Wavelet density estimators for the deconvolution of a component from a mixture

Christophe Chesneau

Noname manuscript No.
(will be inserted by the editor)

Wavelet density estimators for the deconvolution of a component from a mixture

Christophe Chesneau

Received:

Abstract We consider the model: $Y = X + \epsilon$, where X and ϵ are independent random variables. The density of ϵ is known whereas the one of X is a finite mixture with unknown components. Considering the "ordinary smooth case" on the density of ϵ, we want to estimate a component of this mixture. To reach this goal, we develop two wavelet estimators: a nonadaptive based on a projection and an adaptive based on a hard thresholding rule. We evaluate their performances by taking the minimax approach under the mean integrated squared error over Besov balls. We prove that the adaptive one attains a sharp rate of convergence.

Keywords Density deconvolution · Mixture · Minimax approach · Wavelets · Hard thresholding.

1 Motivations

We consider the following model:

$$Y_v = X_v + \epsilon_v, \qquad v \in \{1, \ldots, n\}, \quad n \in \mathbb{N}^*, \tag{1}$$

where $X_1, \ldots, X_n$ are independent random variables and $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. random variables. The density of ϵ_1 is denoted g and, for any $v \in \{1, \ldots, n\}$, the density of X_v is the finite mixture:

$$h_v(x) = \sum_{d=1}^{m} w_d(v) f_d(x), \qquad x \in \mathbb{R},$$

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

Laboratoire de Mathématiques Nicolas Oresme, Université de Caen Basse-Normandie, Campus II, Science 3, 14032 Caen, France. E-mail: chesneau@math.unicaen.fr

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

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

– $(f_d)_{d\in\{1,\ldots,m\}}$ are unknown densities.

For a fixed $d_* \in \{1,\ldots,m\}$, we aim to estimate f_{d_*} when only $Y_1,\ldots,Y_n$ are observed.

In the literature, the model (1) has been recently described for a particular mixture in Lee et al. (2010). In the simplest case where $m = 1$, $w_1(1) = \ldots = w_1(n) = 1$ and $f_{d_*} = f_1 = f$, (1) becomes the standard convolution density model. See e.g. Hall and Carol (1988), Devroye (1989), Fan (1991), Pensky and Vidakovic (1999), Fan and Koo (2002), Butucea and Matias (2005), Comte et al. (2006) and Lacour (2006). The estimation of f_{d_*} when only $X_1,\ldots,X_n$ are observed has been investigated in some 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_{d_*} from $Y_1,\ldots,Y_n$ is a new challenge.

Considering the ordinary smooth case on g (see (3)), we estimate f_{d_*} by two wavelet estimators: a linear nonadaptive and a nonlinear adaptive based on the hard thresholding rule. The construction of our adaptive estimator is "similar" to the one of Pensky and Vidakovic (1999) and Fan and Koo (2002). It has the originality to include some technical tools on mixture and a new version of the "observations thresholding" introduced in wavelet estimation theory by Delyon and Juditsky (1996) in the context of the nonparametric regression. The performances of our estimators are evaluated via the minimax approach under the mean integrated squared error (MISE) over a wide class of functions: the Besov balls $B^s_{p,r}(M)$ (to be defined in Section 3). In particular, under mild assumptions on the weights of the mixture, we prove that our adaptive estimator attains the rate of convergence:

$$r_n = \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)},$$

where z_n depends on these weights and δ is a factor related to the ordinary smooth case. This rate of convergence is sharp in the sense that it is the one attains by the best nonrealistic linear wavelet estimator up to a logarithmic term.

The paper is organized as follows. Assumptions on the model and some notations are introduced in Section 2. Section 3 briefly describes the wavelet basis and the Besov balls. The estimators are presented in Section 4. The minimax results are set in Section 5. Technical proofs are given in Section 6.

2 Assumptions and notations

Assumption on f_{d_*}. Without loss of generality, we assume that the support of f_{d_*} is $[0,1]$.

Assumptions on g. We suppose that there exists a constant $C_* > 0$ such that

$$\sup_{x\in\mathbb{R}} g(x) \leq C_* < \infty. \tag{2}$$

We define the Fourier transform of a function h by

$$\mathcal{F}(h)(x) = \int_{-\infty}^{\infty} h(y)e^{-ixy}dy, \qquad x \in \mathbb{R},$$

whenever this integral exists. The notation $\overline{\cdot}$ will be used for the complex conjugate.

We consider the ordinary smooth case on g: there exist two constants, $c_* > 0$ and $\delta > 1$, such that, for any $x \in \mathbb{R}$, $\mathcal{F}(g)(x)$ satisfies

$$|\,\mathcal{F}(g)(x)| \geq \frac{c_*}{(1+x^2)^{\delta/2}}. \tag{3}$$

This assumption controls the decay of the Fourier coefficients of g, and thus the smoothness of g.

Example: for any $v \in \{1,\dots,n\}$, suppose that $\epsilon_v = \sum_{u=1}^{p} \varepsilon_{u,v}$, where $p \in \mathbb{N}^*$ and $(\varepsilon_{u,v})_{(u,v)\in\{1,\dots,p\}\times\{1,\dots,n\}}$ are i.i.d. random variables having the Laplace density: $f(x) = (1/2)e^{-|x|}$, $x \in \mathbb{R}$. Then $|\,\mathcal{F}(g)(x)| = 1/(1+x^2)^p$. Therefore (3) is satisfied with $c_* = 1$ and $\delta = 2p$.

Assumptions on the weights. We suppose that the matrix

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

is nonsingular i.e. $\det(\Gamma_n) > 0$. For the considered d_* and any $v \in \{1,\dots,n\}$, we set

$$a_{d_*}(v) = \frac{1}{\det(\Gamma_n)}\sum_{k=1}^{m}(-1)^{k+d_*}\gamma^n_{d_*,k}w_k(v), \tag{4}$$

where $\gamma^n_{d_*,k}$ denotes the determinant of the minor (d_*,k) of the matrix Γ_n. Then, for any $k \in \{1,\dots,m\}$,

$$\frac{1}{n}\sum_{v=1}^{n} a_{d_*}(v)w_k(v) = \begin{cases} 1 & \text{if } k = d_*, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

and

$$(a_{d_*}(1),\dots,a_{d_*}(n)) = \underset{(b_1,\dots,b_n)\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{n}\sum_{v=1}^{n} b_v^2.$$

Technical details can be found in Maiboroda (1996).

