Wavelet-based estimation of the derivatives of a function from a heteroscedastic multichannel convolution model

Fabien Navarro

Laboratoire de Mathématiques Nicolas Oresme,
CNRS-UMR 6139, Université de Caen, Campus II, 14032 Caen, France
GREYC CNRS-ENSICAEN-Université de Caen, 14050 Caen France
e-mail: fabien.navarro@greyc.ensicaen.fr

Christophe Chesneau

Laboratoire de Mathématiques Nicolas Oresme,
CNRS-UMR 6139, Université de Caen, Campus II, BP 5186
Boulevard du Maréchal Juin, F-14032 Caen, France
e-mail: chesneau@math.unicaen.fr

Jalal Fadili

GREYC CNRS-ENSICAEN-Université de Caen,
6 Bd du Maréchal Juin, 14050 Caen France
e-mail: Jalal.Fadili@greyc.ensicaen.fr

and

Taoufik Sassi

Laboratoire de Mathématiques Nicolas Oresme,
CNRS-UMR 6139, Université de Caen, Campus II, BP 5186
Boulevard du Maréchal Juin, F-14032 Caen, France
e-mail: taoufik.sassi@math.unicaen.fr

Abstract: We observe n heteroscedastic stochastic processes where, for any $v \in \{1, \ldots, n\}$, a convolution product of an unknown function f and a known function g_v is corrupted by Gaussian noise. Under a particular ordinary smooth assumption on $g_1, \ldots, g_n$, we aim to estimate the d-th derivatives of f from the observations. We consider an adaptive estimator based on a particular wavelet block thresholding: the "BlockJS estimator". Taking the mean integrated squared error (MISE), we prove that it achieves near optimal rates of convergence over a wide range of smoothness classes. The theory is illustrated with some numerical examples. Performance comparisons with some others methods existing in the literature are provided.

AMS 2000 subject classifications: Primary 62G07, 62G20; secondary 62F12.
Keywords and phrases: deconvolution, derivatives function estimation, wavelets, block thresholding, minimax.

Contents

1. Motivations

1.1. Problem statement

We observe n stochastic processes $Y_1(t), \ldots, Y_n(t)$, $t \in [0,1]$ where, for any $v \in \{1, \ldots, n\}$,

$$dY_v(t) = (f \star g_v)(t)dt + \epsilon dW_v(t), \qquad t \in [0,1], \qquad n \in \mathbb{N}^*, \tag{1}$$

$\epsilon > 0$ is the noise level, $(f \star g_v)(t) = \int_0^1 f(t-u)g_v(u)du$ denotes the convolution, $W_1(t), \ldots, W_n(t)$ are n unobserved independent standard Brownian motions, for any $v \in \{1, \ldots, n\}$, $g_v : [0,1] \to \mathbb{R}$ is a known function and $f : [0,1] \to \mathbb{R}$ is an unknown function. We assume that f and $g_1, \ldots, g_n$ belong to $\mathbb{L}^2_{per}([0,1]) = \{h;$ h is 1-periodic on $[0,1]$ and $\int_0^1 h^2(t)dt < \infty\}$. The goal is to estimate f (or an unknown quantity depending on f) from $Y_1(t), \ldots, Y_n(t)$, $t \in [0,1]$.

Remark that, when $g_1 = \cdots = g_n$, we can rewrite (1) as

$$d\widetilde{Y}(t) = (f \star g)(t)dt + \epsilon n^{-1/2} d\widetilde{W}(t), \qquad t \in [0,1], \tag{2}$$

where $\widetilde{Y}(t) = (1/n)\sum_{v=1}^n Y_v(t)$, $g = g_1$ and $\widetilde{W}(t) = (1/n^{1/2})\sum_{v=1}^n W_v(t)$ is standard Brownian motion. Then (2) becomes a standard deconvolution problem in the field of function estimation. Results on the estimation on f can be found in [6], [5], [18], [4] and [9]. When $g_1, \ldots, g_n$ are not necessary equals, the estimation of f has been investigated by [14], [21, 22]. [14], [21, 22] develop adaptive wavelet thresholding estimators (hard thresholding in [14] and block thresholding in [21, 22]) under various assumptions on $g_1, \ldots, g_n$ (typically, ordinary smooth case and supersmooth case). Moreover, [14], [21, 22] establish minimax rates of convergence under the mean integrated squared error (MISE) over Besov balls.

