Sankhyā : The Indian Journal of Statistics
2011, Volume 73-A, Part 2, pp. 245-266

Wavelet Density Estimators for the Deconvolution of a Component from a Mixture

Christophe Chesneau
University of Caen, Caen, France

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 considering the mean integrated squared error over Besov balls. We prove that the adaptive one attains a sharp rate of convergence.

AMS (2000) *subject classification.* Primary 62G07, 62G20.
Keywords and phrases. Density deconvolution, mixture, rate of convergence, wavelets, hard thresholding.

1 Motivations

We consider the following model:

$$Y_v = X_v + \epsilon_v, \tag{1.1}$$

$v \in \{1, \ldots, n\}$, $n \in \mathbb{N}^* = \{1, 2, \ldots\}$, where $X_1, \ldots, X_n$ are independent random variables and $\epsilon_1, \ldots, \epsilon_n$ are *i.i.d.* random variables. For any $v \in \{1, \ldots, n\}$, X_v and ϵ_v are independent. The density of ϵ_1 is known and is denoted by g. For any $v \in \{1, \ldots, n\}$, the density of X_v is the following finite mixture:

$$h_v(x) = \sum_{d=1}^{m} w_d(v) f_d(x), \qquad x \in [-\Omega, \Omega],$$

where

- $\Omega \in (0, \infty)$, $m \in \mathbb{N}^*$,
- $(w_d(v))_{(v,d)\in\{1,\ldots,n\}\times\{1,\ldots,m\}}$ are known positive weights such that, for any $v \in \{1, \ldots, n\}$,

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

- $f_1, \ldots, f_m$ 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 (1.1) has been recently described for a particular mixture in van Es, Gugushvili and Spreij (2008) and Lee et al. (2010). In the simplest case where $m = 1$, $w_1(1) = \ldots = w_1(n) = 1$ and $f_{d_*} = f_1 = f$, (1.1) becomes the standard convolution density model. See e.g. Caroll and Hall (1988), Devroye (1989), Fan (1991), Pensky and Vidakovic (1999), Fan and Koo (2002), Butucea and Matias (2005), Comte, Rozenholc and Taupin (2006), Delaigle and Gijbels (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 (2.2)), 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 mean integrated squared error (MISE) over a wide class of functions: the Besov ball $B^s_{p,r}(M)$ (to be defined in Section 3). 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 attained by the "best" nonadaptive 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 results are set in Section 5. Technical proofs are given in Section 7.

2 Assumptions and Notations

Assumption on $f_1, \ldots, f_m$. Without loss of generality, for any $d \in \{1, \ldots, m\}$, we assume that the support of f_d is $[-\Omega, \Omega]$.

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

$$\sup_{x \in \mathbb{R}} g(x) \leq C_* < \infty. \tag{2.1}$$

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

$$|\mathcal{F}(g)(x)| \geq \frac{c_*}{(1+x^2)^{\delta/2}}. \tag{2.2}$$

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

EXAMPLE 2.1. 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 (2.2) is satisfied with $c_* = 1$ and $\delta = 2p$.

Assumptions on the weights. In the sequel, $a(v)$ is the v-th entry of a vector a. We adopt the notation $\langle a, b\rangle_n = (1/n)\sum_{v=1}^{n} a(v)b(v)$ for the normalized euclidean inner product in $\mathbb{R}^n$, and $\|\cdot\|_n$ the associated norm.

Set $W = (w_1, \cdots, w_m) \in [0,1]^{n\times m}$. Recall that W is known. We also suppose that the Gram matrix

$$\Gamma_n = \frac{1}{n}W^T W = (\langle w_k, w_\ell\rangle_n)_{(k,\ell)\in\{1,\dots,m\}^2}$$

is (symmetric) positive-definite.

For the considered d_* and any $v \in \{1, \dots, n\}$, we set

$$a_{d_*}(v) = \frac{1}{\det(\Gamma_n)} \sum_{d=1}^{m} (-1)^{d+d_*} M_{d_*,d}^{n} w_d(v), \tag{2.3}$$

where $M_{d_*,d}^{n}$ is the minor (d_*, d) of Γ_n.

To get the gist of (2.3), it is useful to view the vector $a_{d_*} = (a_{d_*}(1), \dots, a_{d_*}(n))$ as the solution of the following quadratic objective with linear constraints

$$\min_{b\in\mathbb{R}^n} \|b\|_n^2 \quad \text{such that} \quad \langle w_d, b\rangle_n = \delta_{d_*,d}, \ \text{for } d \in \{1, \cdots, m\}, \tag{2.4}$$

where $\delta_{d_*,d}$ denotes the Kronecker delta. Using the Lagrange multipliers, one can prove that the unique minimizer of (2.4) is given by

$$a_{d_*} = W\Gamma_n^{-1}\Delta_{d_*}, \tag{2.5}$$

where Δ_{d_*} is a vector of zeros except at its d_*-th entry. Using the cofactors of Γ_n to get its inverse, we recover (2.3). Naturally, the positive-definiteness of Γ_n is important for (2.5) to make sense.