We set

$$z_n = \frac{1}{n}\sum_{v=1}^{n} a_{d_*}^2(v) \tag{6}$$

and, for technical reasons, we suppose that $z_n < n/e$.

3 Wavelets and Besov balls

Wavelet basis. Let $N \in \mathbb{N}^*$, ϕ be a father wavelet of a multiresolution analysis on $\mathbb{R}$ and ψ be the associated mother wavelet. Assume that
- $\text{supp}(\phi) = \text{supp}(\psi) = [1-N, N]$,
- $\int_{1-N}^{N} \phi(x)dx = 1$,
- for any $v \in \{0, \ldots, N-1\}$, $\int_{1-N}^{N} x^v \psi(x)dx = 0$,
- ϕ and ψ are of class $\mathcal{C}^v$, $v > 1+\delta$, where δ is the one in (3).

(For instance, the Daubechies wavelets). 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 an integer τ satisfying $2^\tau \geq 2N$ such that the collection

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

(with an appropriate treatments at the boundaries) is an orthonormal basis of $\mathbb{L}^2([0,1]) = \{h : [0,1] \to \mathbb{R};\ \int_0^1 h^2(x)dx < \infty\}$. We refer to Cohen et al. (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

$$\alpha_{j,k} = \int_0^1 h(x)\phi_{j,k}(x)dx, \qquad \beta_{j,k} = \int_0^1 h(x)\psi_{j,k}(x)dx. \tag{7}$$

Besov balls. Let $M > 0$, $s > 0$, $p \geq 1$ and $r \geq 1$. A function h belongs to the Besov balls $B_{p,r}^s(M)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients (7) 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 (1992).

4 Estimators

Wavelet coefficients estimators. The first step to estimate f_{d_*} consists in expanding f_{d_*} on $\mathcal{B}$ and estimating its unknown wavelet coefficients.
For any integer $j \geq \tau$ and any $k \in \{0, \ldots, 2^j - 1\}$, we estimate $\alpha_{j,k} = \int_0^1 f_{d_*}(x)\overline{\phi}_{j,k}(x)dx$ by

$$\widehat{\alpha}_{j,k} = \frac{1}{2\pi n}\sum_{v=1}^{n} a_{d_*}(v)\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixY_v}dx, \tag{8}$$

and $\beta_{j,k} = \int_0^1 f_{d_*}(x)\overline{\psi}_{j,k}(x)dx$ by

$$\widehat{\beta}_{j,k} = \frac{1}{n}\sum_{v=1}^{n} G_v 1_{\{|G_v|\leq \eta_j\}}, \tag{9}$$

where

$$G_v = \frac{1}{2\pi} a_{d_*}(v)\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\psi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixY_v}dx, \tag{10}$$

for any random event $\mathcal{A}$, $1_{\mathcal{A}}$ is the indicator function on $\mathcal{A}$, $a_{d_*}(v)$ is defined by (4),

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

z_n is defined by (6) and $\theta = \sqrt{(C_*/2\pi c_*^2)\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}(\psi)(x)\right|^2 dx}$ (C_*, c_* and δ are those in (2) and (3)).
Note that, since $\psi \in \mathcal{C}^{\upsilon}$, there exists a constant $C > 0$ such that $|\mathcal{F}(\psi)(x)| \leq C(1+|x|)^{-\upsilon}$, $x \in \mathbb{R}$ (see Meyer (1992)). Therefore, since $\upsilon > 1+\delta$, $\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}(\psi)(x)\right|^2 dx \leq C^2\int_{-\infty}^{\infty}(1+x^2)^{\delta}(1+|x|)^{-2\upsilon}dx < \infty$ and θ exists.
The idea of the thresholding in (9) is to do a selection on the observations: when, for $v \in \{1, \ldots, n\}$, G_v is too large, the observation Y_v is neglected. From a technical point of view, this allows us to estimate $\beta_{j,k}$ in an optimal way under mild assumptions on $(a_{d_*}(v))_{v\in\{1,\ldots,n\}}$ (and, a fortiori, on the weights of the mixture). Such a thresholding method has been introduced by Delyon and Juditsky (1996) for regression wavelet estimation.
Statistical properties of $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$ are investigated in Propositions 1, 2, 3 and 4.

We consider two wavelets estimators for f_{d_*}: a linear estimator and a hard thresholding estimator.

Linear estimator. Assuming that $f_{d_*} \in B^s_{p,r}(M)$ with $p \geq 2$, we define the linear estimator $\widehat{f}^L$ by

$$\widehat{f}^L(x) = \sum_{k=0}^{2^{j_0}-1} \widehat{\alpha}_{j_0,k}\phi_{j_0,k}(x), \tag{11}$$

where $\widehat{\alpha}_{j,k}$ is defined by (8), j_0 is the integer satisfying

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

z_n is defined by (6) and δ is the one in (3).
Note that $\widehat{f}^L$ is not adaptive since it depends on s, the smoothness parameter of f_{d_*}.

Hard thresholding estimator. We define the hard thresholding estimator $\widehat{f}^H$ by

$$\widehat{f}^H(x) = \sum_{k=0}^{2^\tau-1} \widehat{\alpha}_{\tau,k}\phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1}\sum_{k=0}^{2^j-1} \widehat{\beta}_{j,k}1_{\{|\widehat{\beta}_{j,k}|\geq\kappa\lambda_j\}}\psi_{j,k}(x), \tag{12}$$

where $\widehat{\alpha}_{\tau,k}$ and $\widehat{\beta}_{j,k}$ are defined by (8) and (9), j_1 is the integer satisfying

$$\frac{1}{2}\left(\frac{n}{z_n}\right)^{1/(2\delta+1)} < 2^{j_1} \leq \left(\frac{n}{z_n}\right)^{1/(2\delta+1)},$$

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

$$\lambda_j = \theta 2^{\delta j}\sqrt{\frac{z_n \ln(n/z_n)}{n}},$$

z_n is defined by (6) and δ is the one in (3).
Contrary to $\widehat{f}^L$, $\widehat{f}^H$ is adaptive. The feature of the hard thresholding estimator is to only estimate the "large" unknown wavelets coefficients of f_{d_*} (those which contain the main characteristics of f_{d_*}). Hard thresholding estimators for other deconvolution problems than (1) can be found in Fan and Koo (2002), Johnstone et al. (2004), Willer (2005) and Cavalier and Raimondo (2007).

