MISE of  $\hat{f}^{(r)}$ .

**Theorem 4.1 (Upper bound for  $\hat{f}^{(r)}$ )** *Let  $X_1, \dots, X_n$  be  $n$  random variables as described in Section 1 under the assumptions of Section 2. Suppose that  $f_\nu^{(r)} \in B_{2,\infty}^s(M)$  with  $s > 0$ . Let  $\hat{f}^{(r)}$  be (4.1). Then there exists a constant  $C > 0$  such that*

$$\mathbb{E} \left( \int_0^1 (\hat{f}^{(r)}(x) - f_\nu^{(r)}(x))^2 dx \right) \leq C \left( \frac{z_n}{n^{1-\theta}} \right)^{2s/(2s+2r+4)}.$$

The proof of Theorem 4.1 uses a moment inequality on (4.2) and a suitable decomposition of the MISE.

Due to the PPQD case (without other assumptions on  $X_1, \dots, X_n$ ), the rate of convergence of  $\hat{f}^{(r)}$  is naturally greater than the optimal one obtained in the independent case i.e.  $(z_n/n)^{2s/(2s+2r+1)}$  (see (Prakasa Rao, 2010, Theorem 6.1 and Remark 6.1)).

Note that Theorem 4.1 can be extended to other kinds of associated  $X_1, \dots, X_n$  as Negative Associated (NA), Pairwise Negative Quadrant Dependence (PNQD),

... This is due to the Newman inequality (Newman, 1980, Lemma 3) used in the proof of Theorem 4.1 which still holds in these cases.

The following proposition investigates the convergence a.s. of  $\hat{f}^{(r)}$  to  $f_\nu^{(r)}$  under the  $\mathbb{L}^2$ -norm.

**Proposition 4.1 (Convergence a.s.)** *Consider the framework of Theorem 4.1. Suppose that  $\sum_{n=1}^{\infty} (z_n/n^{1-\theta})^{2s/(2s+2r+4)}$  exists. Then*

$$\lim_{n \rightarrow \infty} \int_0^1 (\hat{f}^{(r)}(x) - f_\nu^{(r)}(x))^2 dx = 0 \quad \text{a.s.}$$

Finally, remark that the linear wavelet estimator is not adaptive with respect to  $s$ . Adaptivity can perhaps be achieved by using a non-linear wavelet estimator as the hard thresholding one. This approach works in the independent case (see (Pokhyl'ko, 2005, Theorem 4)), but the proof of this fact uses technical tools as the Bernstein inequality and the Rosenthal inequality. It is not immediately clear how to extend this to the PPQD case.

## 5 Proofs

In this section, we consider the density model described in Section 1 under the assumptions of Section 2. Moreover,  $C$  denotes any constant that does not depend on  $j$ ,  $k$  and  $n$ . Its value may change from one term to another and may depends on  $\phi$ .

**Proposition 5.1** *Let  $X_1, \dots, X_n$  be  $n$  random variables as described in Section 1 under the assumptions of Section 2. For any  $k \in \{0, \dots, 2^{j_0} - 1\}$ , let  $\alpha_{j_0, k}^{(r)} = \int_0^1 f_\nu^{(r)}(x) \phi_{j_0, k}(x) dx$  and  $\hat{\alpha}_{j_0, k}^{(r)}$  be (4.2). Then there exists a constant  $C > 0$  such that*

$$\mathbb{E}((\hat{\alpha}_{j_0, k}^{(r)} - \alpha_{j_0, k}^{(r)})^2) \leq C 2^{(3+2r)j_0} \frac{z_n}{n^{1-\theta}}.$$

**Proof of Proposition 5.1.** Proceeding as in (Prakasa Rao, 2010, eq. (4.6)), it follows from (2.6),  $r$  integrations by parts, (2.2) and  $(\phi_{j_0, k})^{(r)}(0) = (\phi_{j_0, k})^{(r)}(1)$  that