In the context of mixture density estimation, Maiboroda (1996) showed that a_{d_*} is the minimal risk weight vector to be used for the empirical measure constructed from the observations to yield an unbiased estimator of the d_*-th density in the mixture.

Further technical details can be found in Maiboroda (1996), Pokhyl'ko (2005) and Prakasa Rao (2010).

We let

$$z_n = \|a_{d_*}\|_n^2 \tag{2.6}$$

and we suppose that $z_n < n/e$.

3 Wavelets and Besov Balls

Wavelet basis on the interval. Throughout the paper, we work with the wavelet basis described below. Let N be a positive integer, and ϕ and ψ be the Daubechies wavelets $db2N$. We chose N such that $\phi, \psi \in \mathcal{C}^w$ with, for technical reasons, $w > 1 + \delta$, where δ is the one in (2.2). 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 and a set of consecutive integers Λ_j with a length proportional to 2^j such that, for any integer $\ell \geq \tau$, the collection

$$\mathcal{B} = \{\phi_{\ell,k}(.),\ k \in \Lambda_\ell;\ \psi_{j,k}(.);\ j \in \mathbb{N} - \{0, \dots, \ell - 1\},\ k \in \Lambda_j\},$$

is an orthonormal basis of $\mathbb{L}^2([-\Omega, \Omega]) = \{h : [-\Omega, \Omega] \to \mathbb{R};\ \int_{-\Omega}^{\Omega} h^2(x)dx < \infty\}$. We refer to Cohen et al. (1993).

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

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

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

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

As is traditional in the wavelet estimation literature, we will investigate the performances of our estimators by assuming that the unknown function belongs to Besov balls. Their definitions in terms of wavelet coefficients are given below.

Wavelet expression of the Besov balls. Let $M > 0$, $s > 0$, $p \geq 1$ and $r \geq 1$. A function h belongs to $B^s_{p,r}(M)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients (3.1) satisfy

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

In this expression, s is a smoothness parameter and p and r are norm parameters. Besov balls contain the Hölder and Sobolev balls. See Meyer (1992).

4 Wavelet Estimators

Estimators of the wavelet coefficients. 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 $k \in \Lambda_j$,

– we estimate $\alpha_{j,k} = \int_{-\Omega}^{\Omega} f_{d_*}(x)\phi_{j,k}(x)dx$ by

$$\widehat{\alpha}_{j,k} = \frac{1}{2\pi n}\sum_{v=1}^{n} a_{d_*}(v)\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixY_v}dx, \tag{4.1}$$

– we estimate $\beta_{j,k} = \int_{-\Omega}^{\Omega} f_{d_*}(x)\psi_{j,k}(x)dx$ by

$$\widehat{\beta}_{j,k} = \frac{1}{n}\sum_{v=1}^{n} G_v \mathbf{1}_{\{|G_v|\leq\eta_j\}}, \tag{4.2}$$

where

$$G_v = \frac{1}{2\pi}a_{d_*}(v)\int_{-\infty}^{\infty}\frac{\overline{\mathcal{F}(\psi_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixY_v}dx, \tag{4.3}$$

for any random event $\mathcal{A}$, $\mathbf{1}_{\mathcal{A}}$ is the indicator function on $\mathcal{A}$, $a_{d_*}(v)$ is defined by (2.3),

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

z_n is defined by (2.6) 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 (2.1) and (2.2)).

Remark 4.1. 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.

Remark 4.2. The definitions of $\widehat{\alpha}_{j,k}$ and $G_1, \dots, G_n$ are such that

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

The proofs are mainly based on the properties of a_{d_*} (described in (2.4)) and the Parseval-Plancherel theorem. See Proposition 7.1 below.

The idea of the thresholding in (4.2) is to do a selection on the observations: when, for $v \in \{1, \dots, n\}$, 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_*} (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.

Moments and concentration inequalities of $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$ are investigated in Propositions 7.2, 7.3 and 7.4 in Section 7.

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

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

$$\widehat{f}^{lin}(x) = \sum_{k \in \Lambda_{j_0}} \widehat{\alpha}_{j_0,k}\phi_{j_0,k}(x), \qquad x \in [-\Omega, \Omega], \tag{4.4}$$

where $\widehat{\alpha}_{j,k}$ is defined by (4.1), 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 (2.6) and δ is the one in (2.2).