5 Minimax results

Upper bounds for $\widehat{f}^L$ and $\widehat{f}^H$ are given in Theorems 1 and 2 below. Further details on the minimax approach (and rates of convergence for various models) can be found in Tsybakov (2004).

Theorem 1 *Consider* (1) *under the assumptions of Section* 2. *Suppose that* $f_{d_*} \in B^s_{p,r}(M)$ *with* $s > 0$, $p \geq 2$ *and* $r \geq 1$. *Let* $\widehat{f}^L$ *be* (11). *Then there exists a constant* $C > 0$ *such that*

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

The proof of Theorem 1 uses a moment inequality on (8) and a suitable decomposition of the MISE. Note that $\widehat{f}^L$ is constructed to minimize the MISE as much as possible. For this reason, our benchmark will be the rate of convergence: $(z_n/n)^{2s/(2s+2\delta+1)}$.

Theorem 2 *Consider* (1) *under the assumptions of Section* 2. *Let* $\widehat{f}^H$ *be* (12). *Suppose that* $f_{d_*} \in B^s_{p,r}(M)$ *with* $r \geq 1$, $\{p \geq 2$ *and* $s > 0\}$ *or* $\{p \in [1, 2)$ *and* $s > (2\delta + 1)/p\}$. *Then there exists a constant* $C > 0$ *such that*

$$\mathbb{E}\left(\int_0^1 \left(\widehat{f}^H(x) - f_{d_*}(x)\right)^2 dx\right) \leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 2 is based on several probability results (moment inequalities, concentration inequality,...) and a suitable decomposition of the MISE.

Theorem 2 shows that, besides being adaptive, $\widehat{f}^H$ attains the same rate of convergence than the one of $\widehat{f}^L$ up to the logarithmic term $(\ln(n/z_n))^{2s/(2s+2\delta+1)}$. Naturally, in the simplest case where $m = 1$, $w_1(1) = \ldots = w_1(n) = 1$, $z_n = 1$ and $f_{d_*} = f_1 = f$, the rate of convergence attained by $\widehat{f}_H$ becomes the standard one for the classical convolution density model: $(\ln n/n)^{2s/(2s+2\delta+1)}$. See (Fan and Koo 2002, Theorem 2).

Main conclusion and perspectives. We have developed a new adaptive estimator $\widehat{f}^H$ for f_{d_*} under mild assumption on the weights of the mixture. It is based on wavelet and thresholding. It has "near-optimal" minimax properties for a wide class of functions f_{d_*}. Possible perspectives of this work are

- to investigate the case where the weights of the mixture are unknown,
- to potentially improve the estimation of f_{d_*} by considering other kinds of thresholding rules as the block thresholding one (BlockJS, ...). See e.g. Cai (1999), Cai (2002), Pensky and Sapatinas (2009), Petsa and Sapatinas (2009) and Chesneau et al. (2010).

6 Proofs

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

6.1 Auxiliary results

Proposition 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 (7) of f_{d_*} . Then*

- *$\widehat{\alpha}_{j,k}$ defined by (8) is an unbiased estimator of $\alpha_{j,k}$,*
- *for $(G_v)_{v\in\{1,\ldots,n\}}$ defined by (10), we have*

$$\mathbb{E}\left(\frac{1}{n}\sum_{v=1}^{n} G_v\right) = \beta_{j,k}.$$

Proof of Proposition 1. Since X_v and ϵ_v are independent, for any $x \in \mathbb{R}$, we have

$$\begin{aligned}\mathbb{E}\left(e^{-ixY_v}\right) &= \mathbb{E}\left(e^{-ixX_v}\right)\mathbb{E}\left(e^{-ix\epsilon_v}\right) = \mathcal{F}(h_v)(x)\mathcal{F}(g)(x)\\ &= \sum_{d=1}^{m} w_d(v)\mathcal{F}(f_d)(x)\mathcal{F}(g)(x). \qquad (13)\end{aligned}$$

It follows from (13), (5) and the Plancherel-Parseval theorem that

$$\begin{aligned}\mathbb{E}\left(\widehat{\alpha}_{j,k}\right) &= \frac{1}{2\pi n}\sum_{v=1}^{n} a_{d_*}(v)\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}\mathbb{E}\left(e^{-ixY_v}\right)dx\\ &= \frac{1}{2\pi n}\sum_{v=1}^{n} a_{d_*}(v)\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}\sum_{d=1}^{m} w_d(v)\mathcal{F}(f_d)(x)\mathcal{F}(g)(x)dx\\ &= \sum_{d=1}^{m}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)\mathcal{F}(f_d)(x)dx\right)\frac{1}{n}\sum_{v=1}^{n} a_{d_*}(v)w_d(v)\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)\mathcal{F}(f_{d_*})(x)dx = \int_0^1 \overline{\phi}_{j,k}(x)f_{d_*}(x)dx = \alpha_{j,k}.\end{aligned}$$

Similarly, taking ψ instead of ϕ, we prove that

$$\mathbb{E}\left(\frac{1}{n}\sum_{v=1}^{n} G_v\right) = \frac{1}{2\pi n}\sum_{v=1}^{n} a_{d_*}(v)\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\psi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}\mathbb{E}\left(e^{-ixY_v}\right)dx = \beta_{j,k}.$$

This complete the proof of Proposition 1.

