Adaptive wavelet estimation of a density from mixtures under multiplicative censoring

Yogendra P. Chaubey, Christophe Chesneau, Hassan Doosti

Adaptive wavelet estimation of a density from mixtures under multiplicative censoring

Yogendra P. Chaubey[1,*] **Christophe Chesneau**[2,†]
and
Hassan Doosti[3]

[1]*Department of Mathematics and Statistics, Concordia University, Montréal, QC H3G 1M8, CANADA. e-mail:* `chaubey@alcor.concordia.ca`

[2]*Laboratoire de Mathématiques Nicolas Oresme, Université de Caen BP 5186, F 14032 Caen Cedex, FRANCE. e-mail:* `chesneau@math.unicaen.fr`

[3]*Department of Mathematics, Kharazmi University, Tehran, IRAN and Department of Mathematics and Statistics, The University of Melbourne, Melbourne, AUSTRALIA. e-mail:* `h.doosti@tmu.ac.ir` - `hassan.doosti@unimelb.edu.au`

Abstract: In this paper, a mixture model under multiplicative censoring is considered. We investigate the estimation of a component of the mixture (a density) from the observations. A new adaptive estimator based on wavelets and a hard thresholding rule is constructed for this problem. Under mild assumptions on the model, we study its asymptotic properties by determining an upper bound of the mean integrated squared error over a wide range of Besov balls. We prove that the obtained upper bound is sharp.

AMS 2000 subject classifications: Primary 62G07, 62G20; secondary 60K35.
Keywords and phrases: Density estimation, mixture, Besov balls, wavelets, hard thresholding.

1. Introduction

First of all, let us present the standard density estimation problem under multiplicative censoring as described in Vardi (1989). We observe n independent random variables $Y_1, \ldots, Y_n$ where

$$Y_i = U_i X_i, \qquad i \in \{1, \ldots, n\}, \tag{1.1}$$

$U_1, \ldots, U_n$ are n unobserved *i.i.d.* random variables having the uniform distribution on $[0, 1]$ and $X_1, \ldots, X_n$ are n unobserved *i.i.d.* random variables. We suppose that $X_1, \ldots, X_n$ and $U_1, \ldots, U_n$ are independent. We aim to estimate f, the unknown density of X_1, from $Y_1, \ldots, Y_n$.

It is shown in Vardi (1989) [see also Vardi and Zhang (1992)] how this model unifies several well-studied statistical problems, including non-para-

---

metric inference for renewal processes, certain non-parametric deconvolution problems and estimation of decreasing densities. Among the methods used to solve this problem, there is the method of singular value decomposition (or orthonormal series based expansion) developed in Andersen and Hansen (2001), the method using the wavelet technique used in Abbaszadeh *et al.* (2012) and the kernel method developed in Asgharian *et al.* (2012).

In this paper, we consider a more general model incorporating mixtures. Here, we observe n independent random variables $Y_1, \ldots, Y_n$ according to (1.1) where $U_1, \ldots, U_n$ are unobserved *i.i.d.* random variables having the uniform distribution on $[0, 1]$ and $X_1, \ldots, X_n$ are unobserved independent random variables but not (necessarily) identically distributed: 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 the following finite mixture:

$$h_i(x) = \sum_{d=1}^{m} w_d(i) f_d(x), \tag{1.2}$$

where $w_d(i) = \mathbb{P}(I_i = d)$ and f_d is the conditional density of X_i given $\{I_i = d\}$. We suppose that all these densities are unknown, the weight $w_d(i)$ is known, and $X_1, \ldots, X_n$ and $U_1, \ldots, U_n$ are independent. For a fixed $\nu \in \{1, \ldots, m\}$, we aim to estimate f_ν i.e. the ν-th component of the mixture, from $Y_1, \ldots, Y_n$.

Note that, when $m = 1$, $w_1(1) = \ldots = w_1(n) = 1$ and $f_\nu = f_1$, we arrive at the density estimation problem under multiplicative censoring described in the first paragraph. The problem of estimating f_ν from $X_1, \ldots, X_n$ (not $Y_1, \ldots, Y_n$) has been considered in several papers [see *e.g.* Maiboroda (1996), Hall and Zhou (2003), Pokhyl'ko (2005) and Prakasa Rao (2010)]. However, to the best of our knowledge, the estimation of f_ν from $Y_1, \ldots, Y_n$ is a new statistical problem that has potential applications in signal processing, biology, industry and telecommunications. For instance, consider the example of a sensor network (i.e. a collection of spatially distributed autonomous sensors intended to measure physical phenomena at diverse locations). Suppose that it contains m nodes, each sensor records only one physical value of the phenomena and only one sensor information is collected at a time. Now, given n multiplicative noisy versions $Y_1, \ldots, Y_n$ of the physical phenomena from non-necessarily identified sensors (i.e. unobserved $I_1, \ldots, I_n$), the goal is to estimate the density function of the physical phenomena at any sensor $\nu \in \{1, \ldots, m\}$. An overview of sensor networks can be found in Akyildiz *et al.* (2002). The reader may be referred to Chabert *et al.* (2004) and references therein for applications in multiplicative noise set-up in signal processing.

Among the numerous existing methods (such as Kernel, Spline, Wavelets and others) [see e.g. Prakasa Rao (1983, 1999) and Tsybakov (2004)], we investigate the estimation of f_ν via wavelet methods. They are attractive for nonparametric function estimation because of their spatial adaptivity, computational efficiency and asymptotic optimality properties. They can achieve near optimal convergence rates over a wide range of function classes (such as Besov balls) and enjoy excellent mean integrated squared error (MISE) properties when used to esti-

mate spatially inhomogeneous function [see Antoniadis (1997) and Härdle *et al.* (1998)].

In this paper, we develop an adaptive wavelet estimator for f_ν based on the hard thresholding rule [introduced in Donoho and Johnstone (1994, 1995)]. In its construction, we use the new methodology developed in Abbaszadeh *et al.* (2012) for the density estimation under bias and multiplicative censoring, the "observations thresholding technique" introduced by Delyon and Juditsky (1996), and some tools on mixtures described in Maiboroda (1996). We evaluate the performance of the proposed density estimator by considering the MISE and assuming that f_ν belongs to a Besov ball. Under mild assumptions on the weights of the mixtures, we prove that our estimator attains a sharp rate of convergence: it is the one attained by the "best" non-adaptive linear wavelet estimator up to a logarithmic term, where "best" is used in the sense of minimum MISE.

The outline of the paper is as follows. We provide detailed assumptions on the mixture model considered here in Section 2. Section 3 describes the wavelet basis on $[0, 1]$ and the Besov balls. The estimators are presented in Section 4 and the main results are provided in Section 5. The proofs are gathered in Section 6.

2. Assumptions

Additional assumptions on our model are presented below.

- Without loss of generality, we assume that the support of f_d is $[0, 1]$ and $f_d \in \mathbb{L}^2([0, 1])$ for any $d \in \{1, \ldots, m\}$, where, for any $p \geq 1$,

$$\mathbb{L}^p([0, 1]) = \left\{ u : [0, 1] \to \mathbb{R}; \ \|u\|_p = \left(\int_0^1 |u(x)|^p dx \right)^{1/p} < \infty \right\}.$$

  Note that, for any $i \in \{1, \ldots, n\}$, the density of Y_i is

$$g_i(x) = \int_x^1 \frac{h_i(y)}{y} dy, \qquad x \in [0, 1],$$

  where h_i is as given in (1.2).
  We suppose that there exists a known constant $C_* > 0$ such that

$$\max_{i \in \{1, \ldots, n\}} \sup_{x \in [0, 1]} g_i(x) \leq C_*. \tag{2.1}$$

  Such a boundedness assumption is standard in the density estimation problems from mixtures [see *e.g.* Maiboroda (1996), Pokhyl'ko (2005) and Prakasa Rao (2010)].
- We now formulate some assumptions on the weights of the mixture. Consider the matrix

$$\Gamma_n = \left(\frac{1}{n} \sum_{i=1}^n w_k(i) w_\ell(i) \right)_{(k, \ell) \in \{1, \ldots, m\}^2}. \tag{2.2}$$

We assume that $\det(\Gamma_n) > 0$.

Then the solution of the following quadratic objective with linear constraints,

$$\min_{b \in \mathbb{R}^n} \frac{1}{n} \sum_{i=1}^{n} b_i^2 \quad \text{such that} \quad \frac{1}{n} \sum_{i=1}^{n} b_i w_d(i) = \delta_{\nu,d}, \text{ for } d \in \{1, \cdots, m\}, \quad (2.3)$$

where $\delta_{\nu,d}$ denotes the Kronecker delta, is unique and is denoted by $(a_\nu(1), \ldots, a_\nu(n))$. Specifically, we have

$$a_\nu(i) = \frac{1}{\det(\Gamma_n)} \sum_{k=1}^{m} (-1)^{k+\nu} \zeta_{\nu,k}^n w_k(i), \quad (2.4)$$

where $\zeta_{\nu,k}^n$ denotes the determinant of the minor (ν, k) of the matrix Γ_n. Details concerning the above results can be found in Maiboroda (1996) and Pokhyl'ko (2005).

