HAL open science

Estimation of the derivatives of a function in a convolution regression model with random design

Christophe Chesneau, Maher Kachour

To cite this version:

Christophe Chesneau, Maher Kachour. Estimation of the derivatives of a function in a convolution regression model with random design. 2015. hal-01054877v2

HAL Id: hal-01054877

https://hal.science/hal-01054877v2

Preprint submitted on 25 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation of the derivatives of a function in a convolution regression model with random design

Christophe Chesneau

Laboratoire de Mathématiques Nicolas Oresme, Université de Caen BP 5186, F 14032 Caen Cedex, FRANCE. e-mail: christophe.chesneau@unicaen.fr

Maher Kachour

École supérieure de commerce IDRAC, 47, rue Sergent Michel Berthet CP 607, 69258 Lyon Cedex 09, FRANCE. e-mail: maher.kachour@idraclyon.com

Abstract: A convolution regression model with random design is considered. We investigate the estimation of the derivatives of an unknown function, element of the convolution product. We introduce new estimators based on wavelet methods and provide theoretical guaranties on their good performances.

AMS 2000 subject classifications: 62G07, 62G20.
Keywords and phrases: Nonparametric regression, Deconvolution, Derivatives function estimation, Wavelets, Hard thresholding.

1. Introduction

We consider the convolution regression model with random design described as follows. Let $(Y_1, X_1), \ldots, (Y_n, X_n)$ be n i.i.d. random variables defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, where

$$Y_v = (f \star g)(X_v) + \xi_v, \qquad v = 1, \ldots, n, \tag{1.1}$$

$(f \star g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt$, $f : \mathbb{R} \to \mathbb{R}$ is an unknown function, $g : \mathbb{R} \to \mathbb{R}$ is a known function, $X_1, \ldots, X_n$ are n i.i.d. random variables with common density $h : \mathbb{R} \to [0, \infty)$, and $\xi_1, \ldots, \xi_n$ are n i.i.d. random variables such that $\mathbb{E}(\xi_1) = 0$ and $\mathbb{E}(\xi_1^2) < \infty$. Throughout this paper, we assume that f, g and h are compactly supported with $\mathrm{supp}(f) = [a, b]$, $\mathrm{supp}(g) = [a', b']$, $\mathrm{supp}(h) = [a_*, b_*]$, $(a, b, a', b') \in \mathbb{R}^4$, $a < b$, $a' < b'$, $a_* = a + a'$, $b_* = b + b'$, f is m times differentiable with $m \in \mathbb{N}$, g is integrable and ordinary smooth (the precise definition is given by **(K2)** in Subsection 3.1), and X_v and ξ_v are independent for any $v = 1, \ldots, n$. We aim to estimate the unknown function f and its m-th derivative, denoted by $f^{(m)}$, from the sample $(Y_1, X_1), \ldots, (Y_n, X_n)$.

The motivation of this problem is the deconvolution of a signal f from $f \star g$ perturbed by noise and randomly observed. The function g can represent a driving force that was applied to a physical system. Such situations naturally

appear in various applied areas, as astronomy, optics, seismology and biology. The model (1.1) can also be viewed as a natural extension of some 1-periodic convolution regression models as those considered by, e.g., Cavalier and Tsybakov (2002), Pensky and Sapatinas (2010) and Loubes and Marteau (2012). In the form (1.1), it has been considered in Birke and Bissantz (2008) and Birke *et al.* (2010) with a deterministic design, and in Hildebrandt *et al.* (2014) with a random design. These last works focus on kernel methods and establish their asymptotic normality. The estimation of $f^{(m)}$, more general to $f = f^{(0)}$, is of interest to examine possible bumps and to study the convexity - concavity properties of f (see, for instance, Prakasa Rao (1983), for standard statistical models).

In this paper, we introduce new estimators for $f^{(m)}$ based on wavelet methods. Through the use of a multiresolution analysis, these methods enjoy local adaptivity against discontinuities and provide efficient estimators for a wide variety of unknown functions $f^{(m)}$. Basics on wavelet estimation can be found in, e.g., Antoniadis (1997), Härdle *et al.* (1998) and Vidakovic (1999). Results on the wavelet estimation of $f^{(m)}$ in other regression frameworks can be found in, e.g., Cai (2002), Petsa and Sapatinas (2011) and Chesneau (2014).

The first part of the study is devoted to the case where h, the common density of $X_1, \ldots, X_n$, is known. We develop a linear wavelet estimator and an adaptive nonlinear wavelet estimator. The second one uses the double hard thresholding technique introduced by Delyon and Juditsky (1996). It does not depend on the smoothness of $f^{(m)}$ in its construction ; it is adaptive. We exhibit their rates of convergence via the mean integrated squared error (MISE) and the assumption that $f^{(m)}$ belongs to Besov balls. The obtained rates of convergence coincide with existing results for the estimation of $f^{(m)}$ in the 1-periodic convolution regression models (see, for instance, Chesneau (2010))

The second part is devoted to the case where h is unknown. We construct a new linear wavelet estimator using a plug-in approach for the estimation of h. Its construction follows the idea of the "NES linear wavelet estimator" introduced by Pensky and Vidakovic (2001) in another regression context. Then we investigate its MISE properties when $f^{(m)}$ belongs to Besov balls, which naturally depend on the MISE of the considered estimator for h. Furthermore, let us mention that all our results are proved with only moments of order 2 on ξ_1, which provides another theoretical contribution to the subject.

The remaining part of this paper is organized as follows. In Section 2 we describe some basics on wavelets, Besov balls and present our wavelet estimation methodology. Section 3 is devoted to our estimators and their performances. The proofs are carried out in Section 4.

2. Preliminaries

This section is devoted to the presentation of the considered wavelet basis, the Besov balls and our wavelet estimation methodology.

2.1. Wavelet basis

Let us briefly present the wavelet basis on the interval $[a, b]$, $(a, b) \in \mathbb{R}^2$, introduced by Cohen *et al.* (1993). Let ϕ and ψ be the initial wavelet functions of the Daubechies wavelets family *db2N* with $N \geq 1$ (see, e.g., Daubechies (1992)). These functions have the distinction of being compactly supported and belong to the class $\mathcal{C}^a$ for $N > 5a$. For any $j \geq 0$ and $k \in \mathbb{Z}$, we set

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

With appropriated treatments at the boundaries, there exists an integer τ and a set of consecutive integers Λ_j of cardinality proportional to 2^j (both depending on a, b and N) such that, for any integer $\ell \geq \tau$,

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

forms an orthonormal basis of the space of squared integrable functions on $[a, b]$, i.e.,

$$\mathbb{L}^2([a, b]) = \left\{ u : [a, b] \to \mathbb{R};\ \left(\int_a^b |u(x)|^2 dx \right)^{1/2} < \infty \right\}.$$

For the case $a = 0$ and $b = 1$, τ is the smallest integer satisfying $2^\tau \geq 2N$ and $\Lambda_j = \{0, \ldots, 2^j - 1\}$.