Note that $\widehat{f}^{lin}$ is not adaptive since it depends on s, the smoothness parameter of f_{d_*}.

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

$$\widehat{f}^{hard}(x) = \sum_{k \in \Lambda_\tau} \widehat{\alpha}_{\tau,k}\phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1}\sum_{k \in \Lambda_j} \widehat{\beta}_{j,k}\mathbf{1}_{\left\{|\widehat{\beta}_{j,k}| \geq \kappa\lambda_j\right\}}\psi_{j,k}(x), \tag{4.5}$$

$x \in [-\Omega, \Omega]$, where $\widehat{\alpha}_{\tau,k}$ is defined by (4.1), $\widehat{\beta}_{j,k}$ by (4.2), 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

$$\lambda_j = \theta 2^{\delta j} \sqrt{\frac{z_n \ln(n/z_n)}{n}}, \tag{4.6}$$

z_n is defined by (2.6) and δ is the one in (2.2).

Contrary to $\widehat{f}^{lin}$, $\widehat{f}^{hard}$ is adaptive. The feature of the hard thresholding estimator is to only estimate the "large" unknown wavelet coefficients of f_{d_*} (those which contain the main characteristics of f_{d_*}). Hard thresholding estimators for other deconvolution problems than (1.1) can be found in Fan and Koo (2002), Johnstone et al. (2004), Willer (2005) and Cavalier and Raimondo (2007).

5 Upper Bounds

Upper bounds for $\widehat{f}^{lin}$ and $\widehat{f}^{hard}$ are given in Theorems 5.1 and 5.2 below. Further details on our statistical approach (and rates of convergence for various models) can be found in Tsybakov (2004).

Theorem 5.1. *Consider* (1.1) *under the assumptions of Section 2. Let* $\widehat{f}^{lin}$ *be* (4.4). *Then, for any* $s > 0$, $p \geq 2$ *and* $r \geq 1$, *there exists a constant* $C > 0$ *such that*

$$\sup_{f_{d_*} \in B^s_{p,r}(M)} \mathbb{E}\left(\int_{-\Omega}^{\Omega} \left(\widehat{f}^{lin}(x) - f_{d_*}(x) \right)^2 dx \right) \leq C \left(\frac{z_n}{n} \right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 5.1 uses a moment inequality on (4.1) and a suitable decomposition of the MISE. Note that $\widehat{f}^{lin}$ 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 5.2. *Consider* (1.1) *under the assumptions of Section 2. Let* $\widehat{f}^{hard}$ *be* (4.5). *Then, for any* $r \geq 1$, *any* $\{p \geq 2$ *and* $s > 0\}$ *or any* $\{p \in [1, 2)$ *and* $s > (2\delta + 1)/p\}$, *there exists a constant* $C > 0$ *such that*

$$\sup_{f_{d_*} \in B^s_{p,r}(M)} \mathbb{E}\left(\int_{-\Omega}^{\Omega} \left(\widehat{f}^{hard}(x) - f_{d_*}(x) \right)^2 dx \right) \leq C \left(\frac{z_n \ln(n/z_n)}{n} \right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 5.2 is based on several probability results (moment inequalities, concentration inequality,...) and a suitable decomposition of the MISE.

Theorem 5.2 shows that, besides being adaptive, $\widehat{f}^{hard}$ attains the same rate of convergence than the one of $\widehat{f}^{lin}$ up to $(\ln(n/z_n))^{2s/(2s+2\delta+1)}$.

Remark 5.1. 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 $\widehat{f}^{hard}$ becomes the standard one for the classical convolution density model i.e. $(\ln n/n)^{2s/(2s+2\delta+1)}$. (See Fan and Koo, 2002, Theorem 2).

Concerning the computational complexity of the proposed methodology, it is domained by that of the wavelet transform on the interval, and by that of computing a_{d_*} (more precisely its ℓ_2-norm) (see (2.5) and (2.6)). More precisely, the latter step can be efficiently accomplished through a singular value decomposition (SVD) of the matrix of probabilities W. Only the right singular vectors and the singular values are needed (there are m of them since $n > m$). The complexity of this step is that of the SVD, which is typically $O(nm^2 + m^3)$.