Remark 2.1. *The definition of $(a_\nu(1), \ldots, a_\nu(n))$ (2.3) allows us to*

- *construct unbiased wavelet coefficients estimators well adapted to the structure of our mixture model (see Proposition 4.1),*
- *obtain sharp rate of convergence (the quantity $(1/n) \sum_{i=1}^{n} a_\nu^2(i)$ will play a central role); see Theorems 5.1 and 5.2.*

We set

$$\rho_n = \frac{1}{n} \sum_{i=1}^{n} a_\nu^2(i) \quad (2.5)$$

and, for technical reasons, we suppose that $\lim_{n \to \infty} n/\rho_n = \infty$.

3. Wavelets and Besov balls

3.1. Wavelets

For the purposes of this paper, we use the compactly supported wavelet basis on $[0, 1]$ briefly described below.

Let $N \geq 10$ and ϕ and ψ be the initial wavelet functions of the Daubechies wavelets $db2N$. In particular, these functions are compactly supported with $\text{supp}(\phi) \subseteq [1 - N, N]$ and $\text{supp}(\psi) \subseteq [1 - N, N]$ (for the sake of simplicity), and belong to $\mathcal{C}^1$.

We define the functions $\phi_{j,k}$ and $\psi_{j,k}$ by dilation and translation of ϕ and ψ respectively, and normalization:

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

Then, with an appropriate treatment at the boundaries, there exists an integer τ such that, for any integer $\ell \geq \tau$,

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

is an orthonormal basis of $\mathbb{L}^2([0,1])$.

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

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

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

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

We refer to Cohen *et al.* (1993) and Mallat (2009) for more details.

3.2. Besov balls

As is traditional in the wavelet estimation literature, we will investigate the performances of our estimators by assuming that the unknown function f_ν belongs to a Besov ball defined below.

Let $M > 0$, $s > 0$, $p \geq 1$ and $r \geq 1$. Set, for every measurable function u on $[0,1]$ and $\epsilon \geq 0$,
$$\Delta_\epsilon(u)(x) = u(x+\epsilon) - u(x),$$
$\Delta_\epsilon^2(u)(x) = \Delta_\epsilon(\Delta_\epsilon(u))(x)$ and identically, $\Delta_\epsilon^N(u)(x) = \Delta_\epsilon^{N-1}(\Delta_\epsilon(u))(x)$ for any positive integer N. Let

$$\rho^N(t,u,p) = \sup_{\epsilon \in [-t,t]} ||\Delta_\epsilon^N(u)||_p.$$

Then, for $s \in [0,N)$, we define the Besov ball $B_{p,r}^s(M)$ of radius $M > 0$ by

$$B_{p,r}^s(M) = \left\{ u \in \mathbb{L}^p([0,1]); \ \left(\int_0^1 \left(\frac{\rho^N(t,u,p)}{t^s} \right)^r \frac{dt}{t} \right)^{1/r} \leq M \right\},$$

with the usual modifications if $p = \infty$ or $r = \infty$.

The parameter s is related to the smoothness of the function u, whereas p and r are norm parameters.

We have the following equivalence (see Härdle *et al.* (1998), Corollary 9.1): $u \in B_{p,r}^s(M)$ with $s \in (0,N)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients given by (3.1) satisfy

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

$$\leq M^*,$$

with the usual modifications if $p = \infty$ or $r = \infty$. For a particular choice of s, p and r, $B_{p,r}^s(M)$ contains the Hölder and Sobolev balls [see Meyer (1990)].

4. Estimators

4.1. Wavelet coefficients estimators

The first step in estimating f_ν consists of expanding f_ν on $\mathcal{B}$ and estimating its unknown wavelet coefficients. Here we adapt the methodology of Abbaszadeh *et al.* (2012). We consider the operator $T : \mathcal{C}^1([0,1]) \to \mathcal{C}^0([0,1])$ defined by

$$T(u)(x) = (xu(x))' = u(x) + xu'(x), \qquad u \in \mathcal{C}^1([0,1]), \quad x \in [0,1]. \quad (4.1)$$

For any integer $j \geq \tau$ and any $k \in \{0, \dots, 2^j - 1\}$, we estimate the wavelet coefficients of f_ν, i.e. $\alpha_{j,k} = \int_0^1 f_\nu(x)\phi_{j,k}(x)dx$ and $\beta_{j,k} = \int_0^1 f_\nu(x)\psi_{j,k}(x)dx$, by respectively

$$\widehat{\alpha}_{j,k} = \frac{1}{n} \sum_{i=1}^n a_\nu(i) T(\phi_{j,k})(Y_i) \quad (4.2)$$

and

$$\widehat{\beta}_{j,k} = \frac{1}{n} \sum_{i=1}^n a_\nu(i) T(\psi_{j,k})(Y_i) \mathbf{1}_{\{|a_\nu(i)T(\psi_{j,k})(Y_i)| \leq \eta_j\}}, \quad (4.3)$$

where Y_i is defined by (1.1), $a_\nu(i)$ by (2.4), T by (4.1), $\mathbf{1}_{\mathcal{A}}$ is the indicator function on $\mathcal{A}$ for any random event $\mathcal{A}$,

$$\eta_j = \theta 2^j \sqrt{\frac{n\rho_n}{\ln(n/\rho_n)}},$$

$$\theta = \sqrt{2C_* \left(1 + \int_{1-N}^N (\psi'(y))^2 dy\right)}, \quad (4.4)$$

C_* is the constant in (2.1) and ρ_n is defined by (2.5).

The wavelet coefficient estimators $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$ are motivated by the properties of the constants $a_\nu(i)$ defined in (2.3) and (2.4). We can write

$$f_\nu(x) = \frac{1}{n} \sum_{i=1}^n a_\nu(i) h_i(x)$$

and use the fact that (see the details in the proof of Proposition 4.1)

$$\int_0^1 h_i(x)\phi_{j,k}(x)dx = \mathbb{E}(T(\phi_{j,k})(Y_i)).$$

Thus $\widehat{\alpha}_{j,k}$ provides an unbiased estimator of $\alpha_{j,k}$ where as $\widehat{\beta}_{j,k}$ is based on the "observations thresholding technique" that had been firstly introduced in Delyon and Juditsky (1996) to estimate the regression function in a regression

model with general errors. In our study, the role of its selection is to provide a good estimator of $\beta_{j,k}$ under mild assumptions on the distributions of $I_1, \ldots, I_n$. To be more specific, we have no boundedness assumptions on the eigenvalues of the matrix Γ_n (as in Autin and Pouet (2011)) and, more generally, on $a_\nu(i)$. This point is developed in Remarks 5.3 and 5.4. Our statistical approach has been used for other mixture models in Chesneau (2011a,b) and Chesneau and Fadili (2012).

Proposition 4.1. *For any integer $j \geq \tau$ and any $k \in \{0, \ldots, 2^j - 1\}$, let $\alpha_{j,k}$ and $\beta_{j,k}$ be the wavelet coefficients of f_ν as given in (3.1). Then*

- *$\widehat{\alpha}_{j,k}$ defined by (4.2) is an unbiased estimator of $\alpha_{j,k}$,*
- *we have*

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n a_\nu(i)T(\psi_{j,k})(Y_i)\right) = \beta_{j,k}.$$

Remark 4.1. *In addition to the unbiased properties described in Proposition 4.1, $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$ satisfy powerful moments and concentration inequalities, i.e.,*

Moments inequality: there exists a constant $C > 0$ such that

$$\mathbb{E}\left((\widehat{\alpha}_{j,k} - \alpha_{j,k})^2\right) \leq C2^{2j}\frac{\rho_n}{n}$$

and

$$\mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k}\right)^4\right) \leq C2^{4j}\frac{(\rho_n \ln(n/\rho_n))^2}{n^2}.$$

Concentration inequality: for a large enough κ, we have

$$\mathbb{P}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq (\kappa/2)\theta 2^j\sqrt{\frac{\rho_n \ln(n/\rho_n)}{n}}\right) \leq 2\left(\frac{\rho_n}{n}\right)^2,$$

where θ is defined by (4.4).

The technical details can be found in the proof of Proposition 6.2.

The next subsections are devoted to the description of our main estimators for f_ν.