For any integer $\ell \geq \tau$ and $u \in \mathbb{L}^2([a, b])$, we have the following wavelet expansion:

$$u(x) = \sum_{k \in \Lambda_\ell} c_{\ell,k}\phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty} \sum_{k \in \Lambda_j} d_{j,k}\psi_{j,k}(x), \qquad x \in [a, b],$$

where

$$c_{j,k} = \int_a^b u(x)\phi_{j,k}(x)dx, \qquad d_{j,k} = \int_a^b u(x)\psi_{j,k}(x)dx. \tag{2.1}$$

An interesting feature of the wavelet basis is to provide a sparse representation of u ; only few wavelet coefficients $d_{j,k}$ caracterized by a high magnitude reveal the main details of u. See, e.g., Cohen *et al.* (1993) and Mallat (2009).

2.2. Besov balls

We say that a function $u \in \mathbb{L}^2([a, b])$ belongs to the Besov ball $B^s_{p,r}(M)$ with $s > 0$, $p \geq 1$, $r \geq 1$ and $M > 0$ if there exists a constant $C > 0$ such that $c_{j,k}$ and $d_{j,k}$ (2.1) satisfy

$$2^{\tau(1/2-1/p)} \left(\sum_{k \in \Lambda_\tau} |c_{\tau,k}|^p \right)^{1/p} + \left(\sum_{j=\tau}^{\infty} \left(2^{j(s+1/2-1/p)} \left(\sum_{k \in \Lambda_j} |d_{j,k}|^p \right)^{1/p} \right)^r \right)^{1/r} \leq C,$$

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

The interest of Besov balls is to contain various kinds of homogeneous and inhomogeneous functions u. See, e.g., Meyer (1992), Donoho *et al.* (1996) and Härdle *et al.* (1998).

2.3. Wavelet estimation

Let f be the unknown function in (1.1) and $\mathcal{B}$ the considered wavelet basis taken with $N > 5m$ (to ensure that ϕ and ψ belongs to the class $\mathcal{C}^m$). Suppose that $f^{(m)}$ exists with $f^{(m)} \in \mathbb{L}^2([a,b])$.

The first step in the wavelet estimation consists in expanding $f^{(m)}$ on $\mathcal{B}$ as

$$f^{(m)}(x) = \sum_{k \in \Lambda_\ell} c_{\ell,k}^{(m)} \phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty} \sum_{k \in \Lambda_j} d_{j,k}^{(m)} \psi_{j,k}(x), \qquad x \in [a,b], \tag{2.2}$$

where $\ell \geq \tau$ and

$$c_{j,k}^{(m)} = \int_a^b f^{(m)}(x) \phi_{j,k}(x) dx, \qquad d_{j,k}^{(m)} = \int_a^b f^{(m)}(x) \psi_{j,k}(x) dx. \tag{2.3}$$

The second step is the estimation of $c_{j,k}^{(m)}$ and $d_{j,k}^{(m)}$ using $(Y_1, X_1), \ldots, (Y_n, X_n)$. The idea of the third step is to exploit the sparse representation of $f^{(m)}$ by selecting the most interesting wavelet coefficients estimators. This selection can be of different natures (truncation, thresholding, ...). Finally, we reconstruct these wavelet coefficients estimators on $\mathcal{B}$, providing an estimator $\hat{f}^{(m)}$ for $f^{(m)}$.

In this study, we evaluate the performance of $\hat{f}^{(m)}$ by studying the asymptotic properties of its MISE under the assumption that $f^{(m)} \in B_{p,r}^s(M)$. More precisely, we aim to determine the sharpest rate of convergence ω_n such that

$$\mathbb{E}\left(\int_a^b |\hat{f}^{(m)}(x) - f^{(m)}(x)|^2 dx \right) \leq C\omega_n,$$

where C denotes a constant independent of n.

3. Rates of convergence

In this section, we list the assumptions on the model, present our wavelet estimators and determine their rates of convergence under the MISE over Besov balls.

3.1. Assumptions

Let us recall that f and g are the functions in (1.1) and h is the density of X_1.

We formulate the following assumptions:

(K1) We have $f^{(q)}(a) = f^{(q)}(b) = 0$ for any $q \in \{0, \dots, m\}$, $f^{(m)} \in \mathbb{L}^2([a,b])$ and there exists a known constant $C_1 > 0$ such that $\sup_{x\in[a,b]} |f(x)| \leq C_1$.

(K2) First of all, let us define the Fourier transform of an integrable function u by

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

The notation $\overline{\cdot}$ will be used for the complex conjugate.
We have $g \in \mathbb{L}^2([a', b'])$ and there exist two constants, $c_1 > 0$ and $\delta \geq 0$, such that

$$|\mathcal{F}(g)(x)| \geq \frac{c_1}{(1+x^2)^{\delta/2}}, \qquad x \in \mathbb{R}. \tag{3.1}$$

(K3) There exists a constant $c_2 > 0$ such that

$$c_2 \leq \inf_{x\in[a_*,b_*]} h(x).$$

The assumptions **(K1)** and **(K3)** are standard in a nonparametric regression framework (see, for instance, Tsybakov (2004)). Remark that we do not need $f(a) = f(b) = 0$ for the estimation of $f = f^{(0)}$. The assumption **(K2)** is the so-called "ordinary smooth case" on g. It is common for the deconvolution - estimation of densities (see, e.g., Fan and Koo (2002) and Pensky and Vidakovic (1999)). An example of compactly supported function g satisfying **(K2)** is $g(x) = \int_1^2 y \max(1 - |x|/y, 0)\, dy$. Then $\text{supp}(g) = [-2, 2]$, $g \in \mathbb{L}^2([-2, 2])$ and **(K2)** is satisfied with $\delta = 2$ and $c_1 = \min(\inf_{x\in[-2\pi,2\pi]} |\mathcal{F}(g)(x)|, 1/(4\pi^2)) > 0$.

3.2. When h the X_i's common density is known

3.2.1. Linear wavelet estimator

We define the linear wavelet estimator $\hat{f}_1^{(m)}$ by

$$\hat{f}_1^{(m)}(x) = \sum_{k\in\Lambda_{j_0}} \hat{c}_{j_0,k}^{(m)} \phi_{j_0,k}(x), \qquad x \in [a,b], \tag{3.2}$$

where

$$\hat{c}_{j,k}^{(m)} = \frac{1}{n} \sum_{v=1}^{n} \frac{Y_v}{h(X_v)} \frac{1}{2\pi} \int_{-\infty}^{\infty} (ix)^m \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_v} dx, \tag{3.3}$$

and j_0 is an integer chosen a posteriori.

Proposition 3.1 presents an elementary property of $\hat{c}_{j,k}^{(m)}$.

Proposition 3.1. *Let $\hat{c}_{j,k}^{(m)}$ be (3.3) and $c_{j,k}^{(m)}$ be (2.3). Suppose that* **(K1)** *holds. Then we have*

$$\mathbb{E}(\hat{c}_{j,k}^{(m)}) = c_{j,k}^{(m)}.$$

Theorem 3.1 below investigates the performance of $\hat{f}_1^{(m)}$ in terms of rates of convergence under the MISE over Besov balls.

