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

Christophe Chesneau

Laboratoire de Mathématiques Nicolas Oresme (LMNO), 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. We prove that it attains a sharp rate of convergence for a wide class of unknown densities.

1. Introduction

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

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

where

- m is a positive integer,

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

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

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

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

Let us now present a brief survey related to this problem under various configurations. When $X_1, \ldots, X_n$ are independent, the estimation of f_v has been considered in e.g. [9], [5] and [14]. The estimation of its r-th derivative $f_v^{(r)}$ has been recently studied by [18] (this is particularly of interest to detect possible bumps, concavity or convexity properties of f_v). When $X_1, \ldots, X_n$ are identically distributed i.e. $h = h_1 = \ldots = h_n$,

the estimation of h for associated $X_1, \ldots, X_n$ (including PPQD) has been investigated in e.g. [1], [4], [11] and [17]. The estimation of $h^{(r)}$ has been explored by [2]. However, to the best of our knowledge, the combination of these two complex statistical frameworks i.e. the estimation of $f_v^{(r)}$, including f_v, under PPQD conditions is a new challenge.

Such a problem occurs in the study of medical, biological and other types of data. The most common situation is the following: for any $i \in \{1, \ldots, n\}$, X_i depends on an unobserved random indicator I_i taking its values in $\{1, \ldots, m\}$. Applying the Bayes theorem, the density of X_i is h_i defined with $w_d(i) = \mathbb{P}(I_i = d)$ and f_d the conditional density of X_i given $\{I_i = d\}$. Naturally, in some situations, $X_1, \ldots, X_n$ are not independent and this motivates the study of various dependence structures as the PPQD one. Further details and applications on the concept of PPQD can be found in [20], [19] and [21].

To estimate $f_v^{(r)}$, several methods are possible as kernel, splines, …(see e.g. [15, 16], [6] and [22]). In this study, we focus our attention on the multiresolution analysis techniques and, more precisely, the wavelet methodology of [14] and [18]. We construct a linear wavelet estimator and explore its asymptotic performance by taking the mean integrated squared error (MISE) and assuming that $f_v^{(r)}$ belongs to a Besov ball. We prove that, under some specific assumptions, it attains a similar rate of convergence to the one obtained in the independent case.

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 v refer to those in $f_v^{(r)}$.

Assumption on $f_1, \ldots, f_m$. Without loss of generality, for any $d \in \{1, \ldots, m\}$, 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, for any $d \in \{1, \ldots, m\}$,

$$\sup_{x \in [0,1]} |f_d^{(r)}(x)| \le C_*. \tag{2.1}$$

We suppose that, for any $d \in \{1, \ldots, m\}$ and $v \in \{0, \ldots, r\}$,

$$f_d^{(v)}(0) = f_d^{(v)}(1) = 0. \tag{2.2}$$

Simple examples are densities of the form $f_d(x) = c^{-1} x^\alpha (1-x)^\beta g_d(x)$, $x \in [0,1]$, where $\alpha > r$, $\beta > r$, g_d is a positive function such that $g_d \in C^r([0,1])$ and $c = \int_0^1 x^\alpha (1-x)^\beta g_d(x) dx$. This includes some usual distributions as $Beta(\alpha, \beta)$, $Beta.mixture(\alpha, \beta)$, ….

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

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

satisfies $\det(\Gamma_n) > 0$. For the considered v and any $i \in \{1, \ldots, n\}$, we set

$$a_v(i) = \frac{1}{\det(\Gamma_n)} \sum_{k=1}^m (-1)^{k+v} \gamma_{v,k}^n w_k(i), \tag{2.3}$$

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

Then $a_v(1), \ldots, a_v(n)$ satisfy

$$(a_v(1), \ldots, a_v(n)) = \underset{(u_1,\ldots,u_n)\in\cap_{d=1}^{m}\mathcal{U}_{v,d}}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n} u_i^2, \tag{2.4}$$

where

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

and $\delta_{v,d}$ denotes the Kronecker delta.

Technical details can be found in [9] and [18].

