



HAL
open science

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

Christophe Chesneau

► **To cite this version:**

Christophe Chesneau. Wavelet linear estimation of a density from observations of mixtures under quadrant dependence. 2010. hal-00510690v1

HAL Id: hal-00510690

<https://hal.science/hal-00510690v1>

Preprint submitted on 20 Aug 2010 (v1), 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 from observations of mixtures under quadrant dependence

Christophe Chesneau

Received:

Abstract We construct a wavelet based linear estimator for the component of a finite mixture under pairwise positive quadrant dependence. We evaluate its performance by determining an upper bound of the mean integrated squared error.

Keywords density estimation · mixtures · quadrant dependence · Besov balls · wavelets.

2000 Mathematics Subject Classification 62G07, 62G20.

1 Motivations

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 $m \in \mathbb{N}^*$,

- $(w_d(i))_{(i,d) \in \{1, \dots, n\} \times \{1, \dots, m\}}$ 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.

We assume that X_1, \dots, X_n are pairwise positive quadrant dependent (PPQD) (to be defined in Section 2). For a fixed $\nu \in \{1, \dots, m\}$, we aim to estimate f_ν from X_1, \dots, X_n .

The estimation of f_ν when X_1, \dots, X_n are independent has been considered in, e.g., Maiboroda (1996), Hall and Zhou (2003), Pokhyl'ko (2005)

and Prakasa Rao (2010). When X_1, \dots, X_n are identically distributed i.e. $h = h_1 = \dots = h_n$, the estimation of h in the PPQD case (and, a fortiori, the positively associated case) 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 estimation of f_ν under dependence conditions (as PPQD) is a new challenge.

To estimate f_ν , several methods are possible (kernel, spline, wavelet, ...) (see e.g. Prakasa Rao (1983, 1999), Härdle et al (1998) and Tsybakov (2004)). In this study, we adopt the wavelet methodology of Pokhyl'ko (2005) and Prakasa Rao (2010). We construct a linear wavelet estimator. We evaluate its performance by taking the mean integrated squared error (MISE) and assuming that f_ν 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 is presented in Section 4. The upper bound theorem is set in Section 5. Section 6 is devoted to the proofs.

2 Assumptions

Additional assumptions on the model are presented below.

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).

Assumptions on X_1, \dots, X_n . Recall 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 kind of dependence was introduced by Lehmann (1966).

We assume that there exist positive real numbers b_0, \dots, b_{n-1} satisfying

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

$$\mathbb{C}(X_i, X_\ell) = b_{|i-\ell|}, \tag{1}$$

2. there exist two constants $C > 0$ and $\theta \in [0, 1)$ such that

$$\sum_{u=0}^{n-1} b_u \leq Cn^\theta. \tag{2}$$

For instance, if there exists $\gamma \geq 0$ such that, for any $(i, \ell) \in \{1, \dots, n\}^2$, $\mathbb{C}(X_i, X_\ell) = 1/(1 + |i - \ell|^\gamma)$, then (2) is satisfied with $\theta = 0$ if $\gamma > 1$ and $\theta = 1 - \gamma$ if $\gamma \in [0, 1)$.

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

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

is nonsingular i.e. $\det(\Gamma_n) > 0$. For the considered ν (the one which refers to the estimation of f_ν) and any $i \in \{1, \dots, n\}$, we set

$$a_\nu(i) = \frac{1}{\det(\Gamma_n)} \sum_{k=1}^m (-1)^{k+\nu} \gamma_{\nu,k}^n w_k(i), \quad (3)$$

where $\gamma_{\nu,k}^n$ denotes the determinant of the minor (ν, k) of the matrix Γ_n . Then, for any $d \in \{1, \dots, m\}$,

$$\frac{1}{n} \sum_{i=1}^n a_\nu(i) w_d(i) = \begin{cases} 1 & \text{if } d = \nu, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

and

$$(a_\nu(1), \dots, a_\nu(n)) = \operatorname{argmin}_{(b_1, \dots, b_n) \in \mathbb{R}^n} \frac{1}{n} \sum_{i=1}^n b_i^2.$$

Technical details can be found in Maiboroda (1996).