6 Conclusion and Perspectives

We have developed a new adaptive estimator $\widehat{f}^{hard}$ for f_{d_*} under mild assumptions on the weights of the mixture. It is based on wavelets and thresholding. It has "near-optimal" properties for a wide class of functions f_{d_*}. See below four possible perspectives of this work:

- examine the case where m is unknown.
- investigate the case where the weights of the mixture are unknown.
- determine the optimal lower bound of the model. This problem is much more complicated that it might seem at the first glance because only the component f_{d_*} is of interest (not the whole unknown density). And this is not immediately clear how to apply the standard lower bounds techniques in this case (these techniques can be found in Tsybakov, 2004, Chapter 2).
- potentially improve the estimation of f_{d_*} by considering other kinds of thresholding rules. This perspective is justified by the following result: in the case where $m = 1$, $w_1(1) = \ldots = w_1(n) = 1$, $z_n = 1$ and $f_{d_*} = f_1 = f$, the wavelet procedure based on a global thresholding rule elaborated by Pensky and Vidakovic (1999) achieves the exact optimal rate of convergence $n^{-2s/(2s+2\delta+1)}$ (without extra logarithmic term) over $B^s_{p,q}(M)$ for $s > 0$, $p \geq 2$ and $q \geq 1$. Naturally, from a technical point of view, the presence of mixtures in our deconvolution problem complicates significantly the situation. Other "block" thresholding rules can also be investigated. See e.g. Cai (1999, 2002), Pensky and Sapatinas (2009), Petsa and Sapatinas (2009) and Chesneau, Fadili and Starck (2010).

7 Proofs

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

7.1. Auxiliary results.

PROPOSITION 7.1. *For any integer $j \geq \tau$ and any $k \in \Lambda_j$, let $\alpha_{j,k}$ and $\beta_{j,k}$ be the wavelet coefficients* (3.1) *of f_{d_*}. Then*

- $\widehat{\alpha}_{j,k}$ *defined by* (4.1) *is an unbiased estimator of* $\alpha_{j,k}$,
- *for* $(G_v)_{v\in\{1,\ldots,n\}}$ *defined by* (4.3), *we have*

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

Proof of Proposition 7.1. Since X_v and ϵ_v are independent, for any $x \in \mathbb{R}$, we have

$$\begin{aligned}\mathbb{E}\left(e^{-ixY_v}\right) &= \mathbb{E}\left(e^{-ixX_v}\right)\mathbb{E}\left(e^{-ix\epsilon_v}\right) = \mathcal{F}(h_v)(x)\mathcal{F}(g)(x)\\ &= \sum_{d=1}^{m} w_d(v)\mathcal{F}(f_d)(x)\mathcal{F}(g)(x). \end{aligned} \tag{7.1}$$

It follows from the Fubini theorem, (7.1), (2.4) and the Parseval-Plancherel theorem that

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

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 completes the proof of Proposition 7.1.

□

Proposition 7.2. *For any integer* $j \geq \tau$ *and any* $k \in \Lambda_j$, *let* $\alpha_{j,k}$ *be the wavelet coefficient* (3.1) *of* f_{d_*} *and* $\widehat{\alpha}_{j,k}$ *be* (4.1). *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 7.2. By Proposition 7.1, $\widehat{\alpha}_{j,k}$ is an unbiased estimator of $\alpha_{j,k}$. Therefore, using the independence of $Y_1, \ldots, Y_n$, we obtain

$$\mathbb{E}\left(\left|\widehat{\alpha}_{j,k} - \alpha_{j,k}\right|^2\right) = \mathbb{V}\left(\widehat{\alpha}_{j,k}\right)$$

$$= \frac{1}{(2\pi)^2 n^2} \sum_{v=1}^{n} a_{d_*}^2(v) \mathbb{V}\left(\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixY_v} dx \right). \tag{7.2}$$

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

$$\mathbb{V}\left(\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixY_v} dx \right) \leq \mathbb{E}\left(\left| \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixY_v} dx \right|^2 \right). \tag{7.3}$$

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

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

Therefore

$$\begin{aligned} \mathbb{E}\left(\left| \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixY_v} dx \right|^2 \right) &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixy} dx \right|^2 q_v(y) dy \\ &= \int_{-\infty}^{\infty} \left| \mathcal{F}\left(\frac{\overline{\mathcal{F}(\phi_{j,k})}(.)}{\mathcal{F}(g)(.)} \right)(y) \right|^2 q_v(y) dy. \end{aligned} \tag{7.4}$$

Since, by (2.1), $\sup_{x\in\mathbb{R}} g(x) \leq C_*$ and h_v is a density, we have

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

The Parseval-Plancherel theorem and (2.2) imply that

$$\begin{aligned} \int_{-\infty}^{\infty} \left| \mathcal{F}\left(\frac{\overline{\mathcal{F}(\phi_{j,k})}(.)}{\mathcal{F}(g)(.)} \right)(y) \right|^2 q_v(y) dy &\leq C_* \int_{-\infty}^{\infty} \left| \mathcal{F}\left(\frac{\overline{\mathcal{F}(\phi_{j,k})}(.)}{\mathcal{F}(g)(.)} \right)(y) \right|^2 dy \\ &= 2\pi C_* \int_{-\infty}^{\infty} \left| \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} \right|^2 dx \\ &\leq 2\pi \frac{C_*}{c_*^2} \int_{-\infty}^{\infty} (1+x^2)^{\delta} \left| \mathcal{F}(\phi_{j,k})(x) \right|^2 dx. \end{aligned} \tag{7.5}$$