4.2. Wavelet linear estimator

We define the linear estimator $\widehat{f}_\nu^l$ by

$$\widehat{f}_\nu^l(x) = \sum_{k=0}^{2^{j_0}-1} \widehat{\alpha}_{j_0,k}\phi_{j_0,k}(x), \qquad x \in [0,1], \tag{4.5}$$

where $\widehat{\alpha}_{j,k}$ is defined by (4.2) and j_0 is an integer which will be chosen later (see Theorem 5.1).

In the case $m = 1$, $w_1(1) = \ldots = w_1(n) = 1$ and $f_\nu = f_1$, $\widehat{f}_\nu^l$ becomes the linear wavelet estimator developed in Abbaszadeh *et al.* (2012).

For a survey on linear wavelet estimators for various density models (including the density model under multiplicative censoring), we refer to Chaubey *et al.* (2011).

4.3. Wavelet hard thresholding estimator

We define the hard thresholding estimator $\widehat{f}_\nu^h$ by

$$\widehat{f}_\nu^h(x) = \sum_{k=0}^{2^\tau - 1} \widehat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j - 1} \widehat{\beta}_{j,k} \mathbf{1}_{\left\{|\widehat{\beta}_{j,k}| \geq \kappa \lambda_j\right\}} \psi_{j,k}(x), \quad x \in [0,1],$$

(4.6)

where $\widehat{\alpha}_{j,k}$ is defined by (4.2), $\widehat{\beta}_{j,k}$ by (4.3), τ is the primary resolution level of the wavelet basis $\mathcal{B}$, j_1 is the integer satisfying

$$\frac{1}{2}\left(\frac{n}{\rho_n}\right)^{1/3} < 2^{j_1} \leq \left(\frac{n}{\rho_n}\right)^{1/3},$$

$\kappa \geq 2 + 8/3 + 2\sqrt{4 + 16/9}$,

$$\lambda_j = \theta 2^j \sqrt{\frac{\rho_n \ln(n/\rho_n)}{n}}.$$

and θ is defined by (4.4).

The definitions of the thresholds η_j and λ_j are based on theoretical considerations; they are chosen in such a way that the MISE of $\widehat{f}_\nu^h$ is as small as possible, while being independent of f_ν (and its degree of smoothness). This independence makes $\widehat{f}_\nu^h$ adaptive. Technical details can be found in the proof of Theorem 6.1.

The idea of the considered thresholding is to set to zero all wavelet coefficients that are less than a particular threshold and to estimate the others. The resulting estimators often possess optimal (or near-optimal) rates of convergence in the minimax sense. Moreover, they are typically implemented via fast algorithms which makes them appealing in practice. For the constructions of hard thresholding wavelet estimators in the standard density model, see e.g. Donoho *et al.* (1996), Delyon and Juditsky (1996) and Härdle *et al.* (1998), and for the standard density mixture model, see Pokhyl'ko (2005). For the multiplicative censoring density model from biased data, see Abbaszadeh *et al.* (2012). To the best of our knowledge, $\widehat{f}_\nu^h$ is the first adaptive wavelet estimator proposed for the density estimation from mixtures under multiplicative censoring.

5. Results

Theorem 5.1 below explores the asymptotic performance of $\widehat{f}_\nu^l$ under the MISE over Besov balls $B_{p,r}^s(M)$. We distinguish the "homogeneous zone" corresponding to $p \geq 2$ and the "inhomogeneous zone" corresponding to $p \in [1,2)$, following the classification of Härdle *et al.* (1998).

Theorem 5.1. *Consider* (1.1) *under the assumptions of Section 2.*

(i) *Suppose that* $f_\nu \in B_{p,r}^s(M)$ *with* $s \in (0, N)$, $p \geq 2$ *and* $r \geq 1$. *Let* $\widehat{f}_\nu^l$ *be given by* (4.5) *with* j_0 *satisfying*

$$\frac{1}{2}\left(\frac{n}{\rho_n}\right)^{1/(2s+3)} < 2^{j_0} \leq \left(\frac{n}{\rho_n}\right)^{1/(2s+3)}. \qquad (5.1)$$

Then there exists a constant $C > 0$ *such that*

$$\mathbb{E}\left(\|\widehat{f}_\nu^l - f_\nu\|_2^2\right) \leq C\left(\frac{\rho_n}{n}\right)^{2s/(2s+3)}.$$

(ii) *Suppose that* $f_\nu \in B_{p,r}^s(M)$ *with* $s \in (1/p, N)$, $p \in [1,2)$ *and* $r \geq 1$. *Let* $\widehat{f}_\nu^l$ *be given by* (4.5) *with* j_0 *satisfying*

$$\frac{1}{2}\left(\frac{n}{\rho_n}\right)^{1/(2s_*+3)} < 2^{j_0} \leq \left(\frac{n}{\rho_n}\right)^{1/(2s_*+3)}, \qquad (5.2)$$

where $s_* = s + 1/2 - 1/p$.
Then there exists a constant $C > 0$ *such that*

$$\mathbb{E}\left(\|\widehat{f}_\nu^l - f_\nu\|_2^2\right) \leq C\left(\frac{\rho_n}{n}\right)^{2s_*/(2s_*+3)}.$$

The proof of Theorem 5.1 (see §6) uses moment inequalities on (4.2) and (4.3), and a suitable decomposition of the MISE.

Let us mention that, in each zone of the Besov balls, the integer j_0 is chosen to minimize the MISE of $\widehat{f}_\nu^l$ in a nonadaptive way.

Remark 5.1. *The variability of the performance of linear estimators according to the homogeneous nature of the zones of the Besov balls is a well known phenomenon. In particular, in the context of the standard nonparametric smoothing (such as for density and regression), the rates are often non-optimal (in the minimax sense) over the inhomogeneous zone. We refer to Härdle* et al. *(1998), p.148.*

Theorem 5.2 given next investigates the asymptotic performance of $\widehat{f}_\nu^h$ under the MISE over Besov balls (the homogeneous and inhomogeneous zones are considered simultaneously).

Theorem 5.2. *Consider* (1.1) *under the assumptions of Section* 2. *Let* $\widehat{f}_\nu^h$ *be given by* (4.6). *Suppose that* $f_\nu \in B_{p,r}^s(M)$ *with* $r \geq 1$, $\{p \geq 2$ *and* $s \in (0, N)\}$ *or* $\{p \in [1, 2)$ *and* $s \in (3/p, N)\}$. *Then there exists a constant* $C > 0$ *such that*

$$\mathbb{E}\left(\|\widehat{f}_\nu^h - f_\nu\|_2^2\right) \leq C \left(\frac{\rho_n \ln(n/\rho_n)}{n}\right)^{2s/(2s+3)}.$$

The proof of Theorem 5.2 uses a general result derived to Chesneau (2011b) (see the proof of his Theorem 2).

Theorem 5.2 shows that, in addition to being adaptive, $\widehat{f}_\nu^h$ attains a sharp rate of convergence. More specifically, adapting the terminology of Härdle *et al.* (1998),

- *for the homogeneous zone* $(p \geq 2)$: $\widehat{f}_\nu^h$ *attains the same rate of convergence to* $\widehat{f}_\nu^l$, *i.e.* $v_n = (\rho_n/n)^{2s/(2s+3)}$, *up to the logarithmic term* $(\ln(n/\rho_n))^{2s/(2s+3)}$.
- *for the inhomogeneous zone* $(p \in [1, 2))$: $\widehat{f}_\nu^h$ *attains a better rate of convergence to* $\widehat{f}_\nu^l$.

Remark 5.2. *Let us mention that, for the standard nonparametric models, other types of adaptive procedures are able to remove the extra logarithmic term in the rate of convergence, for example,*

- *the Lepski procedure [see e.g. Lepski (1990, 1991, 1992) and Lepski and Spokoiny (1995)], and*
- *the BlockJS procedure [see e.g. Cai (1999, 2002)].*

Although the adaptation of these procedures to our statistical problem is possible, their asymptotic properties are not established; new important theoretical difficulties appear and it is not immediately clear how to solve them.