In this paper, considering a particular ordinary smooth case on $g_1, \ldots, g_n$, we propose some extensions of [21, 22, Theorems 1 and 2] from (1). In particular, we focus on a more general problem: estimate the d-th derivative of f: $f^{(d)}$ with $d \in \mathbb{N}$ (we set $f^{(0)} = f$). This is of interest to detect possible bumps, concavity or convexity properties of f. Such derivatives estimation problems have already been investigated from several standard nonparametric models. If we only consider wavelet methods, we refer to [10] for (2) and [24], [7] and [8] for density estimation problems. We develop an adaptive wavelet estimator $\widehat{f}$ of $f^{(d)}$. It is constructed from a periodised Meyer wavelet basis and a particular block thresholding rule known under the named of BlockJS. It can be viewed as a refinement of the one in [22, (2.9)]. Recent developments on BlockJS can be found in [3], [2], [25] and [11]. Adopting the minimax approach under the MISE of Besov balls, we investigate the upper bounds of our estimator. We prove that they are near optimal by the determination of the lower bounds. From a practical point of view, for the case $d = 0$, we prove that it gives better result than the one of [22, (2.9)]. Moreover, when $d \in \{1, 2, \ldots\}$, some numerical examples are provided.

1.2. Paper organization

The paper is organized as follows. Section 2 clarifies the assumptions made on $g_1, \ldots, g_n$. In Section 3, we present wavelets and Besov balls. The BlockJS estimator is defined in Section 4. Section 5 is devoted to the results. Simulations are set in Section 6. The proofs are postponed in Section 7.

2. On our ordinary smooth assumption

First of all, note that any function $h \in \mathbb{L}^2_{per}([0,1])$ can be represented by its Fourier series

$$h(t) = \sum_{\ell \in \mathbb{Z}} \mathcal{F}(h)(\ell) e^{2i\pi\ell t}, \qquad t \in [0,1],$$

where the equality is intended in mean-square convergence sense, and $\mathcal{F}_\ell(h)$ denotes the Fourier coefficient given by

$$\mathcal{F}(h)(\ell) = \int_0^1 h(x) e^{-2i\pi\ell x} dx, \qquad \ell \in \mathbb{Z},$$

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

In this study, we focus on the following particular ordinary smooth assumption on $g_1, \ldots, g_n$: we suppose that there exist three constants, $c_g > 0$, $C_g > 0$

and $\delta > 1$, and n positive real numbers $\sigma_1, \ldots, \sigma_n$ such that, for any $x \in \mathbb{R}$ and any $v \in \{1, \ldots, n\}$,

$$c_g \frac{1}{(1+\sigma_v^2 x^2)^{\delta/2}} \leq |\mathcal{F}(g_v)(x)| \leq C_g \frac{1}{(1+\sigma_v^2 x^2)^{\delta/2}}. \quad (3)$$

This assumption controls the decay of the Fourier coefficients of $g_1, \ldots, g_n$, and thus the smoothness of $g_1, \ldots, g_n$. It is a standard hypothesis usually adopted in the field of nonparametric estimation for deconvolution problems. See e.g. [23], [16] and [18].

Example: let $\upsilon_1, \ldots, \upsilon_n$ be n positive real numbers. For any $v \in \{1, \ldots, n\}$, consider the square integrable 1-periodic function g defined by

$$g_v(x) = \frac{1}{\upsilon_v} \sum_{m \in \mathbb{Z}} e^{-|x+m|/\upsilon_v}, \qquad x \in [0,1].$$

Then, for any $x \in \mathbb{R}$, $\mathcal{F}(g_v)(l) = 2\left(1 + 4\pi^2 l^2 \upsilon_v^2\right)^{-1}$ and (3) is satisfied with $\delta = 2$ and $\sigma_v = 2\pi\upsilon_v$.

In the sequel, we set

$$\rho_n = \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^{\delta}}. \quad (4)$$

For technical reason, we suppose that $\rho_n \geq e$.

3. Wavelets and Besov balls

3.1. Periodized Meyer Wavelets

We consider an orthonormal wavelet basis generated by dilations and translations of a "father" Meyer-type wavelet ϕ and a "mother" Meyer-type wavelet ψ. The features of such wavelets are:

- the Fourier transforms of ϕ and ψ have bounded support. More precisely, we have

$$\begin{cases} \text{supp}\,(\mathcal{F}(\phi)) \subset [-4\pi 3^{-1}, 4\pi 3^{-1}], \\ \text{supp}\,(\mathcal{F}(\psi)) \subset [-8\pi 3^{-1}, -2\pi 3^{-1}] \cup [2\pi 3^{-1}, 8\pi 3^{-1}], \end{cases} \quad (5)$$

  where supp denotes the support and, for any $h \in \mathbb{L}^2_{per}([0,1])$, $\mathcal{F}(h)$ denotes the Fourier transform of h defined by