We set

$$z_n = \frac{1}{n} \sum_{i=1}^n a_\nu^2(i). \quad (5)$$

For technical reasons, we suppose that $z_n < n^{1-\theta}$ where θ refers to (2).

3 Wavelets and Besov balls

Wavelet basis. Let $N \in \mathbb{N}^*$ and (ϕ, ψ) be the initial wavelet functions of the Daubechies wavelets dbN . 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 τ satisfying $2^\tau \geq 2N$ such that the collection

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

(with an appropriate treatments at the boundaries) is an orthonormal basis of $\mathbb{L}^2([0, 1])$, the set of square-integrable functions on $[0, 1]$. We refer to Cohen (1993).

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

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

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

$$\alpha_{j,k} = \int_0^1 h(x) \phi_{j,k}(x) dx, \quad \beta_{j,k} = \int_0^1 h(x) \psi_{j,k}(x) dx. \quad (6)$$

Besov balls. Let $M > 0$, $s > 0$, $p \geq 1$ and $r \geq 1$. A function h belongs to $B_{p,r}^s(M)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients (6) satisfy

$$\left(\sum_{j=\tau-1}^{\infty} \left(2^{j(s+1/2-1/p)} \left(\sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \right)^{1/p} \right)^r \right)^{1/r} \leq M^*.$$

We set $\beta_{\tau-1,k} = \alpha_{\tau,k}$. In this expression, s is a smoothness parameter and p and r are norm parameters. For a particular choice of s , p and r , $B_{p,r}^s(M)$ contain the Hölder and Sobolev balls. See Meyer (1990).

4 Estimator

Assuming that $f_\nu \in B_{p,r}^s(M)$ with $p \geq 2$, we define the linear estimator \widehat{f}^L by

$$\widehat{f}^L(x) = \sum_{k=0}^{2^{j_0}-1} \widehat{\alpha}_{j_0,k} \phi_{j_0,k}(x), \quad (7)$$

where

$$\widehat{\alpha}_{j_0,k} = \frac{1}{n} \sum_{i=1}^n a_\nu(i) \phi_{j_0,k}(X_i), \quad (8)$$

$a_\nu(1), \dots, a_\nu(n)$ are defined by (3), j_0 is the integer satisfying

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

z_n is defined by (5) and θ is the one in (2).

The definition of j_0 , which takes into account the PPQD case, is chosen to minimize the MISE of \widehat{f}^L .

5 Result

Upper bound for \widehat{f}^L is given in Theorem 1 below.

Theorem 1 *Let X_1, \dots, X_n be n random variables as described in Section 1 under the assumptions of Section 2. Suppose that $f_\nu \in B_{p,r}^s(M)$ with $s > 0$, $p \geq 2$ and $r \geq 1$. Let \widehat{f}^L be (7). Then there exists a constant $C > 0$ such that*

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

The proof of Theorem 1 uses moment inequalities on (8) and a suitable decomposition of the MISE.

Due to the PPQD case (without other assumptions on the dependence of X_1, \dots, X_n), the rate of convergence of \widehat{f}^L is naturally greater than the optimal one obtained in the independent case i.e. $(z_n/n)^{2s/(2s+1)}$ (see (Pokhyl'ko 2005, Theorem 1)).

Note 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's inequality and Rosenthal's inequality. It is not immediately clear how to extend this to the PPQD case.

6 Proofs

In this section, C represents a positive constant which may differ from one term to another.