Theorem 3.1. *Suppose that* **(K1)** - **(K3)** *are satisfied and that* $f^{(m)} \in B^s_{p,r}(M)$ *with* $M > 0$, $p \geq 1$, $r \geq 1$, $s \in (\max(1/p - 1/2, 0), N)$ *and* $N > 5(m + \delta + 1)$. *Let* $\hat{f}_1^{(m)}$ *be defined by* (3.2) *with* j_0 *such that*

$$2^{j_0} = [n^{1/(2s_*+2m+2\delta+1)}], \tag{3.4}$$

$s_* = s + \min(1/2 - 1/p, 0)$ *(*$[a]$ *denotes the integer part of* a*).*

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

$$\mathbb{E}\left(\int_a^b |\hat{f}_1^{(m)}(x) - f^{(m)}(x)|^2 dx\right) \leq C n^{-2s_*/(2s_*+2m+2\delta+1)}.$$

Note that the rate of convergence $n^{-2s_*/(2s_*+2m+2\delta+1)}$ corresponds to the one obtained in the estimation of $f^{(m)}$ in the 1-periodic white noise convolution model with an adapted linear wavelet estimator (see, e.g., Chesneau (2010)).

The considered estimator $\hat{f}_1^{(m)}$ depends on s (the smoothness parameter of $f^{(m)}$) ; it is not adaptive. This aspect, as well as the rate of convergence $n^{-2s_*/(2s_*+2m+2\delta+1)}$ can be improved with thresholding methods. The next paragraph is devoting to one of them : the hard thresholding method.

3.2.2. Hard thresholding wavelet estimator

Suppose that **(K2)** is satisfied. We define the hard thresholding wavelet estimator $\hat{f}_2^{(m)}$ by

$$\hat{f}_2^{(m)}(x) = \sum_{k \in \Lambda_\tau} \hat{c}_{\tau,k}^{(m)} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k \in \Lambda_j} \hat{d}_{j,k}^{(m)} \mathbf{1}_{\left\{|\hat{d}_{j,k}^{(m)}| \geq \kappa \lambda_j\right\}} \psi_{j,k}(x), \tag{3.5}$$

$x \in [a, b]$, where $\hat{c}_{j,k}^{(m)}$ is defined by (3.3),

$$\hat{d}_{j,k}^{(m)} = \frac{1}{n} \sum_{v=1}^{n} \frac{Y_v}{h(X_v)} \frac{1}{2\pi} \int_{-\infty}^{\infty} (ix)^m \frac{\overline{\mathcal{F}(\psi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_v} dx \times$$
$$\mathbf{1}_{\left\{\left|\frac{Y_v}{h(X_v)} \frac{1}{2\pi} \int_{-\infty}^{\infty} (ix)^m \frac{\overline{\mathcal{F}\left(\psi_{j,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixX_v} dx\right| \leq \varsigma_j\right\}}, \tag{3.6}$$

$\mathbf{1}$ is the indicator function, $\kappa > 0$ is a large enough constant, j_1 is the integer satisfying

$$2^{j_1} = \left[n^{1/(2m+2\delta+1)}\right],$$

δ refers to (3.1),

$$\varsigma_j = \theta_\psi 2^{mj} 2^{\delta j} \sqrt{\frac{n}{\ln n}}, \qquad \lambda_j = \theta_\psi 2^{mj} 2^{\delta j} \sqrt{\frac{\ln n}{n}}$$

and

$$\theta_\psi = \sqrt{\frac{1}{\pi c_2 c_1^2}\left(C_1^2\left(\int_{-\infty}^{\infty}|g(x)|dx\right)^2 + \mathbb{E}(\xi_1^2)\right)\int_{-\infty}^{\infty} x^m(1+x^2)^\delta|\mathcal{F}(\psi)(x)|^2dx}.$$

The construction of $\hat{f}_2^{(m)}$ uses the double hard thresholding technique introduced by Delyon and Juditsky (1996) and recently improved by Chaubey *et al.* (2014). The main interest of the thresholding using λ_j is to make $\hat{f}_2^{(m)}$ adaptive; the construction (and performance) of $\hat{f}_2^{(m)}$ does not depend on the knowledge of the smoothness of $f^{(m)}$. The role of the thresholding using ς_j in (3.6) is to relax some usual restrictions on the model. To be more specific, it enables us to only suppose that ξ_1 admits finite moments of order 2 (with known $\mathbb{E}(\xi_1^2)$ or a known upper bound of $\mathbb{E}(\xi_1^2)$), relaxing the standard assumption $\mathbb{E}(|\xi_1|^k) < \infty$, for any $k \in \mathbb{N}$.

Further details on the constructions of hard thresholding wavelet estimators can be found in, e.g., Donoho and Johnstone (1994, 1995), Donoho *et al.* (1995, 1996), Delyon and Juditsky (1996) and Härdle *et al.* (1998).

Theorem 3.2 below investigates the performance of $\hat{f}_2^{(m)}$ in terms of rates of convergence under the MISE over Besov balls.

Theorem 3.2. *Suppose that* **(K1)** - **(K3)** *are satisfied and that* $f^{(m)} \in B_{p,r}^s(M)$ *with* $M > 0$, $r \geq 1$, $\{p \geq 2,\ s \in (0, N)\}$ *or* $\{p \in [1, 2),\ s \in ((2m+2\delta+1)/p, N)\}$ *and* $N > 5(m+\delta+1)$. *Let* $\hat{f}_2^{(m)}$ *be defined by* (3.5). *Then there exists a constant* $C > 0$ *such that*

$$\mathbb{E}\left(\int_a^b |\hat{f}_2^{(m)}(x) - f^{(m)}(x)|^2dx\right) \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2m+2\delta+1)}.$$

The proof of Theorem 3.2 is an application of a general result established by (Chaubey *et al.*, 2014, Theorem 6.1). Let us mention that $(\ln n/n)^{2s/(2s+2m+2\delta+1)}$ corresponds to the rate of convergence obtained in the estimation of $f^{(m)}$ in the 1-periodic white noise convolution model with an adapted hard thresholding wavelet estimator (see, e.g., Chesneau (2010)). In the case $m = 0$ and $\delta = 0$, this rate of convergence becomes the optimal one in the minimax sense for the standard density - regression estimation problems (see Härdle *et al.* (1998)).

In comparison to Theorem 3.1, note that

- for the case $p \geq 2$ corresponding to the homogeneous zone of Besov balls, $(\ln n/n)^{2s/(2s+2m+2\delta+1)}$ is equal to the rate of convergence attained by $\hat{f}_1^{(m)}$ up to a logarithmic term,
- for the case $p \in [1, 2)$ corresponding to the inhomogeneous zone of Besov balls, it is significantly better in terms of power.

3.3. When h the X_i's common density is unknown

In the case where h is unknown, we propose a plug-in technique which consists in estimating h in the construction of $\hat{f}_1^{(m)}$ (3.2). This yields the linear wavelet

estimator $\hat{f}_3^{(m)}$ defined by

$$\hat{f}_3^{(m)}(x) = \sum_{k \in \Lambda_{j_2}} \tilde{c}_{j_2,k}^{(m)} \phi_{j_2,k}(x), \qquad x \in [a,b], \tag{3.7}$$

where

$$\tilde{c}_{j,k}^{(m)} = \frac{1}{a_n} \sum_{v=1}^{a_n} \frac{Y_v}{\hat{h}(X_v)} \mathbf{1}_{\{|\hat{h}(X_v)| \geq c_2/2\}} \frac{1}{2\pi} \int_{-\infty}^{\infty} (ix)^m \frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_v} dx,$$

$a_n = [n/2]$, j_2 is an integer chosen a posteriori, c_2 refers to **(K3)** and $\hat{h}$ is an estimator of h constructed from the random variables $U_n = (X_{a_n+1}, \ldots, X_n)$.