$$\mathcal{F}(h)(x) = \int_0^1 h(x) e^{-2i\pi xy} dy, \qquad x \in \mathbb{R}.$$

- (ϕ, ψ) is r-regular for a chosen $r \in \mathbb{N}$, i.e. $\phi \in \mathcal{C}^r$, $\psi \in \mathcal{C}^r$ and, for any $u \in \{0, \ldots, r\}$,

$$\int_{-\infty}^{\infty} x^u \psi(x) dx = 0. \quad (6)$$

A consequence of (5) and (6) is that, for any $m \in \mathbb{N}$ and any $u \in \{0, \dots, r\}$,

$$\sup_{x\in\mathbb{R}}(|\phi^{(u)}(x)| \left(x^2+1\right)^m) < \infty, \qquad \sup_{x\in\mathbb{R}}(|\psi^{(u)}(x)| \left(x^2+1\right)^m) < \infty. \tag{7}$$

For the purposes of this paper, we use the periodised wavelet bases on the unit interval. For any $x \in [0,1]$, any integer j and any $k \in \{0, \dots, 2^j - 1\}$, let

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

be the elements of the wavelet basis, and

$$\phi^{per}_{j,k}(x) = \sum_{\ell\in\mathbb{Z}} \phi_{j,k}(x - l), \qquad \psi^{per}_{j,k}(x) = \sum_{\ell\in\mathbb{Z}} \psi_{j,k}(x - l),$$

their periodised versions. There exists an integer τ such that the collection ζ defined by

$$\zeta = \left\{\phi^{per}_{\tau,k}(.), k \in \{0, \dots, 2^\tau - 1\};\ \psi^{per}_{j,k}(.), \quad j \geq \tau,\ k \in \{0, \dots, 2^j - 1\}\right\}$$

constitutes an orthonormal basis of $\mathbb{L}^2_{per}([0,1])$. In what follows, the superscript "per" will be suppressed from the notations for convenience.

Then, for any $m \geq \tau$, a function $h \in \mathbb{L}^2_{per}([0,1])$ can be expanded into a wavelet series as

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

where

$$\alpha_{m,k} = \int_0^1 h(t)\overline{\phi_{m,k}}(t)dt, \qquad \beta_{j,k} = \int_0^1 h(t)\overline{\psi_{j,k}}(t)dt. \tag{8}$$

For further details about Meyer-type wavelets and wavelet decomposition, see [13], [26] and [27].

3.2. Besov balls

Let $M > 0$, $s > 0$, $p \geq 1$ and $r \geq 1$. We say that a function h belongs to the Besov ball $B^s_{p,r}(M)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients (8) satisfy

$$2^{\tau(1/2-1/p)}\left(\sum_{k=0}^{2^\tau-1}|\alpha_{\tau,k}|^p\right)^{1/p} + \left(\sum_{j=\tau}^{\infty}\left(2^{j(s+1/2-1/p)}\left(\sum_{k=0}^{2^j-1}|\beta_{j,k}|^p\right)^{1/p}\right)^r\right)^{1/r} \leq M^*. \tag{9}$$

with a smoothness parameter $s > 0$, and the norm parameters: $0 < p \leq \infty$ and $0 < r \leq \infty$. For a particular choice of parameters s, p and r, these sets contain the Hölder and Sobolev balls. See [20].

4. BlockJS estimator

We suppose that $f^{(d)} \in \mathbb{L}^2_{per}([0,1])$ and that (3) is satisfied (δ refers to this assumption). We now present the considered adaptive procedure for the estimation of $f^{(d)}$. Let j_1 and j_2 be the integers defined by

$$j_1 = \lfloor \log_2(\log \rho_n) \rfloor, \qquad j_2 = \lfloor (1/(2\delta + 2d + 1)) \log_2(\rho_n / \log \rho_n) \rfloor,$$

where, for any $a \in \mathbb{R}$, $\lfloor a \rfloor$ denotes the whole number part of a. For any $j \in \{j_1, \dots, j_2\}$, set $L = \lfloor \log \rho_n \rfloor$ and $A_j = \{1, \dots, 2^j L^{-1}\}$. For any $K \in A_j$, we consider the set

$$B_{j,K} = \{k \in \{0, \dots, 2^j - 1\}; \, (K-1)L \leq k \leq KL - 1\}.$$

We define the Block James-Stein estimator (BlockJS) by

$$\widehat{f}(x) = \sum_{k=0}^{2^{j_1}-1} \widehat{\alpha}_{j_1,k} \phi_{j_1,k}(x) + \sum_{j=j_1}^{j_2} \sum_{K \in A_j} \sum_{k \in B_{j,K}} \widehat{\beta}^*_{j,k} \psi_{j,k}(x), \qquad x \in [0,1], \tag{10}$$