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

$$\int_{-\infty}^{\infty} (1+x^2)^{\delta} \left| \mathcal{F}(\phi_{j,k})(x) \right|^2 dx = 2^{-j} \int_{-\infty}^{\infty} (1+x^2)^{\delta} \left| \mathcal{F}(\phi)(x/2^j) \right|^2 dx$$

$$= \int_{-\infty}^{\infty} (1+2^{2j}x^2)^{\delta} \left|\mathcal{F}(\phi)(x)\right|^2 dx$$
$$\leq 2^{2\delta j} \int_{-\infty}^{\infty} (1+x^2)^{\delta} \left|\mathcal{F}(\phi)(x)\right|^2 dx. \quad (7.6)$$

It follows from (7.4), (7.5) and (7.6) that

$$\mathbb{E}\left(\left|\int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixY_v} dx\right|^2\right) \leq \theta_* 2^{2\delta j}, \quad (7.7)$$

with $\theta_* = 2\pi(C_*/c_*^2)\int_{-\infty}^{\infty}(1+x^2)^{\delta}\left|\mathcal{F}(\phi)(x)\right|^2 dx$. By (7.2), (7.3) and (7.7), we obtain

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

The proof of Proposition 7.2 is complete. □

Proposition 7.3. *For any integer $j \geq \tau$ and any $k \in \Lambda_j$, let $\beta_{j,k}$ be the wavelet coefficient* (3.1) *of f_{d_*} and $\widehat{\beta}_{j,k}$ be* (4.2). *Then there exists a constant $C > 0$ such that*

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

Proof of Proposition 7.3. Thanks to Proposition 7.1, we have

$$\beta_{j,k} = \mathbb{E}\left(\frac{1}{n}\sum_{v=1}^{n} G_v\right) = \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}(G_v \mathbf{1}_{\{|G_v|\leq \eta_j\}}) + \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}(G_v \mathbf{1}_{\{|G_v|>\eta_j\}}). \quad (7.8)$$

By the elementary inequality $(x+y)^4 \leq 8(x^4+y^4)$, $(x,y) \in \mathbb{R}^2$, we have

$$\mathbb{E}\left(\left|\widehat{\beta}_{j,k}-\beta_{j,k}\right|^4\right)$$
$$= \mathbb{E}\left(\left|\frac{1}{n}\sum_{v=1}^{n}\left(G_v\mathbf{1}_{\{|G_v|\leq\eta_j\}} - \mathbb{E}\left(G_v\mathbf{1}_{\{|G_v|\leq\eta_j\}}\right)\right) - \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}(G_v\mathbf{1}_{\{|G_v|>\eta_j\}})\right|^4\right)$$
$$\leq 8(A+B), \quad (7.9)$$

where

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

and

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

Let us bound A and B, in turn.

Upper bound for A. We need the Rosenthal inequality presented in lemma below (see Rosenthal, 1970).

LEMMA 7.1 (Rosenthal's inequality). *Let $p \geq 2$, $n \in \mathbb{N}^*$ and $(U_v)_{v\in\{1,\dots,n\}}$ be n zero mean independent random variables such that, for any $v \in \{1,\dots,n\}$, $\mathbb{E}(|U_v|^p) < \infty$. Then there exists a constant $C > 0$ such that*

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

Set, for any $v \in \{1,\dots,n\}$,

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

Then, for any $v \in \{1,\dots,n\}$, we have $\mathbb{E}(U_v) = 0$ and, using (7.7) with ψ instead of ϕ, for any $b \in \{2,4\}$,

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

It follows from the Rosenthal inequality and $z_n < n/e$ that

$$\begin{aligned} A &= \frac{1}{n^4}\mathbb{E}\left(\left|\sum_{v=1}^{n} U_v\right|^4\right) \leq C\frac{1}{n^4}\max\left(\sum_{v=1}^{n}\mathbb{E}\left(|U_v|^4\right), \left(\sum_{v=1}^{n}\mathbb{E}\left(|U_v|^2\right)\right)^2\right) \\ &\leq C\frac{1}{n^4}\max\left(2^4\theta^2\eta_j^2 2^{2\delta j}\sum_{v=1}^{n} a_{d_*}^2(v), \left(2^2\theta^2 2^{2\delta j}\sum_{v=1}^{n} a_{d_*}^2(v)\right)^2\right) \\ &\leq C\frac{1}{n^4}\max\left(2^{4\delta j}\frac{n^2 z_n^2}{\ln(n/z_n)}, 2^{4\delta j}n^2 z_n^2\right) = C2^{4\delta j}\frac{z_n^2}{n^2}. \end{aligned} \quad (7.10)$$