There are numerous possibilities for the choice of $\hat{h}$. For instance, $\hat{h}$ can be a kernel density estimator or a wavelet density estimator (see, e.g., Donoho *et al.* (1996), Härdle *et al.* (1998) and Juditsky and Lambert-Lacroix (2004)).

The estimator $\hat{f}_3^{(m)}$ is derived to the "NES linear wavelet estimator" introduced by Pensky and Vidakovic (2001) and recently revisited in a more simple form by Chesneau (2014).

Theorem 3.3 below determines an upper bound of the MISE of $\hat{f}_3^{(m)}$.

Theorem 3.3. *Suppose that* **(K1)** - **(K3)** *are satisfied,* $h \in \mathbb{L}^2([a_*, b_*])$ *and that* $f^{(m)} \in B_{p,r}^s(M)$ *with* $M > 0$, $p \geq 1$, $r \geq 1$, $s \in (\max(1/p - 1/2, 0), N)$ *and* $N > 5(m + \delta + 1)$. *Let* $\hat{f}_3^{(m)}$ *be defined by* (3.7) *with* j_2 *such that* $2^{j_2} \leq n$. *Then there exists a constant* $C > 0$ *such that*

$$\begin{aligned} &\mathbb{E}\left(\int_a^b |\hat{f}_3^{(m)}(x) - f^{(m)}(x)|^2 dx\right) \\ &\quad \leq \quad C\left(2^{(2m+2\delta+1)j_2} \max\left(\mathbb{E}\left(\int_{a_*}^{b_*} |\hat{h}(x) - h(x)|^2 dx\right), \frac{1}{n}\right) + 2^{-2j_2 s_*}\right), \end{aligned}$$

with $s_* = s + \min(1/2 - 1/p, 0)$.

The proof follows the idea of (Chesneau, 2014, Theorem 3) and uses technical operations on Fourier transforms.

From Theorem 3.3,

- if we chose $\hat{h} = h$ and $j_2 = j_0$ (3.4), we obtain Theorem 3.1,
- if $\hat{h}$ and h satisfy : there exists $\upsilon \in [0,1]$ and a constant $C > 0$ such that

$$\mathbb{E}\left(\int_{a_*}^{b_*} |\hat{h}(x) - h(x)|^2 dx\right) \leq C n^{-\upsilon},$$

  then, the optimal integer j_2 is such that $2^{j_2} = [n^{\upsilon/(2s_*+2m+2\delta+1)}]$ and we obtain the following rate of convergence for $\hat{f}_3^{(m)}$:

$$\mathbb{E}\left(\int_a^b |\hat{f}_3^{(m)}(x) - f^{(m)}(x)|^2 dx\right) \leq C n^{-2s_*\upsilon/(2s_*+2m+2\delta+1)}.$$

Naturally the estimation of h has a negative impact on the performance of $\hat{f}_3^{(m)}$. In particular, if $h \in B^{s'}_{p',r'}(M')$, then the standard density linear wavelet estimator $\hat{h}$ attains the rate of convergence $n^{-\upsilon}$ with $\upsilon = 2s_\circ/(2s_\circ+1)$, $s_\circ = s' + \min(1/2 - 1/p', 0)$ (and it is optimal in the minimax sense for $p' \geq 2$, see Härdle *et al.* (1998)). With this choice, the rate of convergence for $\hat{f}_3^{(m)}$ becomes $n^{-4s_*s_\circ/(2s_\circ+1)(2s_*+2m+2\delta+1)}$. Let us mention that $\hat{f}_3^{(m)}$ is not adaptive since it depends of s. However, $\hat{f}_3^{(m)}$ remains an acceptable first approach for the estimation of $f^{(m)}$ with unknown h.

Conclusion and perspectives. This study considers the estimation of $f^{(m)}$ from (1.1). According to the knowledge of h or not, we propose wavelet methods and prove that they attain fast rates of convergence under the MISE over Besov balls. Among the perspectives of this work, we retain :

- The relaxation of the assumption **(K2)**, perhaps by considering : **(K2')** : There exist four constants, $C_1 > 0$, $\omega \in \mathbb{N}$, $\eta > 0$ and $\delta \geq 0$, such that

$$|\mathcal{F}(g)(x)|^{-1} \leq C_1 |\sin\left(\frac{\pi x}{\eta}\right)|^{-\omega}(1+|x|)^{\delta}, \qquad x \in \mathbb{R}.$$

  This condition was first introduced by Delaigle and Meister (2011) in a context of deconvolution - estimation of function. It implies **(K2)** and has the advantage to consider some functions g having zeros in Fourier transform domain as numerous kinds of compactly supported functions.
- The construction of an adaptive version of $\hat{f}_3^{(m)}$ through the use of a thresholding method.
- The extension of our results to the $\mathbb{L}^p$ risk with $p \geq 1$.

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

4. Proofs

In this section, C denotes any constant that does not depend on j, k and n. Its value may change from one term to another and may depend on ϕ or ψ.

Proof of Proposition 3.1. By the independence between X_1 and ξ_1, $\mathbb{E}(\xi_1) = 0$, $\sup(f \star g) = \mathrm{supp}(h) = [a_*, b_*]$ and $\mathcal{F}(f \star g)(x) = \mathcal{F}(f)(x)\mathcal{F}(g)(x)$, we have

$$\begin{aligned}
&\mathbb{E}\left(\frac{Y_1}{h(X_1)}e^{-ixX_1}\right) \\
&= \mathbb{E}\left(\frac{(f\star g)(X_1)}{h(X_1)}e^{-ixX_1}\right) + \mathbb{E}(\xi_1)\mathbb{E}\left(\frac{1}{h(X_1)}e^{-ixX_1}\right) \\
&= \mathbb{E}\left(\frac{(f\star g)(X_1)}{h(X_1)}e^{-ixX_1}\right) = \int_{a_*}^{b_*}\frac{(f\star g)(y)}{h(y)}e^{-ixy}h(y)dy \\
&= \int_{-\infty}^{\infty}(f\star g)(y)e^{-ixy}dy = \mathcal{F}(f\star g)(x) = \mathcal{F}(f)(x)\mathcal{F}(g)(x). \qquad (4.1)
\end{aligned}$$

It follows from **(K1)** and m integration by parts that $(ix)^m\mathcal{F}(f)(x) = \mathcal{F}(f^{(m)})(x)$. Using this equality, (4.1) and the Parseval identity, we obtain

$$\begin{aligned}
\mathbb{E}(\hat{c}_{j,k}^{(m)}) &= \mathbb{E}\left(\frac{Y_1}{h(X_1)}\frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_1}dx\right)\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)}\mathbb{E}\left(\frac{Y_1}{h(X_1)}e^{-ixX_1}\right)dx\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\frac{\overline{\mathcal{F}(\phi_{j,k})}(x)}{\mathcal{F}(g)(x)}\mathcal{F}(f)(x)\mathcal{F}(g)(x)dx\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\mathcal{F}(f)(x)\overline{\mathcal{F}(\phi_{j,k})}(x)dx\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{F}(f^{(m)})(x)\overline{\mathcal{F}(\phi_{j,k})}(x)dx = \int_a^b f^{(m)}(x)\phi_{j,k}(x)dx = c_{j,k}^{(m)}.
\end{aligned}$$

Proposition 3.1 is proved.