$$\begin{aligned} \mathbb{E}(\hat{\alpha}_{j_0, k}^{(r)}) &= \frac{(-1)^r}{n} \sum_{i=1}^n a_\nu(i) \mathbb{E}((\phi_{j_0, k})^{(r)}(X_i)) \\ &= \frac{(-1)^r}{n} \sum_{i=1}^n a_\nu(i) \int_0^1 (\phi_{j_0, k})^{(r)}(x) h_i(x) dx \\ &= (-1)^r \sum_{d=1}^m \int_0^1 f_d(x) (\phi_{j_0, k})^{(r)}(x) dx \left( \frac{1}{n} \sum_{i=1}^n a_\nu(i) w_d(i) \right) \\ &= (-1)^r \int_0^1 f_\nu(x) (\phi_{j_0, k})^{(r)}(x) dx = \int_0^1 f_\nu^{(r)}(x) \phi_{j_0, k}(x) dx = \alpha_{j_0, k}^{(r)}. \end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{E}((\hat{\alpha}_{j_0,k}^{(r)} - \alpha_{j_0,k}^{(r)})^2) &= \mathbb{V}(\hat{\alpha}_{j_0,k}^{(r)}) \\
&= \frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1}^n a_\nu(i) a_\nu(\ell) \mathbb{C}_{ov}((\phi_{j_0,k}^{(r)}(X_i), (\phi_{j_0,k}^{(r)}(X_\ell))) \\
&\leq \frac{1}{n^2} \sum_{i=1}^n a_\nu^2(i) \mathbb{V}((\phi_{j_0,k}^{(r)}(X_i))) + \\
&\quad \frac{1}{n^2} \sum_{i=1}^n \sum_{\substack{\ell=1 \\ \ell \neq i}}^n |a_\nu(i)| |a_\nu(\ell)| |\mathbb{C}_{ov}((\phi_{j_0,k}^{(r)}(X_i), (\phi_{j_0,k}^{(r)}(X_\ell)))|. \quad (5.1)
\end{aligned}$$

Let us bound the first term in (5.1). For any  $i \in \{1, \dots, n\}$ , using (2.1) and  $(\phi_{j_0,k}^{(r)}(x) = 2^{j_0/2} 2^{rj_0} \phi^{(r)}(2^{j_0}x - k)$ , we have

$$\begin{aligned}
\mathbb{V}((\phi_{j_0,k}^{(r)}(X_i))) &\leq \mathbb{E}(((\phi_{j_0,k}^{(r)}(X_i))^2) = \int_0^1 ((\phi_{j_0,k}^{(r)}(x))^2 h_i(x) dx \\
&\leq C_* 2^{2rj_0} \int_0^1 (\phi^{(r)}(x))^2 dx \leq C 2^{2rj_0}.
\end{aligned}$$

Therefore

$$\frac{1}{n^2} \sum_{i=1}^n a_\nu^2(i) \mathbb{V}((\phi_{j_0,k}^{(r)}(X_i))) \leq C 2^{2rj_0} \frac{1}{n^2} \sum_{i=1}^n a_\nu^2(i) = C 2^{2rj_0} \frac{z_n}{n}. \quad (5.2)$$

Let us now bound the second term in (5.1). Since  $X_1, \dots, X_n$  are PPQD, it follows from (Newman, 1980, Lemma 3) that, for any  $(i, \ell) \in \{1, \dots, n\}^2$  with  $i \neq \ell$ ,

$$|\mathbb{C}_{ov}((\phi_{j_0,k}^{(r)}(X_i), (\phi_{j_0,k}^{(r)}(X_\ell)))| \leq \left( \sup_{x \in [0,1]} |(\phi_{j_0,k}^{(r+1)}(x))| \right)^2 \mathbb{C}_{ov}(X_i, X_\ell) \quad (5.3)$$

Using (5.3) and (2.3), we obtain

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{\substack{\ell=1 \\ \ell \neq i}}^n |a_\nu(i)| |a_\nu(\ell)| |\mathbb{C}_{ov}((\phi_{j_0,k}^{(r)}(X_i), (\phi_{j_0,k}^{(r)}(X_\ell)))| \leq \frac{1}{n^2} AB, \quad (5.4)$$

where

$$A = \left( \sup_{x \in [0,1]} |(\phi_{j_0,k}^{(r+1)}(x))| \right)^2, \quad B = \sum_{i=1}^n \sum_{\substack{\ell=1 \\ \ell \neq i}}^n |a_\nu(i)| |a_\nu(\ell)| b_{|i-\ell|}.$$

Let us now bound  $A$  and  $B$  in turn.