□

Proposition 2 *For any integer $j \geq \tau$ and any $k \in \{0, \ldots, 2^j - 1\}$, let $\alpha_{j,k}$ be the wavelet coefficient (7) of f_{d_*} and $\widehat{\alpha}_{j,k}$ be (8). Then there exists a constant $C > 0$ such that*

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

Proof of Proposition 2. By Proposition 1, $\widehat{\alpha}_{j,k}$ is an unbiased estimator of $\alpha_{j,k}$. Therefore, using the independence of $Y_1, \ldots, Y_n$, we obtain

$$\begin{aligned}\mathbb{E}\left(\left(\widehat{\alpha}_{j,k}-\alpha_{j,k}\right)^2\right) &= \mathbb{V}\left(\widehat{\alpha}_{j,k}\right) \\ &= \frac{1}{(2\pi)^2 n^2}\sum_{v=1}^{n} a_{d_*}^2(v)\mathbb{V}\left(\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}e^{-ixY_v}dx\right).\end{aligned} \tag{14}$$

For any $v \in \{1, \ldots, n\}$, we have

$$\mathbb{V}\left(\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}e^{-ixY_v}dx\right) \leq \mathbb{E}\left(\left|\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}e^{-ixY_v}dx\right|^2\right). \tag{15}$$

Since X_v and ϵ_v are independent, the density of Y_v is

$$w_v(x) = (h_v \star g)(x) = \int_{-\infty}^{\infty} h_v(t)g(x-t)dt, \qquad x \in \mathbb{R}.$$

Therefore

$$\begin{aligned}\mathbb{E}\left(\left|\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}e^{-ixY_v}dx\right|^2\right) &= \int_{-\infty}^{\infty}\left|\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}e^{-ixy}dx\right|^2 w_v(y)dy \\ &= \int_{-\infty}^{\infty}\left|\mathcal{F}\left(\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(.)}{\mathcal{F}(g)(.)}\right)(y)\right|^2 w_v(y)dy.\end{aligned} \tag{16}$$

Since, by (2), $\sup_{x\in\mathbb{R}} g(x) \leq C_*$ and h_v is a density, we have

$$\sup_{v\in\{1,\ldots,n\}}\sup_{x\in\mathbb{R}} w_v(x) \leq C_* \sup_{v\in\{1,\ldots,n\}}\int_{-\infty}^{\infty} h_v(t)dt = C_*.$$

The Plancherel-Parseval theorem and (3) imply that

$$\begin{aligned}\int_{-\infty}^{\infty}\left|\mathcal{F}\left(\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(.)}{\mathcal{F}(g)(.)}\right)(y)\right|^2 w_v(y)dy &\leq C_*\int_{-\infty}^{\infty}\left|\mathcal{F}\left(\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(.)}{\mathcal{F}(g)(.)}\right)(y)\right|^2 dy \\ &= 2\pi C_*\int_{-\infty}^{\infty}\left|\frac{\overline{\mathcal{F}\left(\phi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)}\right|^2 dx \\ &\leq 2\pi\frac{C_*}{c_*^2}\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}\left(\phi_{j,k}\right)(x)\right|^2 dx.\end{aligned} \tag{17}$$

Since $|\mathcal{F}(\phi_{j,k})(x)| = 2^{-j/2}\left|\mathcal{F}(\phi)(x/2^j)\right|$, by a change of variables, we have

$$\begin{aligned}\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}(\phi_{j,k})(x)\right|^2dx &= 2^{-j}\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}(\phi)(x/2^j)\right|^2dx\\ &= \int_{-\infty}^{\infty}(1+2^{2j}x^2)^{\delta}\left|\mathcal{F}(\phi)(x)\right|^2dx\\ &\leq 2^{2\delta j}\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}(\phi)(x)\right|^2dx. \quad (18)\end{aligned}$$

It follows from (16), (17) and (18) that

$$\mathbb{E}\left(\left|\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixY_v}dx\right|^2\right)\leq \theta_*2^{2\delta j}, \quad (19)$$

with $\theta_* = 2\pi(C_*/c_*^2)\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}(\phi)(x)\right|^2dx$. By (14), (15) and (19), we obtain

$$\mathbb{E}\left((\widehat{\alpha}_{j,k}-\alpha_{j,k})^2\right)\leq \theta_*2^{2\delta j}\frac{1}{(2\pi)^2n}\left(\frac{1}{n}\sum_{v=1}^{n}a_{d_*}^2(v)\right)\leq C2^{2\delta j}\frac{z_n}{n}.$$

The proof of Proposition 2 is complete.

□

Proposition 3 *For any integer $j\geq\tau$ and any $k\in\{0,\ldots,2^j-1\}$, let $\beta_{j,k}$ be the wavelet coefficient (7) of f_{d_*} and $\widehat{\beta}_{j,k}$ be (9). Then there exists a constant $C>0$ such that*

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

Proof of Proposition 3. Thanks to Proposition 1, we have

$$\beta_{j,k} = \mathbb{E}\left(\frac{1}{n}\sum_{v=1}^{n}G_v\right) = \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}(G_v1_{\{|G_v|\leq\eta_j\}}) + \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}(G_v1_{\{|G_v|>\eta_j\}}) \quad (20)$$

Hence

$$\begin{aligned}&\mathbb{E}\left(\left(\widehat{\beta}_{j,k}-\beta_{j,k}\right)^4\right)\\ &= \mathbb{E}\left(\left(\frac{1}{n}\sum_{v=1}^{n}\left(G_v1_{\{|G_v|\leq\eta_j\}}-\mathbb{E}\left(G_v1_{\{|G_v|\leq\eta_j\}}\right)\right)-\frac{1}{n}\sum_{v=1}^{n}\mathbb{E}(G_v1_{\{|G_v|>\eta_j\}})\right)^4\right)\\ &\leq 8(A+B), \quad (21)\end{aligned}$$

where

$$A = \mathbb{E}\left(\left(\frac{1}{n}\sum_{v=1}^{n}\left(G_v1_{\{|G_v|\leq\eta_j\}}-\mathbb{E}(G_v1_{\{|G_v|\leq\eta_j\}})\right)\right)^4\right)$$

and

$$B = \left(\frac{1}{n} \sum_{v=1}^{n} \mathbb{E}(G_v 1_{\{|G_v|>\eta_j\}}) \right)^4 .$$

Let us bound A and B, in turn. To bound A, we need the Rosenthal inequality presented in lemma below (see Rosenthal (1970)).