Upper bound for B. Using the inequality: $\mathbf{1}_{\{|G_v|>\eta_j\}} \leq (1/\eta_j)|G_v|$, and again (7.7) with ψ instead of ϕ, we obtain

$$\begin{aligned} \frac{1}{n}\sum_{v=1}^{n}\mathbb{E}\left(|G_v|\mathbf{1}_{\{|G_v|>\eta_j\}}\right) &\leq \frac{1}{\eta_j}\left(\frac{1}{n}\sum_{v=1}^{n}\mathbb{E}\left(|G_v|^2\right)\right) \\ &\leq \frac{1}{\theta 2^{\delta j}}\sqrt{\frac{\ln(n/z_n)}{nz_n}}\theta^2 2^{2\delta j} z_n \\ &= \theta 2^{\delta j}\sqrt{\frac{z_n\ln(n/z_n)}{n}}. \end{aligned} \quad (7.11)$$

Hence

$$B \leq C2^{4\delta j}\frac{(z_n\ln(n/z_n))^2}{n^2}. \quad (7.12)$$

It follows from (7.9), (7.10), (7.12) and $z_n < n/e$ that

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

The proof of Proposition 7.3 is complete. □

PROPOSITION 7.4. *For any integer $j \geq \tau$ and any $k \in \Lambda_j$, let $\beta_{j,k}$ be the wavelet coefficient* (3.1) *of f_{d_*}, $\widehat{\beta}_{j,k}$ be* (4.2) *and λ_j be* (4.6). *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 7.4. Using (7.8), we have

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

By (7.11) we obtain

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

Hence

$$\begin{aligned}&\mathbb{P}\left(|\widehat{\beta}_{j,k}-\beta_{j,k}| \geq \kappa\lambda_j/2\right)\\ &\quad\leq \mathbb{P}\left(\left|\frac{1}{n}\sum_{v=1}^{n}\left(G_v\mathbf{1}_{\{|G_v|\leq\eta_j\}}-\mathbb{E}\left(G_v\mathbf{1}_{\{|G_v|\leq\eta_j\}}\right)\right)\right| \geq (\kappa/2-1)\lambda_j\right).\end{aligned} \tag{7.13}$$

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

LEMMA 7.2 (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*

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

Set, for any $v \in \{1, \dots, n\}$,

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

Then, for any $v \in \{1, \dots, n\}$, we have $\mathbb{E}(U_v) = 0$,

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

and, using again (7.7) with ψ instead of ϕ,

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

It follows from the Bernstein inequality that

$$\begin{aligned} &\mathbb{P}\left(\left|\sum_{v=1}^{n} U_v\right| \ge n(\kappa/2 - 1)\lambda_j\right) \\ &\le 2\exp\left(-\frac{n^2(\kappa/2-1)^2\lambda_j^2}{2\left(\theta^2 2^{2\delta j} n z_n + \frac{2n(\kappa/2-1)\lambda_j\eta_j}{3}\right)}\right). \end{aligned} \tag{7.14}$$

Note that

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

and

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

Putting (7.13) and (7.14) together, for any $\kappa \ge 8/3 + 2 + 2\sqrt{16/9 + 4}$, we have

$$\begin{aligned} &\mathbb{P}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}| \ge \kappa\lambda_j/2\right) \\ &\le 2\exp\left(-\frac{(\kappa/2-1)^2 \ln(n/z_n)}{2\left(1 + \frac{2(\kappa/2-1)}{3}\right)}\right) = 2\left(\frac{n}{z_n}\right)^{-\frac{(\kappa/2-1)^2}{2\left(1+\frac{2(\kappa/2-1)}{3}\right)}} \le 2\left(\frac{z_n}{n}\right)^2. \end{aligned}$$

This ends the proof of Proposition 7.4. □

7.2. Proofs of Theorems 5.1 and 5.2. Proof of Theorem 5.1. We expand the function f_{d_*} on $\mathcal{B}$ as

$$f_{d_*}(x) = \sum_{k \in \Lambda_{j_0}} \alpha_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_{k \in \Lambda_j} \beta_{j,k} \psi_{j,k}(x), \qquad x \in [-\Omega, \Omega],$$

where

$$\alpha_{j_0,k} = \int_{-\Omega}^{\Omega} f_{d_*}(x)\phi_{j_0,k}(x)dx, \qquad \beta_{j,k} = \int_{-\Omega}^{\Omega} f_{d_*}(x)\psi_{j,k}(x)dx.$$