Since  $(\phi_{j_0,k})^{(r+1)}(x) = 2^{j_0/2}2^{(r+1)j_0}\phi^{(r+1)}(2^{j_0}x - k)$ , we have

$$A \leq C(2^{j_0/2}2^{(r+1)j_0})^2 = C2^{j_0(3+2r)}. \quad (5.5)$$

We have

$$\begin{aligned} B &= 2 \sum_{i=2}^n \sum_{\ell=1}^{i-1} |a_\nu(i)| |a_\nu(\ell)| b_{i-\ell} \leq \sum_{i=2}^n \sum_{\ell=1}^{i-1} (a_\nu^2(i) + a_\nu^2(\ell)) b_{i-\ell} \\ &= \sum_{i=2}^n a_\nu^2(i) \sum_{u=1}^{i-1} b_u + \sum_{i=2}^n \sum_{u=1}^{i-1} a_\nu^2(i-u) b_u. \end{aligned}$$

Using (2.4), we obtain

$$\sum_{i=2}^n a_\nu^2(i) \sum_{u=1}^{i-1} b_u \leq n z_n \sum_{u=0}^{n-1} b_u \leq C z_n n^{\theta+1}$$

and

$$\sum_{i=2}^n \sum_{u=1}^{i-1} a_\nu^2(i-u) b_u = \sum_{u=1}^{n-1} b_u \sum_{i=u+1}^n a_\nu^2(i-u) \leq n z_n \sum_{u=0}^{n-1} b_u \leq C z_n n^{\theta+1}.$$

Hence

$$B \leq C z_n n^{\theta+1}. \quad (5.6)$$

Putting (5.1), (5.2), (5.4), (5.5) and (5.6) together, we obtain

$$\mathbb{E}((\hat{\alpha}_{j_0,k}^{(r)} - \alpha_{j_0,k}^{(r)})^2) \leq C \left( 2^{2rj_0} \frac{z_n}{n} + 2^{(3+2r)j_0} \frac{z_n}{n^{1-\theta}} \right) \leq C 2^{(3+2r)j_0} \frac{z_n}{n^{1-\theta}}.$$

This ends the proof of Proposition 5.1. □

**Proof of Theorem 4.1.** We expand the function  $f_\nu^{(r)}$  on  $\mathcal{B}$  as

$$f_\nu^{(r)}(x) = \sum_{k=0}^{2^{j_0}-1} \alpha_{j_0,k}^{(r)} \phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^{(r)} \psi_{j,k}(x), \quad x \in [0, 1],$$

where

$$\alpha_{j_0,k}^{(r)} = \int_0^1 f_\nu^{(r)}(x) \phi_{j_0,k}(x) dx, \quad \beta_{j,k}^{(r)} = \int_0^1 f_\nu^{(r)}(x) \psi_{j,k}(x) dx.$$

We have, for any  $x \in [0, 1]$ ,

$$\hat{f}^{(r)}(x) - f_\nu^{(r)}(x) = \sum_{k=0}^{2^{j_0}-1} (\hat{\alpha}_{j_0,k}^{(r)} - \alpha_{j_0,k}^{(r)}) \phi_{j_0,k}(x) - \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^{(r)} \psi_{j,k}(x).$$

Since  $\mathcal{B}$  is an orthonormal basis of  $\mathbb{L}^2([0, 1])$ , we have

$$\mathbb{E} \left( \int_0^1 (\hat{f}^{(r)}(x) - f_\nu^{(r)}(x))^2 dx \right) = A + B,$$

where

$$A = \sum_{k=0}^{2^{j_0}-1} \mathbb{E}((\hat{\alpha}_{j_0,k}^{(r)} - \alpha_{j_0,k}^{(r)})^2), \quad B = \sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} (\beta_{j,k}^{(r)})^2.$$

Using Proposition 5.1 and the definition of  $j_0$ , we obtain

$$A \leq C 2^{j_0} 2^{(3+2r)j_0} \frac{z_n}{n^{1-\theta}} \leq C \left( \frac{z_n}{n^{1-\theta}} \right)^{2s/(2s+2r+4)}.$$

Since  $f_\nu^{(r)} \in B_{2,\infty}^s(M)$ , we have

$$B \leq C 2^{-2j_0 s} \leq C \left( \frac{z_n}{n^{1-\theta}} \right)^{2s/(2s+2r+4)}.$$

Therefore

$$\mathbb{E} \left( \int_0^1 (\hat{f}^{(r)}(x) - f_\nu^{(r)}(x))^2 dx \right) \leq C \left( \frac{z_n}{n^{1-\theta}} \right)^{2s/(2s+2r+4)}.$$