□

Proof of Theorem 3.1. We expand the function $f^{(m)}$ on $\mathcal{B}$ as (2.2) at the level $\ell = j_0$. Since $\mathcal{B}$ forms an orthonormal basis of $\mathbb{L}^2([a,b])$, we get

$$\mathbb{E}\left(\int_a^b|\hat{f}_1^{(m)}(x) - f^{(m)}(x)|^2dx\right) = \sum_{k\in\Lambda_{j_0}}\mathbb{E}\left(|\hat{c}_{j_0,k}^{(m)} - c_{j_0,k}^{(m)}|^2\right) + \sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j}(d_{j,k}^{(m)})^2. \tag{4.2}$$

Using Proposition 3.1, $(Y_1,X_1),\ldots,(Y_n,X_n)$ are i.i.d., the inequalities : $\mathbb{V}(D) \leq \mathbb{E}(|D|^2)$ for any random complex variable D and $(x+y)^2 \leq 2(x^2+y^2)$, $(x,y)\in\mathbb{R}^2$, and **(K1)** and **(K3)**, we have

$$\begin{aligned}
&\mathbb{E}\left(|\hat{c}_{j_0,k}^{(m)} - c_{j_0,k}^{(m)}|^2\right) = \mathbb{V}\left(\hat{c}_{j_0,k}^{(m)}\right)\\
&= \frac{1}{n}\mathbb{V}\left(\frac{Y_1}{h(X_1)}\frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\frac{\overline{\mathcal{F}(\phi_{j_0,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_1}dx\right)\\
&\leq \frac{1}{(2\pi)^2n}\mathbb{E}\left(\left|\frac{Y_1}{h(X_1)}\int_{-\infty}^{\infty}x^m\frac{\overline{\mathcal{F}(\phi_{j_0,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_1}dx\right|^2\right)\\
&\leq \frac{2}{(2\pi)^2n}\mathbb{E}\left(\frac{((f\star g)(X_1))^2+\xi_1^2}{(h(X_1))^2}\left|\int_{-\infty}^{\infty}x^m\frac{\overline{\mathcal{F}(\phi_{j_0,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_1}dx\right|^2\right)\\
&\leq \frac{1}{n}\frac{2}{(2\pi)^2c_2}\left(C_1^2\left(\int_{-\infty}^{\infty}|g(x)|dx\right)^2 + \mathbb{E}(\xi_1^2)\right)\times\\
&\quad\mathbb{E}\left(\frac{1}{h(X_1)}\left|\int_{-\infty}^{\infty}x^m\frac{\overline{\mathcal{F}(\phi_{j_0,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_1}dx\right|^2\right).
\end{aligned} \tag{4.3}$$

The Parseval identity yields

$$\begin{aligned}
&\mathbb{E}\left(\frac{1}{h(X_1)}\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}\left(\phi_{j_0,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx\right|^2\right) \\
&= \int_{a_*}^{b_*} \frac{1}{h(y)}\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}\left(\phi_{j_0,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixy} dx\right|^2 h(y)dy \\
&\leq \int_{-\infty}^{\infty}\left|\mathcal{F}\left(x^m \frac{\overline{\mathcal{F}\left(\phi_{j_0,k}\right)}(x)}{\mathcal{F}(g)(x)}\right)(y)\right|^2 dy = 2\pi \int_{-\infty}^{\infty}\left|x^m \frac{\overline{\mathcal{F}\left(\phi_{j_0,k}\right)}(x)}{\mathcal{F}(g)(x)}\right|^2 dx.
\end{aligned} \tag{4.4}$$

Using **(K2)**, $\left|\mathcal{F}\left(\phi_{j_0,k}\right)(x)\right| = 2^{-j_0/2}\left|\mathcal{F}\left(\phi\right)(x/2^{j_0})\right|$ and a change of variables, we obtain

$$\begin{aligned}
\int_{-\infty}^{\infty}\left|x^m \frac{\overline{\mathcal{F}\left(\phi_{j_0,k}\right)}(x)}{\mathcal{F}(g)(x)}\right|^2 dx &\leq \frac{1}{c_1^2}\int_{-\infty}^{\infty} x^{2m}(1+x^2)^{\delta}\left|\mathcal{F}\left(\phi_{j_0,k}\right)(x)\right|^2 dx \\
&= \frac{1}{c_1^2}2^{-j_0}\int_{-\infty}^{\infty} x^{2m}(1+x^2)^{\delta}\left|\mathcal{F}\left(\phi\right)(x/2^{j_0})\right|^2 dx \\
&= \frac{1}{c_1^2}\int_{-\infty}^{\infty} 2^{2j_0 m} x^{2m}(1+2^{2j_0}x^2)^{\delta}\left|\mathcal{F}\left(\phi\right)(x)\right|^2 dx \\
&\leq \frac{1}{c_1^2}2^{(2m+2\delta)j_0}\int_{-\infty}^{\infty} x^{2m}(1+x^2)^{\delta}\left|\mathcal{F}\left(\phi\right)(x)\right|^2 dx. \\
&\leq C2^{(2m+2\delta)j_0}.
\end{aligned} \tag{4.5}$$

(Let us mention that $\int_{-\infty}^{\infty} x^{2m}(1+x^2)^{\delta}\left|\mathcal{F}\left(\phi\right)(x)\right|^2 dx$ is finite thanks to $N > 5(m+\delta+1)$).

Putting (4.3), (4.4) and (4.5) together, we have

$$\mathbb{E}\left(|\hat{c}_{j_0,k}^{(m)} - c_{j_0,k}^{(m)}|^2\right) \leq C2^{(2m+2\delta)j_0}\frac{1}{n}.$$

For the integer j_0 satisfying (3.4), it holds

$$\sum_{k\in\Lambda_{j_0}} \mathbb{E}\left(|\hat{c}_{j_0,k}^{(m)} - c_{j_0,k}^{(m)}|^2\right) \leq C2^{(2m+2\delta+1)j_0}\frac{1}{n} \leq Cn^{-2s_*/(2s_*+2m+2\delta+1)}. \tag{4.6}$$

Let us now bound the last term in (4.2). Since $f^{(m)} \in B_{p,r}^{s}(M) \subseteq B_{2,\infty}^{s_*}(M)$ (see (Härdle *et al.*, 1998, Corollary 9.2)), we obtain

$$\sum_{j=j_0}^{\infty}\sum_{k\in\Lambda_j}(d_{j,k}^{(m)})^2 \leq C2^{-2j_0 s_*} \leq Cn^{-2s_*/(2s_*+2m+2\delta+1)}. \tag{4.7}$$

Owing to (4.2), (4.6) and (4.7), we have

$$\mathbb{E}\left(\int_a^b |\hat{f}_1^{(m)}(x) - f^{(m)}(x)|^2 dx\right) \leq Cn^{-2s_*/(2s_*+2m+2\delta+1)}.$$

Theorem 3.1 is proved.