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

$$\mathbb{E}\left(\left| \sum_{v=1}^{n} U_v \right|^p \right) \leq C \max \left(\sum_{v=1}^{n} \mathbb{E}\left(|U_v|^p\right), \left(\sum_{v=1}^{n} \mathbb{E}\left(U_v^2\right) \right)^{p/2} \right).$$

Applying the Rosenthal inequality with $p = 4$ and, for any $v \in \{1,\ldots,n\}$,

$$U_v = G_v 1_{\{|G_v|\leq\eta_j\}} - \mathbb{E}\left(G_v 1_{\{|G_v|\leq\eta_j\}}\right),$$

we obtain

$$A = \frac{1}{n^4}\mathbb{E}\left(\left(\sum_{v=1}^{n} U_v \right)^4 \right) \leq C \frac{1}{n^4} \max \left(\sum_{v=1}^{n} \mathbb{E}\left(U_v^4\right), \left(\sum_{v=1}^{n} \mathbb{E}\left(U_v^2\right) \right)^2 \right).$$

Using (19) with ψ instead of ϕ, we have, for any $a \in \{2,4\}$ and any $v \in \{1,\ldots,n\}$,

$$\mathbb{E}\left(U_v^a\right) \leq 2^a \mathbb{E}\left(G_v^a 1_{\{|G_v|\leq\eta_j\}}\right) \leq 2^a \eta_j^{a-2} \mathbb{E}\left(G_v^2\right) \leq 2^a \theta^2 \eta_j^{a-2} 2^{2\delta j} a_{d_*}^2(v).$$

Hence, since $z_n < n/e$,

$$\begin{aligned} A &\leq C \frac{1}{n^4} \max \left(2^4 \theta^2 \eta_j^2 2^{2\delta j} \sum_{v=1}^{n} a_{d_*}^2(v), \left(2^2 \theta^2 2^{2\delta j} \sum_{v=1}^{n} a_{d_*}^2(v) \right)^2 \right) \\ &\leq C \frac{1}{n^4} \max \left(2^{4\delta j} \frac{n^2 z_n^2}{\ln(n/z_n)}, 2^{4\delta j} n^2 z_n^2 \right) = C 2^{4\delta j} \frac{z_n^2}{n^2}. \end{aligned} \tag{22}$$

Let us now bound B. Using the inequality: $1_{\{|G_v|>\eta_j\}} \leq (1/\eta_j)|G_v|$, and again (19) with ψ instead of ϕ, we obtain

$$\begin{aligned} \frac{1}{n} \sum_{v=1}^{n} \mathbb{E}\left(|G_v| 1_{\{|G_v|>\eta_j\}}\right) &\leq \frac{1}{\eta_j} \left(\frac{1}{n} \sum_{v=1}^{n} \mathbb{E}\left(G_v^2\right) \right) \\ &\leq \frac{1}{\theta 2^{\delta j}} \sqrt{\frac{\ln(n/z_n)}{n z_n}} \theta^2 2^{2\delta j} z_n \\ &= \theta 2^{\delta j} \sqrt{\frac{z_n \ln(n/z_n)}{n}}. \end{aligned} \tag{23}$$

Hence

$$B \leq C 2^{4\delta j} \frac{(z_n \ln(n/z_n))^2}{n^2}. \tag{24}$$

It follows from (21), (22), (24) and $z_n < n/e$ that

$$\begin{aligned}\mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k}\right)^4\right) &\leq C\left(2^{4\delta j}\frac{z_n^2}{n^2} + 2^{4\delta j}\frac{(z_n \ln(n/z_n))^2}{n^2}\right) \\ &\leq C 2^{4\delta j}\frac{(z_n \ln(n/z_n))^2}{n^2}.\end{aligned}$$

This complete the proof of Proposition 3.

□

Proposition 4 *For any integer $j \geq \tau$ and any $k \in \{0, \ldots, 2^j - 1\}$, let $\beta_{j,k}$ be the wavelet coefficient (7) of f_{d_*} and $\widehat{\beta}_{j,k}$ be (9). Then, for any $\kappa \geq 8/3 + 2 + 2\sqrt{16/9 + 4}$,*

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

Proof of Proposition 4. Using (20), we have

$$\begin{aligned}&|\widehat{\beta}_{j,k} - \beta_{j,k}| \\ &= \left|\frac{1}{n}\sum_{v=1}^{n}\left(G_v 1_{\{|G_v| \leq \eta_j\}} - \mathbb{E}\left(G_v 1_{\{|G_v| \leq \eta_j\}}\right)\right) - \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}\left(G_v 1_{\{|G_v| > \eta_j\}}\right)\right| \\ &\leq \left|\frac{1}{n}\sum_{v=1}^{n}\left(G_v 1_{\{|G_v| \leq \eta_j\}} - \mathbb{E}\left(G_v 1_{\{|G_v| \leq \eta_j\}}\right)\right)\right| + \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}\left(|G_v| 1_{\{|G_v| > \eta_j\}}\right).\end{aligned}$$

By (23) we obtain

$$\frac{1}{n}\sum_{v=1}^{n}\mathbb{E}\left(|G_v| 1_{\{|G_v| > \eta_j\}}\right) \leq \theta 2^{\delta j}\sqrt{\frac{z_n \ln(n/z_n)}{n}} = \lambda_j.$$

Hence

$$\begin{aligned}S &= \mathbb{P}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}| \geq \kappa\lambda_j/2\right) \\ &\leq \mathbb{P}\left(\left|\frac{1}{n}\sum_{v=1}^{n}\left(G_v 1_{\{|G_v| \leq \eta_j\}} - \mathbb{E}\left(G_v 1_{\{|G_v| \leq \eta_j\}}\right)\right)\right| \geq (\kappa/2 - 1)\lambda_j\right).\end{aligned}$$

Let us now present the Bernstein inequality (see Petrov (1995)).