The proof of Theorem 4.1 is complete. □

**Proof of Proposition 4.1.** It follows from the Markov inequality and Theorem 4.1 that, for any  $\epsilon > 0$ ,

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P} \left( \int_0^1 (\hat{f}^{(r)}(x) - f_\nu^{(r)}(x))^2 dx \geq \epsilon \right) &\leq \frac{1}{\epsilon} \sum_{n=1}^{\infty} \mathbb{E} \left( \int_0^1 (\hat{f}^{(r)}(x) - f_\nu^{(r)}(x))^2 dx \right) \\ &\leq \frac{C}{\epsilon} \sum_{n=1}^{\infty} \left( \frac{z_n}{n^{1-\theta}} \right)^{2s/(2s+2r+4)} < \infty \end{aligned}$$

The Borel-Cantelli theorem yields

$$\lim_{n \rightarrow \infty} \int_0^1 (\hat{f}^{(r)}(x) - f_\nu^{(r)}(x))^2 dx = 0 \quad a.s.$$

The proof of Proposition 4.1 is complete. □

## Acknowledgment.

This work is supported by ANR grant NatImages, ANR-08-EMER-009.

## References

- Cai, Z.W., and Roussas, G.G. (1997). Efficient estimation of a distribution function under quadrant dependence, *Scandinavian Journal of Statistics*, 24, 1-14.
- Chaubey, Y.P., Doosti, H. and Prakasa Rao, B.L.S. (2006). Wavelet based estimation of the derivatives of a density with associated variables, *International Journal of Pure and Applied Mathematics*, 27, 1, 97-106.
- Cohen, A., Daubechies, I., Jawerth, B., and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms, *Applied and Computational Harmonic Analysis*, 24, 1, 54-81.
- Dewan, I., and Prakasa Rao, B.L.S. (1999). A general method of density estimation for associated random variables, *Journal of Nonparametric Statistics*, 10, 405-420.
- Hall, P., and Zhou, X.H. (2003). Nonparametric estimation of component distributions in a multivariate mixture, *The Annals of Statistics*, 31, 1, 201-224.
- Härdle, W., Kerkyacharian, G., Picard, D., and Tsybakov, A.B. (1998). *Wavelet, Approximation and Statistical Applications*. Lectures Notes in Statistics, New York, 129, Springer Verlag.
- Lehmann, E.L. (1996). Some concepts of dependence, *The Annals of Mathematical Statistics*, 37, 1137-1153.
- Maiboroda, R.E. (1996). Estimators of components of mixtures with varying concentrations, *Ukrain. Mat. Zh.*, 48, 4, 618-622.
- Masry, E. (2001). Multivariate probability density estimation for associated processes: Strong consistency and rates, *Statistics and Probability Letters*, 58, 205-219.
- Meyer, Y. (1990). *Ondelettes et Opérateurs*, Hermann, Paris.
- Newman, C.M. (1980). Normal fluctuations and the FKG inequalities, *Commun. Math. Phys.*, 74, 119-128.
- Pokhyl'ko, D. (2005). Wavelet estimators of a density constructed from observations of a mixture, *Theor. Prob. and Math. Statist.*, 70, 135-145.
- Prakasa Rao, B.L.S. (1983). *Nonparametric functional estimation*, Academic Press, Orlando.

- Prakasa Rao, B.L.S. (1999). *Nonparametric functional estimation: an overview*, Asymptotics, Nonparametrics and Time Series, Ed. Subir Ghosh, 461-509, Marcel Dekker Inc. New York.
- Prakasa Rao, B.L.S. (2003). Wavelet linear density estimation for associated sequences, *Journal of the Indian Statistical Association*, 41, 369-379.
- Prakasa Rao, B.L.S. (2010). Wavelet linear estimation for derivatives of a density from observations of mixtures with varying mixing proportions, *Indian Journal of Pure and Applied Mathematics*, 41, 1, 275-291.
- Prakasa Rao, B.L.S., and Dewan, I. (2001). *Associated sequences and related inference problems*, Handbook of Statistics, 19 , Stochastic Processes: Theory and Methods, (eds. C.R. Rao and D.N. Shanbag), 693-728, North Holland, Amsterdam.
- Roussas, G.G. (1999). *Positive and negative dependence with some statistical applications*, Asymptotics, nonparametrics and time series (ed. S. Ghosh), 757-788, Marcel Dekker, New York.
- Sancetta, A. (2009). Strong law of large numbers for pairwise positive quadrant dependent random variables, *Statistical Inference for Stochastic Processes*, 12, 1, 55-64.
- Tsybakov, A.B. (2004). *Introduction à l'estimation non-paramétrique*, Springer.