□

Proof of Theorem 3.2. Observe that, for $\gamma \in \{\phi, \psi\}$, any integer $j \geq \tau$ and $k \in \Lambda_j$,

(a1) using arguments similar to those in Proposition 3.1, we obtain

$$\mathbb{E}\left(\frac{1}{n}\sum_{v=1}^{n}\frac{Y_v}{h(X_v)}\frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\frac{\overline{\mathcal{F}(\gamma_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_v}dx\right) = \int_a^b f^{(m)}(x)\gamma_{j,k}(x)dx.$$

(a2) using (4.3), (4.4) and (4.5) with γ instead of ϕ, we have

$$\begin{aligned}&\sum_{v=1}^{n}\mathbb{E}\left(\left|\frac{Y_v}{h(X_v)}\frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\frac{\overline{\mathcal{F}(\gamma_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_v}dx\right|^2\right)\\ &= n\mathbb{E}\left(\left|\frac{Y_1}{h(X_1)}\frac{1}{2\pi}\int_{-\infty}^{\infty}x^m\frac{\overline{\mathcal{F}(\gamma_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixX_1}dx\right|^2\right) \leq C_*^2 n 2^{(2m+2\delta)j},\end{aligned}$$

with $C_*^2 = (1/(\pi c_2 c_1^2))(C_1^2(\int_{-\infty}^{\infty}|g(x)|dx)^2 + \mathbb{E}(\xi_1^2))\int_{-\infty}^{\infty}x^m(1+x^2)^\delta|\mathcal{F}(\gamma)(x)|^2dx.$

Thanks to **(a1)** and **(a2)**, we can apply (Chaubey *et al.*, 2014, Theorem 6.1) (see Appendix) with $\mu_n = \upsilon_n = n$, $\sigma = m + \delta$, $\theta_\gamma = C_*$, $W_v = (Y_v, X_v)$,

$$q_v(\gamma, (y, z)) = \frac{y}{h(z)}\frac{1}{2\pi}\int_{-\infty}^{\infty}(ix)^m\frac{\overline{\mathcal{F}(\gamma_{j,k})}(x)}{\mathcal{F}(g)(x)}e^{-ixz}dx$$

and $f^{(m)} \in B_{p,r}^s(M)$ with $M > 0$, $r \geq 1$, either $\{p \geq 2 \text{ and } s \in (0, N)\}$ or $\{p \in [1, 2) \text{ and } s \in (1/p, N)\}$, we prove the existence of a constant $C > 0$ such that

$$\mathbb{E}\left(\int_a^b |\hat{f}_2^{(m)}(x) - f^{(m)}(x)|^2dx\right) \leq C\left(\frac{\ln n}{n}\right)^{2s/(2s+2m+2\delta+1)}.$$

Theorem 3.2 is proved.

□

Proof of Theorem 3.3. We expand the function $f^{(m)}$ on $\mathcal{B}$ as (2.2) at the level $\ell = j_2$. Since $\mathcal{B}$ forms an orthonormal basis of $\mathbb{L}^2([a, b])$, we get

$$\mathbb{E}\left(\int_a^b |\hat{f}_3^{(m)}(x) - f^{(m)}(x)|^2dx\right) = \sum_{k\in\Lambda_{j_2}}\mathbb{E}\left(|\tilde{c}_{j_2,k}^{(m)} - c_{j_2,k}^{(m)}|^2\right) + \sum_{j=j_2}^{\infty}\sum_{k\in\Lambda_j}(d_{j,k}^{(m)})^2. \tag{4.8}$$

Using $f^{(m)} \in B_{p,r}^s(M) \subseteq B_{2,\infty}^{s_*}(M)$ (see (Härdle *et al.*, 1998, Corollary 9.2)), we have

$$\sum_{j=j_2}^{\infty}\sum_{k\in\Lambda_j}(d_{j,k}^{(m)})^2 \leq C2^{-2j_2s_*}. \tag{4.9}$$

Let $\hat{c}^{(m)}_{j_2,k}$ be (3.3) with $n = a_n$ and $j = j_2$. The elementary inequality: $(x+y)^2 \leq 2(x^2+y^2)$, $(x,y) \in \mathbb{R}^2$, yields

$$\sum_{k\in\Lambda_{j_2}} \mathbb{E}\left(|\tilde{c}^{(m)}_{j_2,k} - c^{(m)}_{j_2,k}|^2\right) \leq 2(S_1+S_2), \tag{4.10}$$

where

$$S_1 = \sum_{k\in\Lambda_{j_2}} \mathbb{E}\left(|\tilde{c}^{(m)}_{j_2,k} - \hat{c}^{(m)}_{j_2,k}|^2\right), \qquad S_2 = \sum_{k\in\Lambda_{j_2}} \mathbb{E}\left(|\hat{c}^{(m)}_{j_2,k} - c^{(m)}_{j_2,k}|^2\right).$$

Upper bound for S_2. Proceeding as in (4.6), we get

$$S_2 \leq C2^{(2m+2\delta+1)j_2}\frac{1}{a_n} \leq C2^{(2m+2\delta+1)j_2}\frac{1}{n}. \tag{4.11}$$

Upper bound for S_1. The triangular inequality gives

$$\begin{aligned}
&\left|\tilde{c}^{(m)}_{j_2,k} - \hat{c}^{(m)}_{j_2,k}\right| \\
&\leq \frac{1}{(2\pi)a_n}\sum_{v=1}^{a_n}|Y_v|\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_v}dx\right| \times \\
&\quad \left|\frac{1}{\hat{h}(X_v)}\mathbf{1}_{\{|\hat{h}(X_v)|\geq c_2/2\}} - \frac{1}{h(X_v)}\right|.
\end{aligned}$$

Owing to the triangular inequality, the indicator function, **(K3)**, $\left\{|\hat{h}(X_v)| < c_2/2\right\} \subseteq \left\{|\hat{h}(X_v) - h(X_v)| > c_2/2\right\}$ and the Markov inequality, we have

$$\begin{aligned}
&\left|\frac{1}{\hat{h}(X_v)}\mathbf{1}_{\{|\hat{h}(X_v)|\geq c_2/2\}} - \frac{1}{h(X_v)}\right| \\
&= \left|\frac{1}{h(X_v)}\left(\left(\frac{h(X_v)-\hat{h}(X_v)}{\hat{h}(X_v)}\right)\mathbf{1}_{\{|\hat{h}(X_v)|\geq c_2/2\}} - \mathbf{1}_{\{|\hat{h}(X_v)|<c_2/2\}}\right)\right| \\
&\leq \frac{1}{h(X_v)}\left(\frac{2}{c_2}|\hat{h}(X_v)-h(X_v)| + \mathbf{1}_{\{|\hat{h}(X_v)-h(X_v)|>c_2/2\}}\right) \\
&\leq \frac{4}{c_2}\frac{|\hat{h}(X_v)-h(X_v)|}{h(X_v)}.
\end{aligned}$$