Lemma 2 (Bernstein's inequality) *Let $n \in \mathbb{N}^*$ and $(U_v)_{v \in \{1,\ldots,n\}}$ be n zero mean independent random variables such that there exists a constant $M > 0$ satisfying, for any $v \in \{1, \ldots, n\}$, $|U_v| \leq M < \infty$. Then, for any $\lambda > 0$, we have*

$$\mathbb{P}\left(\left|\sum_{v=1}^{n} U_v\right| \geq \lambda\right) \leq 2\exp\left(-\frac{\lambda^2}{2\left(\sum_{v=1}^{n}\mathbb{E}\left(U_v^2\right) + \frac{\lambda M}{3}\right)}\right).$$

Let us set, for any $v \in \{1, \dots, n\}$,

$$U_v = G_v 1_{\{|G_v| \le \eta_j\}} - \mathbb{E}\left(G_v 1_{\{|G_v| \le \eta_j\}}\right).$$

Then $\mathbb{E}(U_v) = 0$,

$$|U_v| \le |G_v|\, 1_{\{|G_v| \le \eta_j\}} + \mathbb{E}\left(|G_v| 1_{\{|G_v| \le \eta_j\}}\right) \le 2\eta_j$$

and, using again (19) with ψ instead of ϕ,

$$\begin{aligned}\sum_{v=1}^{n} \mathbb{E}\left(U_v^2\right) = \sum_{v=1}^{n} \mathbb{V}\left(G_v 1_{\{|G_v| \le \eta_j\}}\right) &\le \sum_{v=1}^{n} \mathbb{E}\left(G_v^2\right) \le \theta^2 2^{2\delta j} \sum_{v=1}^{n} a_{d_*}^2(v) \\ &\le \theta^2 2^{2\delta j} n z_n.\end{aligned}$$

It follows from the Bernstein inequality that

$$S \le 2\exp\left(-\frac{n^2(\kappa/2-1)^2\lambda_j^2}{2\left(\theta^2 2^{2\delta j} n z_n + \frac{2n(\kappa/2-1)\lambda_j\eta_j}{3}\right)}\right).$$

Since

$$\lambda_j\eta_j = \theta 2^{\delta j}\sqrt{\frac{z_n \ln(n/z_n)}{n}}\theta 2^{\delta j}\sqrt{\frac{nz_n}{\ln(n/z_n)}} = \theta^2 2^{2\delta j} z_n$$

and

$$\lambda_j^2 = \theta^2 2^{2\delta j}\frac{z_n \ln(n/z_n)}{n},$$

for any $\kappa \ge 8/3 + 2 + 2\sqrt{16/9+4}$, we have

$$S \le 2\exp\left(-\frac{(\kappa/2-1)^2 \ln(n/z_n)}{2\left(1+\frac{2(\kappa/2-1)}{3}\right)}\right) = 2\left(\frac{n}{z_n}\right)^{-\frac{(\kappa/2-1)^2}{2\left(1+\frac{2(\kappa/2-1)}{3}\right)}} \le 2\left(\frac{z_n}{n}\right)^2.$$

This complete the proof of Proposition 4.

□

6.2 Proofs of the main results

Proof of Theorem 1. We expand the function f_{d_*} on $\mathcal{B}$ as

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

where

$$\alpha_{j_0,k} = \int_0^1 f_{d_*}(x)\overline{\phi}_{j_0,k}(x)dx, \qquad \beta_{j,k} = \int_0^1 f_{d_*}(x)\overline{\psi}_{j,k}(x)dx.$$

We have

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

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

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

where

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

Using Proposition 2, we obtain

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

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

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

Therefore

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

The proof of Theorem 1 is complete.

□

Proof of Theorem 2. We expand the function f_{d_*} on $\mathcal{B}$ as

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

where

$$\alpha_{\tau,k} = \int_0^1 f_{d_*}(x)\overline{\phi}_{\tau,k}(x)dx, \qquad \beta_{j,k} = \int_0^1 f_{d_*}(x)\overline{\psi}_{j,k}(x)dx.$$

We have

$$\begin{aligned}
&\widehat{f}^H(x) - f_{d_*}(x) \\
&= \sum_{k=0}^{2^\tau-1} (\widehat{\alpha}_{\tau,k} - \alpha_{\tau,k})\phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \left(\widehat{\beta}_{j,k} 1_{\left\{|\widehat{\beta}_{j,k}| \geq \kappa\lambda_j\right\}} - \beta_{j,k}\right) \psi_{j,k}(x) \\
&- \sum_{j=j_1+1}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}\psi_{j,k}(x).
\end{aligned}$$

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

$$\mathbb{E}\left(\int_0^1 \left(\widehat{f}^H(x) - f_{d_*}(x)\right)^2 dx\right) = R + S + T, \tag{25}$$

where

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

and

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

Let us bound R, T and S, in turn.

Using Proposition 2 and the inequalities: $z_n < n/e$, $z_n \ln(n/z_n) < n$ and $2s/(2s+2\delta+1) < 1$, we obtain

$$R \leq C 2^{\tau(1+2\delta)} \frac{z_n}{n} \leq C \frac{z_n}{n} \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{26}$$

For $r \geq 1$ and $p \geq 2$, we have $B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M)$. Since $z_n \ln(n/z_n) < n$ and $2s/(2s+2\delta+1) < 2s/(2\delta+1)$, we have

$$\begin{aligned} T &\leq C \sum_{j=j_1+1}^{\infty} 2^{-2js} \leq C 2^{-2j_1 s} \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2\delta+1)} \\ &\leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \end{aligned}$$

For $r \geq 1$ and $p \in [1,2)$, we have $B^s_{p,r}(M) \subseteq B^{s+1/2-1/p}_{2,\infty}(M)$. Since $s > (2\delta+1)/p$, we have $(s+1/2-1/p)/(2\delta+1) > s/(2s+2\delta+1)$. So, using again $z_n \ln(n/z_n) < n$,

$$\begin{aligned} T &\leq C \sum_{j=j_1+1}^{\infty} 2^{-2j(s+1/2-1/p)} \leq C 2^{-2j_1(s+1/2-1/p)} \\ &\leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2(s+1/2-1/p)/(2\delta+1)} \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \end{aligned}$$

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

$$T \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{27}$$

The term S can be decomposed as

$$S = S_1 + S_2 + S_3 + S_4, \tag{28}$$

where

$$S_1 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k} \right)^2 1_{\left\{ |\widehat{\beta}_{j,k}| \geq \kappa\lambda_j \right\}} 1_{\{|\beta_{j,k}| < \kappa\lambda_j/2\}} \right),$$

$$S_2 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k} \right)^2 1_{\left\{ |\widehat{\beta}_{j,k}| \geq \kappa\lambda_j \right\}} 1_{\{|\beta_{j,k}| \geq \kappa\lambda_j/2\}} \right),$$

$$S_3 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E}\left(\beta_{j,k}^2 1_{\left\{ |\widehat{\beta}_{j,k}| < \kappa\lambda_j \right\}} 1_{\{|\beta_{j,k}| \geq 2\kappa\lambda_j\}} \right)$$

and

$$S_4 = \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E}\left(\beta_{j,k}^2 1_{\left\{ |\widehat{\beta}_{j,k}| < \kappa\lambda_j \right\}} 1_{\{|\beta_{j,k}| < 2\kappa\lambda_j\}} \right).$$