Remark 5.3. *Suppose that the smallest eigenvalue of the matrix* Γ_n *(given in* (2.2)*) is larger than or equal to* K *with* $K \in (0, 1)$. *Then Lemma 6.1 of Autin and Pouet (2011) gives*

$$\frac{1}{n} \sum_{d=1}^m \sum_{i=1}^n a_d^2(i) \leq \frac{m}{K}.$$

Hence $\rho_n \leq m/K$ *and the rate of convergence in Theorem 5.2 becomes* $(\ln n/n)^{2s/(2s+3)}$. *It corresponds to the optimal one (up to a logarithmic term) for the standard multiplicative density model [see Andersen and Hansen(2001)].*

Remark 5.4. *Suppose that there exists a constant* $C > 0$ *satisfying*

$$\max_{d \in \{1,\dots,m\}} \max_{i \in \{1,\dots,n\}} |a_d(i)| \leq C.$$

Then, with j_1 *such that*

$$\frac{c_*}{2} \left(\frac{n}{\ln n}\right)^{1/3} \leq 2^{j_1} \leq c_* \left(\frac{n}{\ln n}\right)^{1/3}$$

for a suitable constant c_*, the indicator function in (4.3) is equal to one and Theorem 5.2 can be proved with the rate of convergence $(\ln n/n)^{2s/(2s+3)}$.

Conclusion and perspectives. We construct a new adaptive wavelet estimator to estimate the function f_ν from (1.2). Under mild assumptions, we prove that it attains a sharp rate of convergence for a wide class of functions. Possible perspectives are to

- determine the optimal lower bound of MISE for (1.1), perhaps using the Fano Lemma presented in Tsybakov (2004). However, it is not immediately clear how to extend this to the mixture models.
- improve the estimation of f_ν by removing the extra logarithmic term (see Remark 5.2),
- investigate the case where the distributions of $I_1, \ldots, I_n$ are unknown,
- consider other estimation problems as the adaptive estimation of the q-th derivatives of f_ν (as in Prakasa Rao (1996) for the standard density model and in Prakasa Rao (2010) for the standard mixture model).

All these aspects need further investigations that we leave for a future work.

6. Proofs

In this section, we consider (1.1) 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 depend on ϕ or ψ.

6.1. Proof of Proposition 4.1

Proof of Proposition 4.1. Recall that the density of Y_i can be expressed in terms of h_i in the following manner:

$$g_i(x) = \int_x^1 \frac{h_i(y)}{y} dy, \qquad x \in [0,1].$$

So $h_i(x) = -xg_i'(x)$, $x \in [0,1]$ a.s.. It follows from integration by parts, $g_i(1) = 0$, and $\phi_{j,k}(0)$ and $\phi_{j,k}(1)$ exist, that

$$\int_0^1 h_i(x)\phi_{j,k}(x)dx = -\int_0^1 xg_i'(x)\phi_{j,k}(x)dx$$

$$= -[xg_i(x)\phi_{j,k}(x)]_0^1 + \int_0^1 g_i(x)\left(\phi_{j,k}(x) + x(\phi_{j,k})'(x)\right)dx$$

$$= \int_0^1 g_i(x)T(\phi_{j,k})(x)dx = \mathbb{E}(T(\phi_{j,k})(Y_i)), \qquad (6.1)$$

where T denotes the operator in (4.1).

Using (6.1), (1.2) and (2.3), we obtain

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

Similarly, taking ψ instead of ϕ, we prove the second point that completes the proof of Proposition 4.1. $\qquad\square$

6.2. Proofs of Theorems 5.1 and 5.2

Proof of Theorem 5.1.

(i) We expand the function f_ν on $\mathcal{B}$ at the level j_0 given in (5.1) :

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

where $\alpha_{j_0,k} = \int_0^1 f_\nu(x)\phi_{j_0,k}(x)dx$ and $\beta_{j,k} = \int_0^1 f_\nu(x)\psi_{j,k}(x)dx$.
We can write

$$
\mathbb{E}\left(\|\widehat{f}_\nu^l - f_\nu\|_2^2\right) = \sum_{k=0}^{2^{j_0}-1}\mathbb{E}\left((\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k})^2\right) + \sum_{j=j_0}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^2. \qquad (6.2)
$$

Let us now bound the first term. Since, by Proposition 4.1, $\widehat{\alpha}_{j,k}$ is an unbiased estimator of $\alpha_{j,k}$, we have

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

By the definition of T in (4.1), for any $i \in \{1,\ldots,n\}$, we have

$$
\mathbb{E}\left((T(\phi_{j_0,k})(Y_i))^2\right) \leq 2(A+B),
$$

where
$$
A = \mathbb{E}\left(\phi_{j_0,k}^2(Y_i)\right), \qquad B = \mathbb{E}\left(Y_i^2((\phi_{j_0,k})'(Y_i))^2\right).
$$

Since, by (2.1), $\max_{i \in \{1,\dots,n\}} \sup_{x \in [0,1]} g_i(x) \leq C_*$, we have

$$A = \int_0^1 \phi_{j_0,k}^2(x) g_i(x) dx \leq C_* \int_0^1 \phi_{j_0,k}^2(x) dx = C_*.$$

On the other hand, using again (2.1), $(\phi_{j_0,k})'(x) = 2^{3j_0/2} \phi'(2^{j_0}x - k)$, making the change of variables $y = 2^{j_0}x - k$ and owing to $\mathrm{supp}(\phi) \subseteq [1 - N, N]$, we obtain

$$B = \int_0^1 x^2((\phi_{j_0,k})'(x))^2 g_i(x) dx \leq C_* \int_0^1 x^2((\phi_{j_0,k})'(x))^2 dx$$

$$\leq\ C_* \int_0^1 ((\phi_{j_0,k})'(x))^2 dx = C_* 2^{3j} \int_0^1 (\phi'(2^{j_0}x - k))^2 dx$$

$$\leq\ C_* 2^{2j_0} \int_{1-N}^{N} (\phi'(y))^2 dy.$$

Hence

$$\mathbb{E}\left((T(\phi_{j_0,k})(Y_i))^2\right) \leq C_o 2^{2j_0}, \tag{6.3}$$

where $C_o = 2C_* \left(1 + \int_{1-N}^{N} (\phi'(y))^2 dy\right)$.
Therefore

$$\mathbb{E}\left((\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k})^2\right) \leq C 2^{2j_0} \frac{1}{n} \left(\frac{1}{n} \sum_{i=1}^{n} a_\nu^2(i)\right) = C 2^{2j_0} \frac{\rho_n}{n}.$$

For j_0 satisfying (5.1) it holds

$$\sum_{k=0}^{2^{j_0}-1} \mathbb{E}\left((\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k})^2\right) \leq C 2^{j_0} 2^{2j_0} \frac{\rho_n}{n} \leq C \left(\frac{\rho_n}{n}\right)^{2s/(2s+3)}. \tag{6.4}$$

On the other hand, for $p \geq 2$, we have $f_\nu \in B_{p,r}^s(M) \subseteq B_{2,\infty}^s(M)$ [see Härdle *et al.* (1998), Corollary 9.2], which implies

$$\sum_{j=j_0}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \leq C 2^{-2j_0 s} \leq C \left(\frac{\rho_n}{n}\right)^{2s/(2s+3)}. \tag{6.5}$$

It follows from (6.2), (6.4) and (6.5) that

$$\mathbb{E}\left(\|\widehat{f}_\nu^l - f_\nu\|_2^2\right) \leq C \left(\frac{\rho_n}{n}\right)^{2s/(2s+3)}.$$

The part (i) of Theorem 5.1 is proved.

(ii) Let $p \in [1, 2)$. The proof is similar to (i) with s_* instead of s: we use the expansion (6.2) at the level j_0 given in (5.2), and the inclusion $B_{p,r}^s(M) \subseteq B_{2,\infty}^{s_*}(M)$ (thanks to $p \in [1, 2)$ and the definition of s_*, i.e. $s_* = s + 1/2 - 1/p$, [see Härdle *et al.* (1998), Corollary 9.2]), we obtain the desired result.