We have, for any $x \in [-\Omega, \Omega]$,

$$\widehat{f}^{lin}(x) - f_{d_*}(x) = \sum_{k\in\Lambda_{j_0}} (\widehat{\alpha}_{j_0,k} - \alpha_{j_0,k})\, \phi_{j_0,k}(x) - \sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j} \beta_{j,k}\psi_{j,k}(x).$$

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

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

where

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

Using Proposition 7.2, we obtain

$$A \le C2^{j_0(1+2\delta)}\frac{z_n}{n} \le C\left(\frac{z_n}{n}\right)^{2s/(2s+2\delta+1)}.$$

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

$$B \le C2^{-2j_0 s} \le C\left(\frac{z_n}{n}\right)^{2s/(2s+2\delta+1)}.$$

Therefore

$$\sup_{f_{d_*}\in B^s_{p,r}(M)} \mathbb{E}\left(\int_{-\Omega}^{\Omega}\left(\widehat{f}^{lin}(x) - f_{d_*}(x)\right)^2 dx\right) \le C\left(\frac{z_n}{n}\right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 5.1 is complete. □

PROOF OF THEOREM 5.2. We expand the function f_{d_*} on $\mathcal{B}$ as

$$f_{d_*}(x) = \sum_{k\in\Lambda_\tau} \alpha_{\tau,k}\phi_{\tau,k}(x) + \sum_{j=\tau}^{\infty}\sum_{k\in\Lambda_j} \beta_{j,k}\psi_{j,k}(x), \qquad x \in [-\Omega, \Omega],$$

where

$$\alpha_{\tau,k} = \int_{-\Omega}^{\Omega} f_{d_*}(x)\phi_{\tau,k}(x)dx, \qquad \beta_{j,k} = \int_{-\Omega}^{\Omega} f_{d_*}(x)\psi_{j,k}(x)dx.$$

We have, for any $x \in [-\Omega, \Omega]$,

$$\widehat{f}^{hard}(x) - f_{d_*}(x)$$

$$= \sum_{k\in\Lambda_\tau}(\widehat{\alpha}_{\tau,k}-\alpha_{\tau,k})\phi_{\tau,k}(x)+\sum_{j=\tau}^{j_1}\sum_{k\in\Lambda_j}\left(\widehat{\beta}_{j,k}\mathbf{1}_{\left\{|\widehat{\beta}_{j,k}|\geq\kappa\lambda_j\right\}}-\beta_{j,k}\right)\psi_{j,k}(x)$$
$$-\sum_{j=j_1+1}^{\infty}\sum_{k\in\Lambda_j}\beta_{j,k}\psi_{j,k}(x).$$

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

$$\mathbb{E}\left(\int_{-\Omega}^{\Omega}\left(\widehat{f}^{hard}(x)-f_{d_*}(x)\right)^2dx\right)=R+S+T, \tag{7.15}$$

where

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

and

$$T=\sum_{j=j_1+1}^{\infty}\sum_{k\in\Lambda_j}|\beta_{j,k}|^2.$$

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

Using Proposition 7.2 and the inequalities: $z_n<n/e$, $z_n\ln(n/z_n)<n$ and $2s/(2s+2\delta+1)<1$, we obtain

$$R\leq C2^{\tau(1+2\delta)}\frac{z_n}{n}\leq C\frac{z_n}{n}\leq C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{7.16}$$

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

$$\begin{aligned}T &\leq C\sum_{j=j_1+1}^{\infty}2^{-2js}\leq C2^{-2j_1s}\leq C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2s/(2\delta+1)}\\ &\leq C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.\end{aligned}$$

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 $z_n\ln(n/z_n)<n$,

$$\begin{aligned}T &\leq C\sum_{j=j_1+1}^{\infty}2^{-2j(s+1/2-1/p)}\leq C2^{-2j_1(s+1/2-1/p)}\\ &\leq C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2(s+1/2-1/p)/(2\delta+1)}\leq C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.\end{aligned}$$

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

$$T \leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{7.17}$$

We can express S as

$$S = S_1 + S_2 + S_3 + S_4, \tag{7.18}$$

where

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

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

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

and

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

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

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

It follows from the Cauchy-Schwarz inequality and Propositions 7.3 and 7.4 that

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

$$\leq \quad C2^{2\delta j}\frac{z_n^2\ln(n/z_n)}{n^2}.$$

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

$$\begin{aligned} S_1+S_3 &\leq C\frac{z_n^2\ln(n/z_n)}{n^2}\sum_{j=\tau}^{j_1}2^{j(1+2\delta)} \leq C\frac{z_n^2\ln(n/z_n)}{n^2}2^{j_1(1+2\delta)} \\ &\leq C\frac{z_n\ln(n/z_n)}{n} \leq C\left(\frac{z_n\ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \end{aligned} \tag{7.19}$$