Let us analyze each term S_1, S_2, S_3 and S_4 in turn.
Upper bounds for S_1 and S_3. We have

$$\left\{ |\widehat{\beta}_{j,k}| < \kappa\lambda_j,\ |\beta_{j,k}| \geq 2\kappa\lambda_j \right\} \subseteq \left\{ |\widehat{\beta}_{j,k} - \beta_{j,k}| > \kappa\lambda_j/2 \right\},$$

$$\left\{ |\widehat{\beta}_{j,k}| \geq \kappa\lambda_j,\ |\beta_{j,k}| < \kappa\lambda_j/2 \right\} \subseteq \left\{ |\widehat{\beta}_{j,k} - \beta_{j,k}| > \kappa\lambda_j/2 \right\}$$

and

$$\left\{ |\widehat{\beta}_{j,k}| < \kappa\lambda_j,\ |\beta_{j,k}| \geq 2\kappa\lambda_j \right\} \subseteq \left\{ |\beta_{j,k}| \leq 2|\widehat{\beta}_{j,k} - \beta_{j,k}| \right\}.$$

So

$$\max(S_1, S_3) \leq C \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k} \right)^2 1_{\left\{ |\widehat{\beta}_{j,k} - \beta_{j,k}| > \kappa\lambda_j/2 \right\}} \right).$$

It follows from the Cauchy-Schwarz inequality and Propositions 3 and 4 that

$$\begin{aligned}
&\mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k} \right)^2 1_{\left\{ |\widehat{\beta}_{j,k} - \beta_{j,k}| > \kappa\lambda_j/2 \right\}} \right) \\
&\leq \left(\mathbb{E}\left(\left(\widehat{\beta}_{j,k} - \beta_{j,k} \right)^4 \right) \right)^{1/2} \left(\mathbb{P}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}| > \kappa\lambda_j/2 \right) \right)^{1/2} \\
&\leq C 2^{2\delta j} \frac{z_n^2 \ln(n/z_n)}{n^2}.
\end{aligned}$$

Since $z_n \ln(n/z_n) < n$ and $2s/(2s + 2\delta + 1) < 1$, we have

$$\begin{aligned}
\max(S_1, S_3) &\leq C \frac{z_n^2 \ln(n/z_n)}{n^2} \sum_{j=\tau}^{j_1} 2^{j(1+2\delta)} \leq C \frac{z_n^2 \ln(n/z_n)}{n^2} 2^{j_1(1+2\delta)} \\
&\leq C \frac{z_n \ln(n/z_n)}{n} \leq C \left(\frac{z_n \ln(n/z_n)}{n} \right)^{2s/(2s+2\delta+1)}. \qquad (29)
\end{aligned}$$

Upper bound for S_2. Using Proposition 3 and the Cauchy-Schwarz inequality, we obtain

$$\mathbb{E}\left(\left(\widehat{\beta}_{j,k}-\beta_{j,k}\right)^2\right) \le \left(\mathbb{E}\left(\left(\widehat{\beta}_{j,k}-\beta_{j,k}\right)^4\right)\right)^{1/2} \le C2^{2\delta j}\frac{z_n\ln(n/z_n)}{n}.$$

Hence

$$S_2 \le C\frac{z_n\ln(n/z_n)}{n}\sum_{j=\tau}^{j_1}2^{2\delta j}\sum_{k=0}^{2^j-1}1_{\{|\beta_{j,k}|>\kappa\lambda_j/2\}}.$$

Let j_2 be the integer defined by

$$\frac{1}{2}\left(\frac{n}{z_n\ln(n/z_n)}\right)^{1/(2s+2\delta+1)} < 2^{j_2} \le \left(\frac{n}{z_n\ln(n/z_n)}\right)^{1/(2s+2\delta+1)}. \quad (30)$$

We have

$$S_2 \le S_{2,1}+S_{2,2},$$

where

$$S_{2,1} = C\frac{z_n\ln(n/z_n)}{n}\sum_{j=\tau}^{j_2}2^{2\delta j}\sum_{k=0}^{2^j-1}1_{\{|\beta_{j,k}|>\kappa\lambda_j/2\}}$$

and

$$S_{2,2} = C\frac{z_n\ln(n/z_n)}{n}\sum_{j=j_2+1}^{j_1}2^{2\delta j}\sum_{k=0}^{2^j-1}1_{\{|\beta_{j,k}|>\kappa\lambda_j/2\}}.$$

We have

$$\begin{aligned}S_{2,1} &\le C\frac{z_n\ln(n/z_n)}{n}\sum_{j=\tau}^{j_2}2^{j(1+2\delta)} \le C\frac{z_n\ln(n/z_n)}{n}2^{j_2(1+2\delta)}\\ &\le C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.\end{aligned}$$

For $r\ge 1$ and $p\ge 2$, since $B^s_{p,r}(M)\subseteq B^s_{2,\infty}(M)$,

$$\begin{aligned}S_{2,2} &\le C\frac{z_n\ln(n/z_n)}{n}\sum_{j=j_2+1}^{j_1}2^{2\delta j}\frac{1}{\lambda_j^2}\sum_{k=0}^{2^j-1}\beta_{j,k}^2 = C\sum_{j=j_2+1}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^2 \le C2^{-2j_2s}\\ &\le C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.\end{aligned}$$

For $r \geq 1$, $p \in [1,2)$ and $s > (2\delta+1)/p$, since $B^s_{p,r}(M) \subseteq B^{s+1/2-1/p}_{2,\infty}(M)$ and $(2s+2\delta+1)(2-p)/2 + (s+1/2-1/p+\delta-2\delta/p)p = 2s$, we have

$$\begin{aligned}
S_{2,2} &\leq C\frac{z_n \ln(n/z_n)}{n} \sum_{j=j_2+1}^{j_1} 2^{2\delta j} \frac{1}{\lambda_j^p} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \\
&\leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{(2-p)/2} \sum_{j=j_2+1}^{\infty} 2^{j\delta(2-p)} 2^{-j(s+1/2-1/p)p} \\
&\leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{(2-p)/2} 2^{-j_2(s+1/2-1/p+\delta-2\delta/p)p} \\
&\leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.
\end{aligned}$$