We set

$$z_n = \frac{1}{n}\sum_{i=1}^{n} a_v^2(i). \tag{2.5}$$

For technical reasons, we suppose that $(n/z_n)^{1/(2r+1)} \in (1, n)$.

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

$$\mathbb{P}(X_i > x, X_\ell > y) \geq \mathbb{P}(X_i > x)\mathbb{P}(X_\ell > y). \tag{2.6}$$

This kind of dependence has been introduced by [8]. Examples of PPQD variables can be found in [21].

We suppose that, for any $(i, \ell) \in \{1, \ldots, n\}^2$, there exists a constant $C > 0$ such that

$$\sup_{(x,y)\in[0,1]^2} |h_{i,\ell}(x, y) - h_i(x)h_\ell(y)| \leq C, \tag{2.7}$$

where $h_{i,\ell}$ denotes the density of (X_i, X_ℓ).

We suppose that there exists a constant $C > 0$ such that

$$\sum_{i=2}^{n} i^3 \sum_{\ell=1}^{i-1} |a_v(i)||a_v(\ell)|\mathbb{C}_{ov}(X_i, X_\ell) \leq Cnz_n, \tag{2.8}$$

where $\mathbb{C}_{ov}(.,.)$ denotes the covariance, $a_v(1), \ldots, a_v(n)$ are (2.3) and z_n is (2.5).

In particular, if there exist a constant $C > 0$ and a sequence of positive real numbers $(b_n)_{n\in\mathbb{N}}$ satisfying

- $\sum_{n=1}^{\infty} b_n < \infty$,
- for any $(i, \ell) \in \{1, \ldots, n\}^2$ with $i \neq \ell$, $i^3\mathbb{C}_{ov}(X_i, X_\ell) \leq Cb_{|i-\ell|}$,

then (2.8) is satisfied.

Note that (2.6), (2.7) and (2.8) include the independent case.

3. Wavelets and Besov balls

Throughout the paper, we work with the wavelet basis on $[0, 1]$ described below. Let ϕ and ψ be the initial wavelet functions of the Daubechies wavelets $db2N$ with N such that ϕ and ψ belong to C^{r+1}. Furthermore, mention that ϕ and ψ are compactly supported. Set

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

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

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

with an appropriate treatment at the boundaries, is an orthonormal basis of $\mathbb{L}^2([0, 1])$ (the set of square-integrable functions on $[0, 1]$).

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

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

where

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

Details can be found in [3] and [10].

In this paper we assume that $f_\nu^{(r)}$ belongs to a subset of Besov space defined below. We say that a function $h \in \mathbb{L}^2([0, 1])$ belongs to $B_{2,\infty}^s(M)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that (3.1) satisfy

$$\sup_{j \geq \eta} 2^{2js} \sum_{k=0}^{2^j - 1} \beta_{j,k}^2 \leq M^*.$$

For more details about wavelet basis, see [12] and [10].

The next section is devoted to our estimator and its asymptotic performances in term of MISE.

4. Estimator and results

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

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

where

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

$a_\nu(1), \dots, a_\nu(n)$ are (2.3), j_0 is the integer satisfying

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

and z_n is defined by (2.5).

The definitions of $\hat{\alpha}_{j_0,k}^{(r)}$ and j_0, which take into account the PPQD case, are chosen to minimize the MISE of $\hat{f}_\nu^{(r)}$.

Note that $\hat{f}_\nu^{(r)}$ is close to the estimator considered by [18, equation (4.5)] in the independent case. Further details on derivatives density estimation via wavelet methods can also be found in [2] and [7].

Theorem 4.1 below investigates the MISE of $\hat{f}_\nu^{(r)}$ when $f_\nu^{(r)} \in B_{2,\infty}^s(M)$.

Theorem 4.1. *Let $X_1, \ldots, X_n$ 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 $\hat{f}_\nu^{(r)}$ be (4.1). Then there exists a constant $C > 0$ such that*

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

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

The obtained rate of convergence is the one related to the independent case i.e. $(z_n/n)^{2s/(2s+2r+1)}$ (see [18, Theorem 6.1 and Remark 6.1]).