This completes the proof of Theorem 5.1. □

Proof of Theorem 5.2. First of all, we present a general result (see Theorem 6.1) derived from Theorem 2 of Chesneau (2011b). Its proof is given in §6.2. In this theorem, we consider a more general form of the hard thresholding estimator $\hat{f}_\nu^h$ denoted by $\hat{f}_H$ for estimating an unknown function $f \in \mathbb{L}^2([0,1])$ from n independent random variables $W_1, \ldots, W_n$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ (We use the notations of Section 3):

$$\widehat{f}_H(x) = \sum_{k=0}^{2^\tau - 1} \widehat{\alpha}_{\tau,k}\phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \widehat{\beta}_{j,k}\mathbf{1}_{\{|\widehat{\beta}_{j,k}| \geq \kappa\vartheta_j\}}\psi_{j,k}(x), \tag{6.6}$$

where

$$\widehat{\alpha}_{j,k} = \frac{1}{\upsilon_n}\sum_{i=1}^{n} q_i(\phi_{j,k}, W_i), \tag{6.7}$$

$$\widehat{\beta}_{j,k} = \frac{1}{\upsilon_n}\sum_{i=1}^{n} q_i(\psi_{j,k}, W_i)\mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)|\leq\varsigma_j\}}, \tag{6.8}$$

$$\varsigma_j = \theta_\psi 2^{\omega j}\frac{\upsilon_n}{\sqrt{\mu_n\ln\mu_n}} \tag{6.9}$$

$$\vartheta_j = \theta_\psi 2^{\omega j}\sqrt{\frac{\ln\mu_n}{\mu_n}}, \tag{6.10}$$

$\kappa \geq 2 + 8/3 + 2\sqrt{4 + 16/9}$ and j_1 is the integer satisfying

$$\frac{1}{2}\mu_n^{1/(2\omega+1)} < 2^{j_1} \leq \mu_n^{1/(2\omega+1)}. \tag{6.11}$$

Here, we suppose that there exist

- n functions $q_1, \ldots, q_n$ with $q_i : \mathbb{L}^2([0,1]) \times \mathbb{R} \to \mathbb{R}$ for any $i \in \{1, \ldots, n\}$,
- two sequences of real numbers $(\upsilon_n)_{n\in\mathbb{N}}$ and $(\mu_n)_{n\in\mathbb{N}}$ satisfying $\lim_{n\to\infty} \upsilon_n = \infty$ and $\lim_{n\to\infty} \mu_n = \infty$

such that, for $\gamma \in \{\phi, \psi\}$,

(A1). any integer $j \geq \tau$ and any $k \in \{0, \ldots, 2^j - 1\}$,

$$\mathbb{E}\left(\frac{1}{\upsilon_n}\sum_{i=1}^{n} q_i(\gamma_{j,k}, W_i)\right) = \int_0^1 f(x)\gamma_{j,k}(x)dx.$$

(A2). there exist two constants, $\theta_\gamma > 0$ and $\omega \geq 0$, such that, for any integer $j \geq \tau$ and any $k \in \{0, \ldots, 2^j - 1\}$,

$$\sum_{i=1}^{n} \mathbb{E}\left((q_i(\gamma_{j,k}, W_i))^2\right) \leq \theta_\gamma^2 2^{2\omega j}\frac{\upsilon_n^2}{\mu_n}.$$

Remark 6.1. *The estimators $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$ defined in the above theorem coincide with those in (4.2) and (4.3), respectively by choosing $v_n = n$. The particular form of the estimator chosen will be clear from the context.*

Now we are ready to state the theorem required for proving Theorem 5.2.

Theorem 6.1. *Let $\hat{f}_H$ be (6.6) under (A1) and (A2). Suppose that $f \in B^s_{p,r}(M)$ with $r \geq 1$, $\{p \geq 2$ and $s \in (0, N)\}$ or $\{p \in [1, 2)$ and $s \in ((2\omega + 1)/p, N)\}$. Then there exists a constant $C > 0$ such that*

$$\mathbb{E}\left(\|\widehat{f}_H - f\|_2^2\right) \leq C \left(\frac{\ln \mu_n}{\mu_n}\right)^{2s/(2s+2\omega+1)}.$$

Remark 6.2. *Theorem 6.1 is still valid*

- *with j_1 such that*

$$\frac{c_*}{2}\left(\frac{\mu_n}{\ln \mu_n}\right)^{1/(2\omega+1)} < 2^{j_1} \leq c_*\left(\frac{\mu_n}{\ln \mu_n}\right)^{1/(2\omega+1)},$$

  $c_* > 0$, instead of (6.11).
- *without θ_ψ in the definitions of (6.9) and (6.10), but with $\kappa \geq 2\theta_\psi^2 + 8/3 + 2\sqrt{4\theta_\psi^2 + 16/9}$.*

Let us now investigate the assumptions (A1) and (A2) of Theorem 6.1 for the model (1.1). Set, for any $\gamma \in \{\phi, \psi\}$,

$$q_i(\gamma_{j,k}, y) = a_\nu(i) T(\gamma_{j,k})(y).$$

By Proposition 4.1, for any $\gamma \in \{\phi, \psi\}$, we have

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n q_i(\gamma_{j,k}, Y_i)\right) = \int_0^1 f_\nu(x)\gamma_{j,k}(x)dx.$$

Proceeding as in (6.3), we obtain

$$\frac{1}{n^2}\sum_{i=1}^n \mathbb{E}\left((q_i(\gamma_{j,k}, Y_i))^2\right) \leq \theta_\gamma^2 2^{2j} \frac{1}{n^2}\sum_{i=1}^n a_\nu^2(i) = \theta_\gamma^2 2^{2j} \frac{\rho_n}{n},$$

with

$$\theta_\gamma = \sqrt{2C_*\left(1 + \int_{1-N}^N (\gamma'(y))^2 dy\right)}. \tag{6.12}$$

Thus the assumptions (A1) and (A2) in Theorem 6.1 are satisfied with, for any $i \in \{1, \ldots, n\}$,

$$W_i = Y_i = U_i X_i, \qquad q_i(\gamma_{j,k}, y) = a_\nu(i) T(\gamma_{j,k})(y),$$

$v_n = n$, $\omega = 1$, $\mu_n = n/\rho_n$ and θ_γ is (6.12).

This implies the following result:

Let $f_\nu \in B^s_{p,r}(M)$ with $r \geq 1$, $\{p \geq 2$ and $s \in (0, N)\}$ or $\{p \in [1, 2)$ and $s \in (3/p, N)\}$, then the hard thresholding estimator $\widehat{f}^h_\nu$ (4.6) satisfies

$$\mathbb{E}\left(\|\widehat{f}^h_\nu - f_\nu\|^2_2 \right) \leq C \left(\frac{\rho_n \ln(n/\rho_n)}{n} \right)^{2s/(2s+3)}.$$

This completes the proof of Theorem 5.2. □

6.3. Proof of Theorem 6.1

The proof of the theorem is broken into two major parts. In the first part, we present two auxiliary results on the statistical properties of (6.7) and (6.8). Then, following the idea of Härdle *et al.* (1998), Proof of Proposition 10.3, these results are used in the second part on a suitable decomposition of the MISE of $\widehat{f}_H$.

Part I: Auxiliary results

Proposition 6.1. *Consider $\widehat{\beta}_{j,k}$ as defined in (6.8). We have*

$$|\widehat{\beta}_{j,k} - \beta_{j,k}| \leq \frac{1}{v_n} \left| \sum_{i=1}^n U_i \right| + \vartheta_j,$$

where

$$U_i = q_i(\psi_{j,k}, W_i)\mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)| \leq \varsigma_j\}} - \mathbb{E}(q_i(\psi_{j,k}, W_i)\mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)| \leq \varsigma_j\}}). \quad (6.13)$$

Proof of Proposition 6.1. By (A1), we have

$$\begin{aligned}
\beta_{j,k} &= \mathbb{E}\left(\frac{1}{v_n} \sum_{i=1}^n q_i(\psi_{j,k}, W_i) \right) \\
&= \frac{1}{v_n} \sum_{i=1}^n \mathbb{E}(q_i(\psi_{j,k}, W_i)\mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)| \leq \varsigma_j\}}) \\
&\quad + \frac{1}{v_n} \sum_{i=1}^n \mathbb{E}(q_i(\psi_{j,k}, W_i)\mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)| > \varsigma_j\}}).
\end{aligned}$$

Hence

$$\begin{aligned}
&|\widehat{\beta}_{j,k} - \beta_{j,k}| \\
&\leq \frac{1}{v_n} \left| \sum_{i=1}^n U_i \right| + \frac{1}{v_n} \sum_{i=1}^n \mathbb{E}\left(|q_i(\psi_{j,k}, W_i)|\mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)| > \varsigma_j\}} \right),
\end{aligned}$$

where $U_1, \ldots, U_n$ are given by (6.13).

The Markov inequality and (A2) imply that

$$\frac{1}{\upsilon_n} \sum_{i=1}^{n} \mathbb{E}\left(|q_i(\psi_{j,k}, W_i)| \mathbf{1}_{\{|q_i(\psi_{j,k}, W_i)| > \varsigma_j\}}\right)$$

$$\leq \frac{1}{\varsigma_j \upsilon_n} \sum_{i=1}^{n} \mathbb{E}\left((q_i(\psi_{j,k}, W_i))^2\right) \leq \frac{1}{\varsigma_j} \theta_\psi^2 2^{2\omega j} \frac{\upsilon_n}{\mu_n}$$

Using the definitions of ς_j and ϑ_j from (6.9) and (6.10) respectively, the right hand side of the above inequality becomes

$$\frac{1}{\theta_\psi 2^{\omega j} \frac{\upsilon_n}{\sqrt{\mu_n \ln \mu_n}}} \theta_\psi^2 2^{2\omega j} \frac{\upsilon_n}{\mu_n} = \theta_\psi 2^{\omega j} \sqrt{\frac{\ln \mu_n}{\mu_n}} = \vartheta_j$$

that completes the proof of Proposition 6.1.