Proposition 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 $\alpha_{j_0,k}$ be the wavelet coefficient (6) of f_ν and $\widehat{\alpha}_{j_0,k}$ be (8). Then there exists a constant $C > 0$ such that*

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

Proof of Proposition 1. It follows from (4) that

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

Therefore

$$\begin{aligned}
\mathbb{E} \left((\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k})^2 \right) &= \mathbb{V}(\widehat{\alpha}_{j_0,k}) \\
&= \frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1}^n a_\nu(i) a_\nu(\ell) \mathbb{C}(\phi_{j_0,k}(X_i), \phi_{j_0,k}(X_\ell)) \\
&\leq \frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1}^n |a_\nu(i)| |a_\nu(\ell)| |\mathbb{C}(\phi_{j_0,k}(X_i), \phi_{j_0,k}(X_\ell))|. \tag{9}
\end{aligned}$$

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 $i \neq \ell$,

$$|\mathbb{C}(\phi_{j_0,k}(X_i), \phi_{j_0,k}(X_\ell))| \leq \left(\sup_{x \in [0,1]} |(\phi_{j_0,k}(x))'| \right)^2 \mathbb{C}(X_i, X_\ell). \tag{10}$$

Using (10) and (1), we obtain

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{\ell=1}^n |a_\nu(i)| |a_\nu(\ell)| |\mathbb{C}(\phi_{j_0,k}(X_i), \phi_{j_0,k}(X_\ell))| \leq \frac{1}{n^2} AB, \tag{11}$$

where

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

Let us now bound A and B in turn.

Upper bound for A . Since $\phi \in \mathcal{C}^1([0, 1])$, we have $(\phi_{j_0,k}(x))' = 2^{3j_0/2} \phi'(2^{j_0}x - k)$, so

$$\sup_{x \in [0,1]} |(\phi_{j_0,k}(x))'| \leq 2^{3j_0/2} \sup_{x \in [0,1]} |\phi'(x)| = C 2^{3j_0/2}.$$

Hence

$$A \leq C 2^{3j_0}. \tag{12}$$

Upper bound for B . We have

$$\begin{aligned}
B &= b_0 n z_n + 2 \sum_{i=2}^n \sum_{\ell=1}^{i-1} |a_\nu(i)| |a_\nu(\ell)| b_{i-\ell} \\
&\leq b_0 n z_n + \sum_{i=2}^n \sum_{\ell=1}^{i-1} (a_\nu^2(i) + a_\nu^2(\ell)) b_{i-\ell} \\
&= b_0 n z_n + \sum_{i=2}^n \sum_{u=1}^{i-1} (a_\nu^2(i) + a_\nu^2(i-u)) b_u \\
&= b_0 n z_n + \sum_{i=2}^n a_\nu^2(i) \sum_{u=1}^{i-1} b_u + \sum_{i=2}^n \sum_{u=1}^{i-1} a_\nu^2(i-u) b_u.
\end{aligned}$$

Using (2), we obtain

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

and

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

Hence

$$B \leq C z_n n^{\theta+1}. \quad (13)$$

Putting (9), (11), (12) and (13) together, we obtain

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

This complete the proof of Proposition 1. □

Proof of Theorem 1. We expand the function f_ν on \mathcal{B} as

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

where

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

We have

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

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

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

where

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

Using Proposition 1 and the definition of j_0 , we obtain

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

Since $p \geq 2$, we have $B_{p,r}^s(M) \subseteq B_{2,\infty}^s(M)$. Hence

$$B \leq C2^{-2j_0s} \leq C \left(\frac{z_n}{n^{1-\theta}} \right)^{2s/(2s+4)}.$$

Therefore

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

The proof of Theorem 1 is complete. \square

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. *Scand. J. Statist.*, 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, *J. Nonparametric Statistics*, 10, 405-420.
- Hall, P. and Zhou, X.H. (2003). Nonparametric estimation of component distributions in a multivariate mixture. *Ann. Statist.*, 31 (1), 201-224.
- Härdle, W., Kerkycharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelet, Approximation and Statistical Applications*. Lectures Notes in Statistics, New York 129, Springer Verlag.
- Lehmann, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.*, 37, 1137-1153.
- Maiboroda, R. E. (1996). Estimators of components of a mixture with varying concentrations. *Ukrain. Mat. Zh.*, 48, 4, 562-566.
- Masry, E. (2001). Multivariate probability density estimation for associated processes: Strong consistency and rates. *Statist. Probab. Lett.*, 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.
- Tsybakov, A. B. (2004). *Introduction à l'estimation non-paramétrique*. Springer.