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

$$S_2 \leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{31}$$

Upper bound for S_4. We have

$$S_4 \leq \sum_{j=\tau}^{j_1} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 1_{\{|\beta_{j,k}|<2\kappa\lambda_j\}}.$$

Let j_2 be the integer (30). We have

$$S_4 \leq S_{4,1} + S_{4,2},$$

where

$$S_{4,1} = \sum_{j=\tau}^{j_2} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 1_{\{|\beta_{j,k}|<2\kappa\lambda_j\}}, \qquad S_{4,2} = \sum_{j=j_2+1}^{j_1} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 1_{\{|\beta_{j,k}|<2\kappa\lambda_j\}}.$$

We have

$$\begin{aligned}
S_{4,1} &\leq C\sum_{j=\tau}^{j_2} 2^j \lambda_j^2 = C\frac{z_n \ln(n/z_n)}{n} \sum_{j=\tau}^{j_2} 2^{j(1+2\delta)} \leq C\frac{z_n \ln(n/z_n)}{n} 2^{j_2(1+2\delta)} \\
&\leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.
\end{aligned}$$

For $r \geq 1$ and $p \geq 2$, since $B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M)$, we have

$$S_{4,2} \leq \sum_{j=j_2+1}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \leq C2^{-2j_2 s} \leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.$$

For $r \geq 1$, $p \in [1,2)$ and $s > (2\delta+1)/p$, since $B^s_{p,r}(M) \subseteq B^{s+1/2-1/p}_{2,\infty}(M)$ and $(2s+2\delta+1)(2-p)/2 + (s+1/2-1/p+\delta-2\delta/p)p = 2s$, we have

$$\begin{aligned}
S_{4,2} &\leq C \sum_{j=j_2+1}^{j_1} \lambda_j^{2-p} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \\
&= C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{(2-p)/2} \sum_{j=j_2+1}^{j_1} 2^{j\delta(2-p)} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \\
&\leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{(2-p)/2} \sum_{j=j_2+1}^{\infty} 2^{j\delta(2-p)} 2^{-j(s+1/2-1/p)p} \\
&\leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{(2-p)/2} 2^{-j_2(s+1/2-1/p+\delta-2\delta/p)p} \\
&\leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.
\end{aligned}$$

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

$$S_4 \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{32}$$

It follows from (28), (29), (31) and (32) that

$$S \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{33}$$

Combining (25), (26), (27) and (33), we have, for $r \geq 1$, $\{p \geq 2 \text{ and } s > 0\}$ or $\{p \in [1,2) \text{ and } s > (2\delta+1)/p\}$,

$$\mathbb{E}\left(\int_0^1 \left(\widehat{f}^H(x) - f_{d_*}(x)\right)^2 dx\right) \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 2 is complete.

□

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

References

Butucea, C. and Matias, C. (2005). Minimax estimation of the noise level and of the signal density in a semiparametric convolution model. *Bernoulli*, 11, **2**, 309-340.

Cai, T. (1999). Adaptive Wavelet Estimation: A Block Thresholding And Oracle Inequality Approach. *The Annals of Statistics*, 27, 898–924.

Cai, T. (2002). On adaptive wavelet estimation of a derivative and other related linear inverse problems. *Journal of Statistical Planning and Inference*, 108, 329–349.

Cavalier, L. and Raimondo, M. (2007). Wavelet deconvolution with noisy eigenvalues. *IEEE Transactions on signal processing*, 55, 2414–2424.

Chesneau, C., Fadili, J. and Starck, J.-L. (2010). Stein block thresholding for image denoising. *Applied and Computational Harmonic Analysis*, 28, **1**, 67–88.

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.

Comte, F., Rozenholc, Y. and Taupin, M.-L. (2006). Penalized contrast estimator for density deconvolution. *The Canadian Journal of Statistics*, 34, 431-452.

Delyon, B. and Juditsky, A. (1996). On minimax wavelet estimators. *Applied Computational Harmonic Analysis*, **3**, 215–228.

Devroye, L. (1989). Consistent deconvolution in density estimation. *Canad. Journ. Statist.*, 17, 235-239.

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problem. *Ann. Statist.*, 19, 1257-1272.

Fan, J. and Koo, J.Y. (2002). Wavelet deconvolution. *IEEE transactions on information theory*, 48, 734–747.

Hall, P. and Carol, R.J. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.*, 83, 1184-1186.

Hall, P. and Zhou, X.H. (2003). Nonparametric estimation of component distributions in a multivariate mixture. *Ann. Statist.*, 31 (1), 201-224.

Johnstone, I., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004). Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society, Serie B*, 66 (3), 547–573.

Lacour, C. (2006). Rates of convergence for nonparametric deconvolution. *C. R. Acad. Sci. Paris Ser. I Math.*, 342 (11), 877-882.

Lee, M., Shen, H., Burch, C. and Marron, J.S. (2010). Direct deconvolution density estimation of a mixture, distribution motivated by mutation effects distribution. *Journal of nonparametric statistics*, **22**, 1 1-22.

Maiboroda, R. E. (1996). Estimators of components of a mixture with varying concentrations. *Ukrain. Mat. Zh.*, **48**, 4, 562-566.

Meyer, Y. (1992). Wavelets and Operators. *Cambridge University Press*, Cambridge.

Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *The Annals of Statistics*, 27, 2033–2053.

Pensky, M. and Sapatinas, T. (2009). Functional deconvolution in a periodic setting: Uniform Case. *The Annals of Statistics*, 37, **1**, 73–104.

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

Petsa, A. and Sapatinas, T. (2009). Minimax convergence rates under the $\mathbb{L}^p$-risk in the functional deconvolution model. *Statistics and Probability Letters*, 79, 1568-1576.

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. (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.

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*, **8**, 273-303.

Tsybakov, A. (2004). *Introduction à l'estimation nonparametrique*. Springer Verlag, Berlin.

Willer, T. (2005). Deconvolution in white noise with a random blurring effect. Preprint LPMA.