Note that Theorem 4.1 can be extended to Pairwise Negative Quadrant Dependence $X_1, \ldots, X_n$ (this is due to the Newman inequality [13, Lemma 3] used in the proof of Theorem 4.1 which still holds in this case).

Remark that $\hat{f}_\nu^{(r)}$ 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 [14, Theorem 4]). However, the proof of this fact uses technical tools as the Bernstein and the Rosenthal inequalities and it is not immediately clear how to extend this to the PPQD case.

5. Proofs

In this section, 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 ϕ.

Proposition 5.1. *Let $X_1, \ldots, X_n$ be n random variables as described in Section 1 under the assumptions of Section 2. For any $k \in \{0, \ldots, 2^{j_0} - 1\}$, 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*

$$\mathbb{E}((\hat{\alpha}_{j_0,k}^{(r)} - \alpha_{j_0,k}^{(r)})^2) \le C 2^{2rj_0} \frac{z_n}{n}.$$

Proof of Proposition 5.1. Proceeding as in [18, equation (4.6)], it follows from (2.4) and r integrations by parts with (2.2) 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

$$\mathbb{E}((\hat{\alpha}_{j_0,k}^{(r)} - \alpha_{j_0,k}^{(r)})^2) = \mathbb{V}(\hat{\alpha}_{j_0,k}^{(r)})$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{\ell=1}^{n} a_v(i) a_v(\ell) \mathbb{C}_{ov}((\phi_{j_0,k})^{(r)}(X_i), (\phi_{j_0,k})^{(r)}(X_\ell))$$

$$\leq \frac{1}{n^2} \sum_{i=1}^{n} a_v^2(i) \mathbb{V}((\phi_{j_0,k})^{(r)}(X_i)) + \frac{2}{n^2} \sum_{i=2}^{n} \sum_{\ell=1}^{i-1} |a_v(i)||a_v(\ell)||\mathbb{C}_{ov}((\phi_{j_0,k})^{(r)}(X_i), (\phi_{j_0,k})^{(r)}(X_\ell))|. \qquad (5.1)$$

Let us bound the first term in (5.1). For any $i \in \{1, \dots, n\}$, using (2.1) which implies $\sup_{x \in [0,1]} h_i(x) \leq C_*$, the equality $(\phi_{j_0,k})^{(r)}(x) = 2^{j_0/2} 2^{r j_0} \phi^{(r)}(2^{j_0} x - k)$ and making the change of variable $y = 2^{j_0} x - k$, we have

$$\mathbb{V}((\phi_{j_0,k})^{(r)}(X_i)) \leq \mathbb{E}(((\phi_{j_0,k})^{(r)}(X_i))^2) = \int_0^1 ((\phi_{j_0,k})^{(r)}(x))^2 h_i(x) dx$$

$$\leq C_* 2^{2r j_0} \int (\phi^{(r)}(y))^2 dy \leq C 2^{2r j_0}.$$

Therefore

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

Let us now investigate the bound of the covariance term in (5.1) via two different approaches.

Bound 1. By a standard covariance equality and (2.7), for any $(i, \ell) \in \{1, \dots, n\}^2$ with $i \neq \ell$, we have

$$|\mathbb{C}_{ov}((\phi_{j_0,k})^{(r)}(X_i), (\phi_{j_0,k})^{(r)}(X_\ell))|$$

$$= \left| \int_0^1 \int_0^1 (h_{i,\ell}(x,y) - h_i(x)h_\ell(y))(\phi_{j_0,k})^{(r)}(x)(\phi_{j_0,k})^{(r)}(y) dx dy \right|$$

$$\leq \int_0^1 \int_0^1 |h_{i,\ell}(x,y) - h_i(x)h_\ell(y)||(\phi_{j_0,k})^{(r)}(x)||(\phi_{j_0,k})^{(r)}(y)| dx dy$$

$$\leq C \left(\int_0^1 |(\phi_{j_0,k})^{(r)}(x)| dx \right)^2.$$

Moreover, since $(\phi_{j_0,k})^{(r)}(x) = 2^{j_0/2} 2^{r j_0} \phi^{(r)}(2^{j_0} x - k)$, by the change of variables $y = 2^{j_0} x - k$, we obtain

$$\int_0^1 |(\phi_{j_0,k})^{(r)}(x)| dx = 2^{r j_0} 2^{-j_0/2} \int |\phi^{(r)}(y)| dy.$$