Therefore

$$\left|\tilde{c}^{(m)}_{j_2,k} - \hat{c}^{(m)}_{j_2,k}\right| \leq CA_{j_2,k,n},$$

where

$$A_{j,k,n} = \frac{1}{a_n}\sum_{v=1}^{a_n}|Y_v|\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_v}dx\right| \frac{|\hat{h}(X_v)-h(X_v)|}{h(X_v)}.$$

Let us now consider $U_n = (X_{a_n+1}, \ldots, X_n)$. For any complex random variable D, we have the equality:

$$\mathbb{E}(D^2) = \mathbb{E}(\mathbb{E}(D^2|U_n)) = \mathbb{E}(\mathbb{V}(D|U_n)) + \mathbb{E}((\mathbb{E}(D|U_n))^2),$$

where $\mathbb{E}(D|U_n)$ denotes the expectation of D conditionally to U_n and $\mathbb{V}(D|U_n)$, the variance of D conditionally to U_n. Therefore

$$S_1 \leq C \sum_{k \in \Lambda_{j_2}} \mathbb{E}(A^2_{j_2,k,n}) = C(W_{j_2,n} + Z_{j_2,n}), \tag{4.12}$$

where

$$W_{j_2,n} = \sum_{k \in \Lambda_{j_2}} \mathbb{E}\left(\mathbb{V}\left(A_{j_2,k,n}|U_n\right)\right), \qquad Z_{j_2,n} = \sum_{k \in \Lambda_{j_2}} \mathbb{E}\left(\left(\mathbb{E}\left(A_{j_2,k,n}|U_n\right)\right)^2\right).$$

Let us now observe that, owing to the independence of $(Y_1, X_1), \ldots, (Y_n, X_n)$, the random variables $|Y_1| \left| \int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx \right| |\hat{h}(X_1) - h(X_1)|/h(X_1), \ldots,$ $|Y_{a_n}| \left| \int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_{a_n}} dx \right| |\hat{h}(X_{a_n}) - h(X_{a_n})|/h(X_{a_n})$ conditionally to U_n are independent. Using this property with the inequalities: $\mathbb{V}(D|U_n) \leq \mathbb{E}(D^2|U_n)$ for any complex random variable D and $(x+y)^2 \leq 2(x^2+y^2)$, $(x,y) \in \mathbb{R}^2$, the independence between X_1 and ξ_1, **(K1)** and **(K3)**, we get

$$\begin{aligned}
\mathbb{V}\left(A_{j_2,k,n}|U_n\right) &= \frac{1}{a_n} \mathbb{V}\left(|Y_1| \left| \int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}\left(\phi_{j_2,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx \right| \frac{|\hat{h}(X_1) - h(X_1)|}{h(X_1)} \middle| U_n \right) \\
&\leq \frac{1}{a_n} \mathbb{E}\left(Y_1^2 \left| \int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}\left(\phi_{j_2,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx \right|^2 \left| \frac{\hat{h}(X_1) - h(X_1)}{h(X_1)} \right|^2 \middle| U_n \right) \\
&\leq \frac{1}{a_n} \frac{2}{c_2} \left(C_1^2 \left(\int_{-\infty}^{\infty} |g(x)| dx \right)^2 + \mathbb{E}(\xi_1^2) \right) \times \\
&\quad \mathbb{E}\left(\left| \int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}\left(\phi_{j_2,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx \right|^2 \frac{|\hat{h}(X_1) - h(X_1)|^2}{h(X_1)} \middle| U_n \right) \\
&= \frac{2}{c_2} \left(C_1^2 \left(\int_{-\infty}^{\infty} |g(x)| dx \right)^2 + \mathbb{E}(\xi_1^2) \right) \times \\
&\quad \frac{1}{a_n} \int_{a_*}^{b_*} \left| \int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}\left(\phi_{j_2,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixy} dx \right|^2 \frac{|\hat{h}(y) - h(y)|^2}{h(y)} h(y) dy \\
&\leq C \frac{1}{n} \int_{a_*}^{b_*} \left| \int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}\left(\phi_{j_2,k}\right)}(x)}{\mathcal{F}(g)(x)} e^{-ixy} dx \right|^2 |\hat{h}(y) - h(y)|^2 dy.
\end{aligned}$$

Owing to **(K2)**, $|\mathcal{F}(\phi_{j_2,k})(x)| = 2^{-j_2/2}\left|\mathcal{F}(\phi)(x/2^{j_2})\right|$ and a change of variables, we obtain

$$\begin{aligned}
&\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixy} dx\right| \leq \int_{-\infty}^{\infty} |x|^m \frac{|\mathcal{F}(\phi_{j_2,k})(x)|}{|\mathcal{F}(g)(x)|} dx \\
&\leq \frac{1}{c_1}\int_{-\infty}^{\infty} |x|^m (1+x^2)^{\delta/2} |\mathcal{F}(\phi_{j_2,k})(x)| dx \\
&= \frac{1}{c_1} 2^{-j_2/2}\int_{-\infty}^{\infty} |x|^m (1+x^2)^{\delta/2} \left|\mathcal{F}(\phi)(x/2^{j_2})\right| dx \\
&= \frac{1}{c_1} 2^{j_2/2}\int_{-\infty}^{\infty} 2^{j_2 m}|x|^m (1+2^{2j_2}x^2)^{\delta/2} |\mathcal{F}(\phi)(x)| dx \\
&\leq \frac{1}{c_1} 2^{(m+\delta+1/2)j_2}\int_{-\infty}^{\infty} |x|^m (1+x^2)^{\delta/2} |\mathcal{F}(\phi)(x)| dx. \\
&\leq C 2^{(m+\delta+1/2)j_2}.
\end{aligned}$$

Therefore, using $\mathrm{Card}(\Lambda_{j_2}) \leq C2^{j_2}$ and $2^{j_2} \leq n$, we obtain

$$\begin{aligned}
W_{j_2,n} &\leq C 2^{(2m+2\delta+1)j_2} 2^{j_2} \frac{1}{n}\mathbb{E}\left(\int_{a_*}^{b_*} |\hat{h}(y)-h(y)|^2 dy\right) \\
&\leq C 2^{(2m+2\delta+1)j_2}\mathbb{E}\left(\int_{a_*}^{b_*} |\hat{h}(y)-h(y)|^2 dy\right). \qquad (4.13)
\end{aligned}$$

Now, by the Hölder inequality for conditional expectations, arguments similar to (4.3), (4.4) and (4.5), we get

$$\begin{aligned}
&\mathbb{E}(A_{j_2,k,n}|U_n) = \mathbb{E}\left(|Y_1|\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx\right| \frac{|\hat{h}(X_1)-h(X_1)|}{h(X_1)}\middle| U_n\right) \\
&\leq \left(\mathbb{E}\left(\frac{Y_1^2}{h(X_1)}\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx\right|^2 \middle| U_n\right)\right)^{1/2} \times \\
&\quad \left(\mathbb{E}\left(\frac{|\hat{h}(X_1)-h(X_1)|^2}{h(X_1)}\middle| U_n\right)\right)^{1/2} \\
&= \left(\mathbb{E}\left(\frac{Y_1^2}{h(X_1)}\left|\int_{-\infty}^{\infty} x^m \frac{\overline{\mathcal{F}(\phi_{j_2,k})}(x)}{\mathcal{F}(g)(x)} e^{-ixX_1} dx\right|^2\right)\right)^{1/2} \times \\
&\quad \left(\int_{a_*}^{b_*} \frac{|\hat{h}(y)-h(y)|^2}{h(y)} h(y) dy\right)^{1/2} \\
&\leq C 2^{(m+\delta)j_2}\left(\int_{a_*}^{b_*} |\hat{h}(y)-h(y)|^2 dy\right)^{1/2}.
\end{aligned}$$