Upper bound for S_2. Using the Cauchy-Schwarz inequality and Proposition 7.3, we obtain

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

Hence

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

Let j_2 be the integer defined by

$$\frac{1}{2}\left(\frac{n}{z_n\ln(n/z_n)}\right)^{1/(2s+2\delta+1)} < 2^{j_2} \leq \left(\frac{n}{z_n\ln(n/z_n)}\right)^{1/(2s+2\delta+1)}. \tag{7.20}$$

We have

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

where

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

and

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

We have

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

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

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

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

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

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

$$S_2 \leq C \left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{7.21}$$

Upper bound for S_4. We have

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

Let j_2 be the integer (7.20). We have

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

where

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

We have

$$S_{4,1} \leq C \sum_{j=\tau}^{j_2} 2^j \lambda_j^2 = C\frac{z_n \ln(n/z_n)}{n} \sum_{j=\tau}^{j_2} 2^{j(1+2\delta)} \leq C\frac{z_n \ln(n/z_n)}{n} 2^{j_2(1+2\delta)}$$

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

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

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

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

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

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

$$S_4 \leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{7.22}$$

It follows from (7.18), (7.19), (7.21) and (7.22) that

$$S \leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}. \tag{7.23}$$

Combining (7.15), (7.16), (7.17) and (7.23), we have, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1,2)$ and $s > (2\delta+1)/p\}$,

$$\sup_{f_{d_*}\in B^s_{p,r}(M)} \mathbb{E}\left(\int_{-\Omega}^{\Omega} \left(\widehat{f}^{hard}(x) - f_{d_*}(x)\right)^2 dx\right) \leq C\left(\frac{z_n \ln(n/z_n)}{n}\right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 5.2 is complete. □

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

We thank the two referees and the associated editor for detailed and constructive comments which have lead to significant improvement of the paper.

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**, 309–340.

Cai, T. (1999). Adaptive Wavelet Estimation: A Block Thresholding And Oracle Inequality Approach. *Ann. Statist.*, **27**, 898–924.

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

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

Cavalier, L. and Raimondo, M. (2007). Wavelet deconvolution with noisy eigenvalues. *IEEE Trans. Signal Process.*, **55**, 2414–2424.

Chesneau, C., Fadili, J. and Starck, J.-L. (2010). Stein block thresholding for image denoising. *Appl. Comput. Harmon. Anal.*, **28**, 67–88.

Cohen, A., Daubechies, I., Jawerth, B. and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Appl. Computat. Harmon. Anal.*, **1**, 54–81.

Comte, F., Rozenholc, Y. and Taupin, M.-L. (2006). Penalized contrast estimator for density deconvolution. *Canad. J. Statist.*, **34**, 431–452.

Delaigle, A. and Gijbels, I. (2006). Estimation of boundary and discontinuity points in deconvolution problems. *Statist. Sinica*, **16**, 773–788.

Delyon, B. and Juditsky, A. (1996). On minimax wavelet estimators. *Appl. Comput. Harmon. Anal.*, **3**, 215–228.

Devroye, L. (1989). Consistent deconvolution in density estimation. *Canad. J. 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 Trans. Inform. Theory*, **48**, 734–747.

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

Johnstone, I., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004). Wavelet deconvolution in a periodic setting. *J. Roy. Stat. Soc. Ser. B*, **66**, 547–573.

Lacour, C. (2006). Rates of convergence for nonparametric deconvolution. *C. R. Math. Acad. Sci. Paris*, **342**, 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. *J. Nonparametr. Stat.*, **22**, 1–22.

Maiboroda, R. E. (1996). Estimators of components of a mixture with varying concentrations. *Ukraïn. Mat. Zh.*, **48**, 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. *Ann. Statist.*, **27**, 2033–2053.

Pensky, M. and Sapatinas, T. (2009). Functional deconvolution in a periodic setting: Uniform case. *Ann. Statist.*, **37**, 73–104.

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

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

Pokhyl'ko, D. (2005). Wavelet estimators of a density constructed from observations of a mixture. *Theory Probab. 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 J. Pure Appl. Math.*, **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 J. Math.*, **8**, 273–303.

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

Van Es, B., Gugushvili, S. and Spreij, P. (2008), Deconvolution for an atomic distribution, *Electron. J. Stat.*, **2**, 265–297.

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

Christophe Chesneau
University of Caen
LMNO, Caen, 14000
France
E-mail: chesneau@math.unicaen.fr

Paper received May 2010; revised January 2011.