Therefore

$$|\mathbb{C}_{ov}((\phi_{j_0,k})^{(r)}(X_i), (\phi_{j_0,k})^{(r)}(X_\ell))| \leq C 2^{2r j_0} 2^{-j_0}. \qquad (5.3)$$

Bound 2. Since $X_1, \dots, X_n$ are PPQD, it follows from [13, Lemma 3] that, for any $(i, \ell) \in \{1, \dots, n\}^2$ with $i \neq \ell$,

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

Since $(\phi_{j_0,k})^{(r+1)}(x) = 2^{(2r+3)j_0/2}\phi^{(r+1)}(2^{j_0}x - k)$ and $\sup_{x\in[0,1]}|\phi^{(r+1)}(x)| \leq C$, we have

$$\left(\sup_{x\in[0,1]}|(\phi_{j_0,k})^{(r+1)}(x)|\right)^2 \leq C2^{j_0(2r+3)}.$$

Therefore

$$|\mathbb{C}_{ov}((\phi_{j_0,k})^{(r)}(X_i), (\phi_{j_0,k})^{(r)}(X_\ell))| \leq C2^{j_0(2r+3)}\mathbb{C}_{ov}(X_i, X_\ell). \tag{5.4}$$

Combining (5.3) and (5.4), for any $(i, \ell) \in \{1, \dots, n\}^2$ with $i \neq \ell$, we obtain

$$|\mathbb{C}_{ov}((\phi_{j_0,k})^{(r)}(X_i), (\phi_{j_0,k})^{(r)}(X_\ell))| \leq C\min(2^{j_0(2r+3)}\mathbb{C}_{ov}(X_i, X_\ell), 2^{2rj_0}2^{-j_0}). \tag{5.5}$$

It follows from (5.5) and $2^{j_0} < n$ that the second term in (5.1) can be bound as

$$\frac{2}{n^2}\sum_{i=2}^{n}\sum_{\ell=1}^{i-1}|a_\nu(i)||a_\nu(\ell)||\mathbb{C}_{ov}((\phi_{j_0,k})^{(r)}(X_i), (\phi_{j_0,k})^{(r)}(X_\ell))| \leq C(E + F), \tag{5.6}$$

where

$$E = \frac{1}{n^2}2^{2rj_0}2^{-j_0}\sum_{i=2}^{2^{j_0}-1}\sum_{\ell=1}^{i-1}|a_\nu(i)||a_\nu(\ell)|$$

and

$$F = \frac{1}{n^2}2^{j_0(2r+3)}\sum_{i=2^{j_0}}^{n}\sum_{\ell=1}^{i-1}|a_\nu(i)||a_\nu(\ell)|\mathbb{C}_{ov}(X_i, X_\ell).$$

Using $|xy| \leq (1/2)(x^2 + y^2)$, $(x, y) \in \mathbb{R}^2$, we have

$$
\begin{aligned}
E &\leq C\frac{1}{n^2}2^{2rj_0}2^{-j_0}\sum_{i=2}^{2^{j_0}-1}\sum_{\ell=1}^{i-1}\left(a_\nu^2(i) + a_\nu^2(\ell)\right) \\
&\leq C\frac{1}{n^2}2^{2rj_0}2^{-j_0}2^{j_0}\sum_{i=1}^{n}a_\nu^2(i) = C2^{2rj_0}\frac{z_n}{n}.
\end{aligned}
\tag{5.7}
$$