Hence

$$Z_{j_2,n} \leq C2^{(2m+2\delta+1)j_2}\mathbb{E}\left(\int_{a_*}^{b_*}|\hat{h}(y)-h(y)|^2dy\right). \tag{4.14}$$

It follows from (4.12), (4.13) and (4.14) that

$$S_1 \leq C2^{(2m+2\delta+1)j_2}\mathbb{E}\left(\int_{a_*}^{b_*}|\hat{h}(y)-h(y)|^2dy\right). \tag{4.15}$$

Putting (4.10), (4.11) and (4.15) together, we get

$$\begin{aligned}&\sum_{k\in\Lambda_{j_2}}\mathbb{E}\left(|\tilde{c}_{j_2,k}^{(m)}-c_{j_2,k}^{(m)}|^2\right)\\ &\leq\quad C2^{(2m+2\delta+1)j_2}\max\left(\mathbb{E}\left(\int_{a_*}^{b_*}|\hat{h}(y)-h(y)|^2dy\right),\frac{1}{n}\right).\end{aligned} \tag{4.16}$$

Combining (4.8), (4.9) and (4.16), we obtain the desired result, i.e.,

$$\begin{aligned}&\mathbb{E}\left(\int_a^b|\hat{f}_3^{(m)}(x)-f^{(m)}(x)|^2dx\right)\\ &\leq\quad C\left(2^{(2m+2\delta+1)j_2}\max\left(\mathbb{E}\left(\int_{a_*}^{b_*}|\hat{h}(y)-h(y)|^2dy\right),\frac{1}{n}\right)+2^{-2j_2s_*}\right).\end{aligned}$$

Theorem 3.3 is proved.

□

Appendix

Let us now present in details the general result of (Chaubey *et al.*, 2014, Theorem 6.1) used in the proof of Theorem 3.2.

We consider the wavelet basis presented in Section 2 and a general form of the hard thresholding wavelet estimator denoted by $\hat{f}_H$ for estimating an unknown function $f \in \mathbb{L}^2([a,b])$ from n independent random variables $W_1, \ldots, W_n$:

$$\hat{f}_H(x) = \sum_{k \in \Lambda_\tau} \hat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k \in \Lambda_j} \hat{\beta}_{j,k} \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \kappa \vartheta_j\}} \psi_{j,k}(x), \tag{4.17}$$

where

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

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

$$\varsigma_j = \theta_\psi 2^{\sigma j} \frac{\upsilon_n}{\sqrt{\mu_n \ln \mu_n}}, \qquad \vartheta_j = \theta_\psi 2^{\sigma j} \sqrt{\frac{\ln \mu_n}{\mu_n}},$$

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

$$2^{j_1} = [\mu_n^{1/(2\sigma+1)}].$$

Here, we suppose that there exist

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

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

(A1) any integer $j \geq \tau$ and any $k \in \Lambda_j$,

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

(A2) there exist two constants, $\theta_\gamma > 0$ and $\sigma \geq 0$, such that, for any integer $j \geq \tau$ and any $k \in \Lambda_j$,

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

Let $\hat{f}_H$ be (4.17) *under (A1) and (A2). Suppose that $f \in B^s_{p,r}(M)$ with $r \geq 1$, $\{p \geq 2$ and $s \in (0, N)\}$ or $\{p \in [1, 2)$ and $s \in ((2\sigma + 1)/p, N)\}$. Then there exists a constant $C > 0$ such that*

$$\mathbb{E}\left(\int_a^b |\hat{f}_H(x) - f(x)|^2 dx\right) \leq C \left(\frac{\ln \mu_n}{\mu_n}\right)^{2s/(2s+2\sigma+1)}.$$

□

Acknowledgments

The author is thankful to the reviewers for their comments which have helped in improving the presentation.

References

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

Birke, M. and Bissantz, N. (2008). Asymptotic normality and confidence intervals for inverse regression models with convolution-type operators, *Journal of Multivariate Statistics*, 100, 10, 2364-2375.

Birke, M., Bissantz, N. and Holzmann, H. (2010). Confidence bands for inverse regression models, *Inverse Problems*, 26, 115020.

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

Cavalier, L. and Tsybakov, A. (2002). Sharp adaption for inverse problems with random noise, *Prob. Theory Related Fields*, 123, 323-354.

Chaubey, Y.P., Chesneau, C. and Doosti, H. (2014). Adaptive wavelet estimation of a density from mixtures under multiplicative censoring, preprint: http://hal.archives-ouvertes.fr/hal-00918069.

Chesneau, C. (2010). Wavelet estimation of the derivatives of an unknown function from a convolution model, *Current Development in Theory and Applications of Wavelets*, 4, 2, 131-151.

Chesneau, C. (2014). A note on wavelet estimation of the derivatives of a regression function in a random design setting, *International Journal of Mathematics and Mathematical Sciences*, Volume 2014, Article ID 195765, 8 pages.

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.

Daubechies, I. (1992). *Ten lectures on wavelets*, SIAM.

Delaigle, A. and Meister, A. (2011). Nonparametric function estimation under Fourier-oscillating noise, *Statistica Sinica*, 21, 1065-1092.

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

Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage, *Biometrika*, 81, 425-455.

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

Donoho, D., Johnstone, I., Kerkyacharian, G. and Picard, D. (1995). Wavelet shrinkage: Asymptopia?, *Journal of the Royal Statistical Society*, Ser. B, 57, 301-369.

Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding, *Annals of Statistics*, 24, 508-539.

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

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

Hildebrandt, T., Bissantz, N. and Dette, H. (2014). Additive inverse regression models with convolution-type operators, *Electronic Journal of Statistics*, 8, 1-40.

Juditsky, A. and Lambert-Lacroix, S. (2004). On minimax density estimation on $\mathbb{R}$, *Bernoulli*, 10, 2, 187-220.

Loubes, J.M. and C. Marteau, C. (2012). Adaptive estimation for an inverse regression model with unknown operator, *Statistics and Risk Modeling*, 29, 215-242.

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

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

Pensky, M. and Sapatinas, T. (2010). On convergence rates equivalency and sampling strategies in functional deconvolution models, *Annals of Statistics*, 38, 1793-1844.

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

Pensky, M. and Vidakovic, B. (2001). On non-equally spaced wavelet regression, *Annals of the Institute of Statistical Mathematics*, 53, 681-690.

Petsa, A. and Sapatinas, T. (2011). On the estimation of the function and its derivatives in nonparametric regression: a Bayesian estimation approach, *Sankhya, Series A*, 73, 231-244.

Prakasa Rao, B.L.S. (1983). *Nonparametric Functional Estimation*, Academic Press, Orlando.

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

Vidakovic, B. (1999). *Statistical Modeling by Wavelets*, John Wiley & Sons, Inc., New York, 384 pp.