Proposition 6.2. *Consider $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$ as defined in (6.7) and (6.8), respectively.*

(i) *There exists a constant $C > 0$ such that*

$$\mathbb{E}\left((\widehat{\alpha}_{j,k} - \alpha_{j,k})^2\right) \leq C 2^{2\omega j} \frac{1}{\mu_n}.$$

(ii) *There exists a constant $C > 0$ such that*

$$\mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k}\right)^4\right) \leq C 2^{4\omega j} \left(\frac{\ln \mu_n}{\mu_n}\right)^2.$$

(iii) *For any $\kappa \geq 2 + 8/3 + 2\sqrt{4 + 16/9}$, we have*

$$\mathbb{P}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \kappa \vartheta_j/2\right) \leq 2 \frac{1}{\mu_n^2}. \tag{6.14}$$

Proof of Proposition 6.2.

(i) Noting that $\widehat{\alpha}_{j,k}$ is an unbiased estimator of $\alpha_{j,k}$, by (A1), the independence of $W_1, \ldots, W_n$ and (A2) yield

$$\mathbb{E}\left((\widehat{\alpha}_{j,k} - \alpha_{j,k})^2\right) = \mathbb{V}(\widehat{\alpha}_{j,k}) = \frac{1}{\upsilon_n^2} \sum_{i=1}^{n} \mathbb{V}(q_i(\phi_{j,k}, W_i))$$

$$\leq \frac{1}{\upsilon_n^2} \sum_{i=1}^{n} \mathbb{E}\left((q_i(\phi_{j,k}, W_i))^2\right)$$

$$\leq \theta_\phi^2 2^{2\omega j} \frac{1}{\mu_n}.$$

(ii) Using Proposition 6.1 we have

$$\mathbb{E}\left(\left(\widehat{\beta}_{j,k}-\beta_{j,k}\right)^4\right) \leq C\left(\frac{1}{\upsilon_n^4}\mathbb{E}\left(\left(\sum_{i=1}^{n}U_i\right)^4\right)+\vartheta_j^4\right). \qquad (6.15)$$

where $U_1,\ldots,U_n$ are given by (6.13).
Note that $\mathbb{E}(U_i)=0$ and

$$\mathbb{E}\left((U_i)^4\right) \leq 16\varsigma_j^2\mathbb{E}\left((q_i(\psi_{j,k},W_i))^2\right).$$

Hence using the Rosenthal inequality (see the Appendix), (A2) and the definition of ς_j from (6.9), we have

$$\mathbb{E}\left(\left(\sum_{i=1}^{n}U_i\right)^4\right) \leq C\left(\sum_{i=1}^{n}\mathbb{E}\left(U_i^4\right)+\left(\sum_{i=1}^{n}\mathbb{E}\left(U_i^2\right)\right)^2\right)$$

$$\leq C\left(\varsigma_j^2\sum_{i=1}^{n}\mathbb{E}\left((q_i(\psi_{j,k},W_i))^2\right)+\left(\sum_{i=1}^{n}\mathbb{E}\left((q_i(\psi_{j,k},W_i))^2\right)\right)^2\right)$$

$$\leq C\left(\varsigma_j^2\theta_\psi^2 2^{2\omega j}\frac{\upsilon_n^2}{\mu_n}+\theta_\psi^4 2^{4\omega j}\frac{\upsilon_n^4}{\mu_n^2}\right)$$

$$\leq C\left(2^{4\omega j}\frac{\upsilon_n^4}{\mu_n^2\ln\mu_n}+2^{4\omega j}\frac{\upsilon_n^4}{\mu_n^2}\right) \leq C2^{4\omega j}\frac{\upsilon_n^4}{\mu_n^2}. \qquad (6.16)$$

It follows from (6.15), (6.16) and the definition of ϑ_j from (6.10) that

$$\mathbb{E}\left(\left(\widehat{\beta}_{j,k}-\beta_{j,k}\right)^4\right) \leq C2^{4\omega j}\left(\frac{\ln\mu_n}{\mu_n}\right)^2.$$

(iii) Proposition 6.1 yields

$$\mathbb{P}\left(|\widehat{\beta}_{j,k}-\beta_{j,k}|\geq\kappa\vartheta_j/2\right) \leq \mathbb{P}\left(\left|\sum_{i=1}^{n}U_i\right|\geq\kappa_*\upsilon_n\vartheta_j\right), \qquad (6.17)$$

where $U_1,\ldots,U_n$ are given by (6.13) and $\kappa_*=\kappa/2-1$.
Observe that $\mathbb{E}(U_1)=0$, $|U_1|\leq 2\varsigma_j$ and, by (A2),

$$\sum_{i=1}^{n}\mathbb{E}\left(U_i^2\right) \leq \sum_{i=1}^{n}\mathbb{E}\left((q_i(\psi_{j,k},W_i))^2\right) \leq \theta_\psi^2 2^{2\omega j}\frac{\upsilon_n^2}{\mu_n}.$$

Hence, using the Bernstein inequality (see the Appendix) and the defini-

tions of ϑ_j and ς_j from (6.9) (6.10), respectively, we have

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} U_i\right| \geq \kappa_* \upsilon_n \vartheta_j\right)$$

$$\leq 2\exp\left(-\frac{\kappa_*^2 \upsilon_n^2 \vartheta_j^2}{2\left(\sum_{i=1}^{n} \mathbb{E}\left(U_i^2\right) + \frac{2}{3}\kappa_* \upsilon_n \vartheta_j \varsigma_j\right)}\right)$$

$$\leq 2\exp\left(-\frac{\kappa_*^2 \upsilon_n^2 \theta_\psi^2 2^{2\omega j}\frac{\ln \mu_n}{\mu_n}}{2\left(\theta_\psi^2 2^{2\omega j}\frac{\upsilon_n^2}{\mu_n} + \frac{2}{3}\kappa_* \theta_\psi^2 2^{2\omega j}\frac{\upsilon_n^2}{\mu_n}\right)}\right)$$

$$= 2\mu_n^{-\frac{\kappa_*^2}{2\left(1+\frac{2}{3}\kappa_*\right)}}.$$

Taking κ such that $\kappa_*^2/\left(2\left(1+(2/3)\kappa_*\right)\right) \geq 2$, i.e. $\kappa \geq 2+8/3+2\sqrt{4+16/9}$, we obtain

$$\mathbb{P}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \kappa\vartheta_j/2\right) \leq 2\frac{1}{\mu_n^2}.$$

This completes the proof of Proposition 6.2. $\qquad\square$

Part II: Proof of the theorem

We decompose f on the wavelet basis $\mathcal{B}$ as

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

where $\alpha_{\tau,k} = \int_0^1 f(x)\phi_{\tau,k}(x)dx$ and $\beta_{j,k} = \int_0^1 f(x)\psi_{j,k}(x)dx$. Using the orthonormality of the wavelet basis $\mathcal{B}$, we can decompose the MISE of $\widehat{f}_H$ in three terms as:

$$\mathbb{E}\left(\|\widehat{f}_H - f\|_2^2\right) = E + F + G, \tag{6.18}$$

where

$$E = \sum_{k=0}^{2^\tau - 1} \mathbb{E}\left((\widehat{\alpha}_{\tau,k} - \alpha_{\tau,k})^2\right), \qquad F = \sum_{j=\tau}^{j_1}\sum_{k=0}^{2^j - 1} \mathbb{E}\left(\left(\widehat{\beta}_{j,k}\mathbf{1}_{\left\{|\widehat{\beta}_{j,k}|\geq \kappa\vartheta_j\right\}} - \beta_{j,k}\right)^2\right)$$

and

$$G = \sum_{j=j_1+1}^{\infty}\sum_{k=0}^{2^j - 1} \beta_{j,k}^2.$$

We now exhibit the bounds of each of these terms.
Bound for E :

Proposition 6.2 gives

$$E \leq C \frac{1}{\mu_n} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}. \tag{6.19}$$

Bound for F:

For the sake of simplicity, set $\hat{B}_{j,k} = \widehat{\beta}_{j,k} - \beta_{j,k}$. The term F can be decomposed as

$$F = F_1 + F_2 + F_3 + F_4, \tag{6.20}$$

where

$$F_1 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E} \left(\hat{B}_{j,k}^2 \mathbf{1}_{\left\{ |\widehat{\beta}_{j,k}| \geq \kappa \vartheta_j, \ |\beta_{j,k}| < \kappa \vartheta_j/2 \right\}} \right),$$

$$F_2 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E} \left(\hat{B}_{j,k}^2 \mathbf{1}_{\left\{ |\widehat{\beta}_{j,k}| \geq \kappa \vartheta_j, \ |\beta_{j,k}| \geq \kappa \vartheta_j/2 \right\}} \right),$$

$$F_3 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E} \left(\beta_{j,k}^2 \mathbf{1}_{\left\{ |\widehat{\beta}_{j,k}| < \kappa \vartheta_j, \ |\beta_{j,k}| \geq 2\kappa \vartheta_j \right\}} \right)$$

and

$$F_4 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E} \left(\beta_{j,k}^2 \mathbf{1}_{\left\{ |\widehat{\beta}_{j,k}| < \kappa \vartheta_j, \ |\beta_{j,k}| < 2\kappa \vartheta_j \right\}} \right).$$

Bound for $F_1 + F_3$:

Observe that

$$\mathbf{1}_{\left\{ |\widehat{\beta}_{j,k}| < \kappa \vartheta_j, \ |\beta_{j,k}| \geq 2\kappa \vartheta_j \right\}} \leq \mathbf{1}_{\left\{ |\hat{B}_{j,k}| > \kappa \vartheta_j/2 \right\}},$$

$$\mathbf{1}_{\left\{ |\widehat{\beta}_{j,k}| \geq \kappa \vartheta_j, \ |\beta_{j,k}| < \kappa \vartheta_j/2 \right\}} \leq \mathbf{1}_{\left\{ |\hat{B}_{j,k}| > \kappa \vartheta_j/2 \right\}}$$

and

$$\mathbf{1}_{\left\{ |\widehat{\beta}_{j,k}| < \kappa \vartheta_j, \ |\beta_{j,k}| \geq 2\kappa \vartheta_j \right\}} \leq \mathbf{1}_{\left\{ |\beta_{j,k}| \leq 2|\hat{B}_{j,k}| \right\}}.$$

Owing to the previous inequalities, the Cauchy-Schwarz inequality and Proposition 6.2, we obtain

$$F_1 + F_3 \leq C \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E} \left(\hat{B}_{j,k}^2 \mathbf{1}_{\left\{ |\hat{B}_{j,k}| > \kappa \vartheta_j/2 \right\}} \right)$$

$$\leq C \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \left(\mathbb{E} \left(\hat{B}_{j,k}^4 \right) \right)^{1/2} \left(\mathbb{P} \left(|\hat{B}_{j,k}| > \kappa \vartheta_j/2 \right) \right)^{1/2}$$

$$\leq C \frac{\ln \mu_n}{\mu_n^2} \sum_{j=\tau}^{j_1} 2^{j(1+2\omega)} \leq C \frac{\ln \mu_n}{\mu_n} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}. \tag{6.21}$$

Bound for F_2 :

Using the Cauchy-Schwarz inequality and Proposition 6.2, we obtain

$$
\begin{aligned}
F_2 &\leq \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \left(\mathbb{E}\left(\hat{B}_{j,k}^4 \right) \right)^{1/2} \mathbf{1}_{\{|\beta_{j,k}| \geq \kappa \vartheta_j/2\}} \\
&\leq C \frac{\ln \mu_n}{\mu_n} \sum_{j=\tau}^{j_1} 2^{2\omega j} \sum_{k=0}^{2^j-1} \mathbf{1}_{\{|\beta_{j,k}| > \kappa \vartheta_j/2\}}.
\end{aligned}
$$

Let j_* be the integer defined by

$$
\frac{1}{2} \left(\frac{\mu_n}{\ln \mu_n} \right)^{1/(2s+2\omega+1)} < 2^{j_*} \leq \left(\frac{\mu_n}{\ln \mu_n} \right)^{1/(2s+2\omega+1)}. \tag{6.22}
$$

Then F_2 can be bounded as

$$
F_2 \leq F_{2,1} + F_{2,2},
$$

where

$$
F_{2,1} = C \frac{\ln \mu_n}{\mu_n} \sum_{j=\tau}^{j_*} 2^{2\omega j} \sum_{k=0}^{2^j-1} \mathbf{1}_{\{|\beta_{j,k}| > \kappa \vartheta_j/2\}}
$$

and

$$
F_{2,2} = C \frac{\ln \mu_n}{\mu_n} \sum_{j=j_*+1}^{j_1} 2^{2\omega j} \sum_{k=0}^{2^j-1} \mathbf{1}_{\{|\beta_{j,k}| > \kappa \vartheta_j/2\}}.
$$

We have

$$
F_{2,1} \leq C \frac{\ln \mu_n}{\mu_n} \sum_{j=\tau}^{j_*} 2^{j(1+2\omega)} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
$$

For $r \geq 1$ and $p \geq 2$, the Markov inequality, definition of ϑ_j (from (6.10)) and $f \in B_{p,r}^s(M) \subseteq B_{2,\infty}^s(M)$ (see Härdle *et al.* (1998), Corollary 9.2) yield

$$
\begin{aligned}
F_{2,2} &\leq C \frac{\ln \mu_n}{\mu_n} \sum_{j=j_*+1}^{j_1} 2^{2\omega j} \frac{1}{\vartheta_j^2} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \leq C \sum_{j=j_*+1}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \\
&\leq C \sum_{j=j_*+1}^{\infty} 2^{-2js} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
\end{aligned}
$$

For $r \geq 1$, $p \in [1, 2)$ and $s > (2\omega+1)/p$, the Markov inequality, the definition of ϑ_j, $f \in B_{p,r}^s(M)$ and $(2s+2\omega+1)(2-p)/2 + (s+1/2-1/p+\omega-2\omega/p)p = 2s$

imply that

$$
\begin{aligned}
F_{2,2} &\leq C \frac{\ln \mu_n}{\mu_n} \sum_{j=j_*+1}^{j_1} 2^{2\omega j} \frac{1}{\vartheta_j^p} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \\
&\leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{(2-p)/2} \sum_{j=j_*+1}^{\infty} 2^{j\omega(2-p)} 2^{-j(s+1/2-1/p)p} \\
&\leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{(2-p)/2} 2^{-j_*(s+1/2-1/p+\omega-2\omega/p)p} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
\end{aligned}
$$

So, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1,2)$ and $s > (2\omega+1)/p\}$, we have

$$
F_2 \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}. \tag{6.23}
$$

Bound for F_4 :
We have

$$
F_4 \leq \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \mathbf{1}_{\{|\beta_{j,k}|<2\kappa\vartheta_j\}}.
$$

Let j_* be the integer (6.22). Then F_4 can be bound as

$$
F_4 \leq F_{4,1} + F_{4,2},
$$

where

$$
F_{4,1} = \sum_{j=\tau}^{j_*} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \mathbf{1}_{\{|\beta_{j,k}|<2\kappa\vartheta_j\}}, \qquad F_{4,2} = \sum_{j=j_*+1}^{j_1} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \mathbf{1}_{\{|\beta_{j,k}|<2\kappa\vartheta_j\}}.
$$

Observe that, using the definition of ϑ_j, we have

$$
F_{4,1} \leq C \sum_{j=\tau}^{j_*} 2^j \vartheta_j^2 = C \frac{\ln \mu_n}{\mu_n} \sum_{j=\tau}^{j_*} 2^{j(1+2\omega)} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
$$

Further, for $r \geq 1$ and $p \geq 2$, since $f \in B_{p,r}^s(M) \subseteq B_{2,\infty}^s(M)$ (see Härdle *et al.* (1998), Corollary 9.2), we have

$$
F_{4,2} \leq \sum_{j=j_*+1}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \leq C \sum_{j=j_*+1}^{\infty} 2^{-2js} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
$$

For $r \geq 1$, $p \in [1,2)$ and $s > (2\omega+1)/p$, owing to the Markov inequality, the definition of ϑ_j (from (6.10)), $f \in B_{p,r}^s(M)$ and $(2s+2\omega+1)(2-p)/2 + (s+$

$1/2 - 1/p + \omega - 2\omega/p)p = 2s$, we obtain

$$
\begin{aligned}
F_{4,2} &\leq C \sum_{j=j_*+1}^{j_1} \vartheta_j^{2-p} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \\
&= C \left(\frac{\ln \mu_n}{\mu_n} \right)^{(2-p)/2} \sum_{j=j_*+1}^{j_1} 2^{j\omega(2-p)} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \\
&\leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{(2-p)/2} \sum_{j=j_*+1}^{\infty} 2^{j\omega(2-p)} 2^{-j(s+1/2-1/p)p} \\
&\leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{(2-p)/2} 2^{-j_*(s+1/2-1/p+\omega-2\omega/p)p} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
\end{aligned}
$$

Thus, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1,2)$ and $s > (2\omega+1)/p\}$, we have

$$
F_4 \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}. \tag{6.24}
$$

Therefore it follows from (6.20), (6.21), (6.23) and (6.24) that

$$
F \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}. \tag{6.25}
$$

Bound for G:

For $r \geq 1$ and $p \geq 2$, we have $f \in B_{p,r}^s(M) \subseteq B_{2,\infty}^s(M)$. Since $2s/(2s+2\omega+1) < 2s/(2\omega+1)$, we have

$$
G \leq C \sum_{j=j_1+1}^{\infty} 2^{-2js} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2\omega+1)} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
$$

For $r \geq 1$ and $p \in [1,2)$, we have $f \in B_{p,r}^s(M) \subseteq B_{2,\infty}^{s+1/2-1/p}(M)$ (see Härdle *et al.* (1998), Corollary 9.2). Since $s > (2\omega+1)/p$, we have $(s+1/2-1/p)/(2\omega+1) > s/(2s+2\omega+1)$. Hence

$$
G \leq C \sum_{j=j_1+1}^{\infty} 2^{-2j(s+1/2-1/p)} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2(s+1/2-1/p)/(2\omega+1)} \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
$$

Hence, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1,2)$ and $s > (2\omega+1)/p\}$, we have

$$
G \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}. \tag{6.26}
$$

Combining (6.18), (6.19), (6.25) and (6.26), we have, for any $r \geq 1$, $\{p \geq 2$ and $s \in (0, N)\}$ or $\{p \in [1,2)$ and $s \in ((2\omega+1)/p, N)\}$,

$$
\mathbb{E} \left(\|\widehat{f}_H - f\|_2^2 \right) \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\omega+1)}.
$$

This completes the proof of Theorem 6.1. □

Appendix

Lemma 6.1 (Rosenthal's inequality). *Let $p \geq 2$, $n \in \mathbb{N}^*$ and $(A_i)_{i \in \{1,\ldots,n\}}$ be n zero mean independent random variables such that, for any $i \in \{1,\ldots,n\}$, $\mathbb{E}(|A_i|^p) < \infty$. Then there exists a constant $C > 0$ such that*

$$\mathbb{E}\left(\left|\sum_{i=1}^n A_i\right|^p\right) \leq C\left(\sum_{v=1}^n \mathbb{E}\left(|A_i|^p\right) + \left(\sum_{i=1}^n \mathbb{E}\left(A_i^2\right)\right)^{p/2}\right).$$

See Rosenthal (1970) for details.

Lemma 6.2 (Bernstein's inequality). *Let $n \in \mathbb{N}^*$ and $(A_i)_{i \in \{1,\ldots,n\}}$ be n zero mean independent random variables such that there exists a constant $M > 0$ satisfying, for any $i \in \{1,\ldots,n\}$, $|A_i| \leq M < \infty$. Then, for any $y > 0$, we have*

$$\mathbb{P}\left(\left|\sum_{i=1}^n A_i\right| \geq y\right) \leq 2\exp\left(-\frac{y^2}{2\left(\sum_{i=1}^n \mathbb{E}\left(A_i^2\right) + \frac{yM}{3}\right)}\right).$$

See Petrov (1995) for details.

Acknowledgements

References

ABBASZADEH, M., CHESNEAU, C. AND DOOSTI, H. (2012). Nonparametric estimation of a density under bias and multiplicative censoring via wavelet methods. *Statistics and Probability Letters* **82** 932-941.

ASGHARIAN, M., CARONE, M. AND FAKOOR, V. (2012). Large-sample study of the kernel density estimators under multiplicative censoring. *Annals of Statistics* **40** 159-187.

AKYILDIZ, I.F., SU, W., SANKARASUBRAMANIAM, Y. AND CAYIRCI, E. (2002). Wireless sensor networks: a survey. *Computer Networks* **38** 393-422.

ANDERSEN, K. AND HANSEN, M. (2001). Multiplicative censoring: density estimation by a series expansion approach. *Journal of Statistical Planning and Inference* **98** 137-155.

ANTONIADIS, A. (1997). Wavelets in statistics: a review (with discussion). *Journal of the Italian Statistical Society* **6** 97-144.

AUTIN, F. AND POUET, C. (2011). Test on the components of mixture densities. *Stat. & Risk Modeling* **28** 389-410.

CAI, T. (1999). Adaptive wavelet estimation: a block thresholding and oracle inequality approach. *The Annals of Statistics* **27** 898-924.

CAI, T. (2002). On block thresholding in wavelet regression: adaptivity, block size and threshold level. *Statistica Sinica* **12** 1241-1273.

CHABERT, M., RUIZ, D. AND TOURNERET, J.-Y. (2004). Optimal wavelet for abrupt change detection in multiplicative noise. In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004) **2** 1089-1092.

CHAUBEY, Y.P., CHESNEAU, C. AND DOOSTI, H. (2011). Wavelet Linear Density Estimation: A review. *Journal of the Indian Society of Agricultural Statistics* **65** 169-179.

CHESNEAU, C. (2011a). Wavelet density estimators for the deconvolution of a component from a mixture, *Sankhyā* **73-A** 245-266.

CHESNEAU, C. (2011b). Adaptive wavelet estimator for a function and its derivatives in an indirect convolution model, *Journal of Statistical Theory and Practice* **5** 303-326.

CHESNEAU, C. AND FADILI, J. (2012). Adaptive wavelet estimation of a function in an indirect regression model. *Advances in Statistical Analysis*, **96**, 1, 25-46.

COHEN, A., DAUBECHIES, I., JAWERTH, B. AND VIAL, P. (1993). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis* **24** 54–81.

DELYON, B. AND JUDITSKY, A. (1996). On minimax wavelet estimators, *Applied Computational Harmonic Analysis.* **3** 215–228.

DONOHO, D. L. AND JOHNSTONE, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81** 425-455.

DONOHO, D. L. AND JOHNSTONE, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association* **90** 1200-1224.

DONOHO, D.L., JOHNSTONE, I.M., KERKYACHARIAN, G. AND PICARD, D. (1996). Density estimation by wavelet thresholding. *The Annals of Statistics* **24** 508-539.

HALL, P. AND ZHOU, X.H. (2003). Nonparametric estimation of component distributions in a multivariate mixture. *The Annals of Statistics* **31** 201-224.

HÄRDLE, W., KERKYACHARIAN, G., PICARD, D. AND TSYBAKOV, A. (1998). *Wavelet, Approximation and Statistical Applications.* Lectures Notes in Statistics, 129, *Springer Verlag,* New York.

LEPSKI, O. (1990). On a problem of adaptive estimation in Gaussian white noise. *Th. Prob.* **35** 454-466.

LEPSKI, O. (1991). Asymptotically minimax adaptive estimation I : upper bounds. Optimally adaptive estimates. *Th. Prob. Appl.* **36** 682-697.

LEPSKI, O. (1992). Asymptotically minimax adaptive estimation II : statistical models without optimal adaptation. Adaptive estimators. *Th. Prob. Appl.* **37** 433-448.

LEPSKI, O. AND SPOKOINY V.G. (1995). Local adaptation to inhomogeneous smoothness : resolution level. *Math. Methods Statist.* **4** 239-258

MAIBORODA, R. E. (1996). Estimators of components of a mixture with varying concentrations. *Ukrain. Mat. Zh.* **48** 562-566.

MALLAT, S. (2009). *A wavelet tour of signal processing.* Elsevier/ Academic Press, Amsterdam, third edition. The sparse way, With contributions from Gabriel Peyré.

MEYER, Y. (1990). *Ondelettes et Opérateurs.* Hermann, Paris.

PETROV, V. V. (1995). *Limit Theorems of Probability Theory: Sequences of Independent Random Variables.* Oxford: Clarendon Press.

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. (1996). Nonparametric estimation of the derivatives of a density by the method of wavelets. *Bull. Inform. Cyb.* **28** 91-100.

PRAKASA RAO, B. L. S. (1999). Nonparametric functional estimation: an overview. In *Asymptotics, Nonparametrics and Time Series*, Ed.: Subir Ghosh, 461-509, Marcel Dekker Inc., New York.

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** 275-291.

ROSENTHAL, H. P. (1970). On the subspaces of $\mathbb{L}^p$ ($p \geq 2$) spanned by sequences of independent random variables. *Israel Journal of Mathematics* **20** 273-303.

TSYBAKOV, A.B. (2004). *Introduction à l'estimation non-paramétrique.* Springer, New York.

VARDI, Y. (1989). Multiplicative censoring, renewal processes, deconvolution and decreasing density: Nonparametric estimation. *Biometrika* **76** 751-761.

VARDI, Y. AND ZHANG, C. H. (1992). Large sample study of empirical distributions in a random multiplicative censoring model. *The Annals of Statistics* **20** 1022-1039.