Using (2.8), it comes

$$F \leq \frac{1}{n^2}2^{2rj_0}\sum_{i=2}^{n}i^3\sum_{\ell=1}^{i-1}|a_\nu(i)||a_\nu(\ell)|\mathbb{C}_{ov}(X_i, X_\ell) \leq C2^{2rj_0}\frac{z_n}{n}. \tag{5.8}$$

Putting (5.1), (5.2), (5.6), (5.7) and (5.8) together, we obtain

$$\mathbb{E}((\hat{\alpha}_{j_0,k}^{(r)} - \alpha_{j_0,k}^{(r)})^2) \leq C2^{2rj_0}\frac{z_n}{n}.$$

This ends the proof of Proposition 5.1. □

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

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

where

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

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

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

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

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

where

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

Using Proposition 5.1 and the definition of j_0, we obtain

$$A \le C 2^{j_0} 2^{2rj_0} \frac{z_n}{n} \le C\left(\frac{z_n}{n}\right)^{2s/(2s+2r+1)}.$$

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

$$B \le C \sum_{j=j_0}^{\infty} 2^{-2js} \le C 2^{-2j_0 s} \le C\left(\frac{z_n}{n}\right)^{2s/(2s+2r+1)}.$$

Therefore

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

The proof of Theorem 4.1 is complete. □

References

[1] Cai, Z.W., Roussas, G.G. (1997) Efficient estimation of a distribution function under quadrant dependence, Scandinavian Journal of Statistics 24, 1–14.

[2] Chaubey, Y.P., Doosti, H., 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.

[3] Cohen, A., Daubechies, I., Jawerth, B., Vial, P. (1993) Wavelets on the interval and fast wavelet transforms, Applied and Computational Harmonic Analysis 24(1), 54-81.

[4] Dewan, I., Prakasa Rao, B.L.S. (1999) A general method of density estimation for associated random variables, Journal of Nonparametric Statistics 10, 405-420.

[5] Hall, P., Zhou, X.H. (2003) Nonparametric estimation of component distributions in a multivariate mixture, The Annals of Statistics 31, 201-224.

[6] Härdle, W., Kerkyacharian, G., Picard, D., Tsybakov, A.B. (1998) Wavelet, Approximation and Statistical Applications, Lectures Notes in Statistics, New York, 129, Springer Verlag.

[7] Hosseinioun, N., Doosti, H., Nirumand, H.A. (2010) Nonparametric estimation of the derivatives of a density by the method of wavelet for mixing sequences, Statistical papers, to appear.

[8] Lehmann, E.L. (1996) Some concepts of dependence, The Annals of Mathematical Statistics 37, 1137-1153.

[9] Maiboroda, R.E. (1996) Estimators of components of mixtures with varying concentrations, Ukrain. Mat. Zh. 48(4) 618-622.

[10] Mallat, S. (2009) A wavelet tour of signal processing, 3rd edition, Elsevier/Academic Press, Amsterdam.

[11] Masry, E. (2001) Multivariate probability density estimation for associated processes: Strong consistency and rates, Statistics and Probability Letters 58, 205-219.

[12] Meyer, Y. (1990) Ondelettes et Opérateurs, Hermann, Paris.

[13] Newman, C.M. (1980) Normal fluctuations and the FKG inequalities, Commun. Math. Phys. 74, 119-128.

[14] Pokhyl'ko, D. (2005) Wavelet estimators of a density constructed from observations of a mixture, Theor. Prob. and Math. Statist. 70, 135-145.

[15] Prakasa Rao, B.L.S. (1983) Nonparametric functional estimation, Academic Press, Orlando.

[16] 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.

[17] Prakasa Rao, B.L.S. (2003) Wavelet linear density estimation for associated sequences, Journal of the Indian Statistical Association 41, 369-379.

[18] 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.

[19] Prakasa Rao, B.L.S., 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.

[20] 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.

[21] Sancetta, A. (2009) Strong law of large numbers for pairwise positive quadrant dependent random variables, Statistical Inference for Stochastic Processes 12(1) 55-64.

[22] Tsybakov, A.B. (2004) Introduction à l'estimation non-paramétrique, Springer.