where

$$\widehat{\beta}^*_{j,k} = \widehat{\beta}_{j,k} \left(1 - \frac{\lambda \epsilon^2 {\rho_n}^{-1} 2^{2j(\delta+d)}}{\frac{1}{L} \sum_{k \in B_{j,K}} |\widehat{\beta}_{j,k}|^2} \right)_+ ,$$

with, for any $a \in \mathbb{R}$, $(a)_+ = \max(a, 0)$, $\lambda > 0$,

$$\widehat{\alpha}_{j_1,k} = \frac{1}{\rho_n} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^\delta} \sum_{\ell \in \mathcal{D}_{j_1}} (2\pi i \ell)^d \frac{\overline{\mathcal{F}(\phi_{j_1,k})}(\ell)}{\mathcal{F}(g_v)(\ell)} \int_0^1 e^{-2\pi i \ell t} dY_v(t)$$

and

$$\widehat{\beta}_{j,k} = \frac{1}{\rho_n} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^\delta} \sum_{\ell \in \mathcal{C}_j} (2\pi i \ell)^d \frac{\overline{\mathcal{F}(\psi_{j,k})}(\ell)}{\mathcal{F}(g_v)(\ell)} \int_0^1 e^{-2\pi i \ell t} dY_v(t).$$

Here,

$$\mathcal{D}_{j_1} = \mathrm{supp}\left(\mathcal{F}(\phi_{j_1,0})\right) = \mathrm{supp}\left(\mathcal{F}\left(\phi_{j_1,k}\right)\right),$$
$$\mathcal{C}_j = \mathrm{supp}\left(\mathcal{F}(\psi_{j,0})\right) = \mathrm{supp}\left(\mathcal{F}\left(\psi_{j,k}\right)\right).$$

For the original construction of BlockJS (i.e. in the standard Gaussian noise model), we refer to [1].

Remark 4.1. *The sets A_j and $B_{j,K}$ are chosen such that $\bigcup_{K \in A_j} B_{j,K} = \{0, \dots, 2^j - 1\}$, for any $(K, K') \in A_j^2$ with $K \neq K'$, $B_{j,K} \cap B_{j,K'} = \varnothing$ and* $\mathrm{Card}(B_{j,K}) = L = \lfloor \log \rho_n \rfloor$.

Remark 4.2. *Notice that, thanks to (5), for any $j \in \{j_1, \dots, j_2\}$, we have*

$$\begin{cases} \mathcal{D}_{j_1} \subset [-4\pi 3^{-1} 2^{j_1}, 4\pi 3^{-1} 2^{j_1}], \\ \mathcal{C}_j \subset [-8\pi 3^{-1} 2^j, -2\pi 3^{-1} 2^j] \cup [2\pi 3^{-1} 2^j, 8\pi 3^{-1} 2^j]. \end{cases} \tag{11}$$

5. Results

5.1. Main results

Theorem 5.1 below determines the rates of convergence achieved by $\widehat{f}$ under the MISE over Besov balls.

Theorem 5.1. *Consider the model (1) and recall that we want to estimate $f^{(d)}$ with $d \in \mathbb{N}$. Assume that (ϕ, ψ) is r-regular for some $r \geq d$ and (3) is satisfied. Let $\widehat{f}$ be the estimator defined by (10) with a large enough λ. Then there exists a constant $C > 0$ such that, for any $M > 0$, $p \geq 1$, $r \geq 1$, $s > 1/p$ and n large enough, we have*

$$\sup_{f^{(d)} \in B^s_{p,r}(M)} \mathbb{E}\left(\int_0^1 \left(\widehat{f}(x) - f^{(d)}(x)\right)^2 dx\right) \leq C\varphi_n,$$

where

$$\varphi_n = \begin{cases} \rho_n^{-2s/(2s+2\delta+2d+1)}, & \text{if } p \geq 2, \\ (\log \rho_n/\rho_n)^{2s/(2s+2\delta+2d+1)}, & \text{if } p \in [1,2),\, s > (1/p - 1/2)(2\delta + 2d + 1). \end{cases}$$

Theorem 5.1 can be proved by using a more general theorem: [11, Theorem 3.1]. To apply this result, two conditions on the wavelet coeffcients estimators are required: a moment condition and a concentration condition. They are presented in Propositions 5.2 and 5.3 below.

It is natural to address the following question: is it φ_n the optimal rate of convergence ? Theorem 5.2 below gives the answer.

Theorem 5.2. *Consider the model (1) and recall that we want to estimate $f^{(d)}$ with $d \in \mathbb{N}$. Assume that (3) is satisfied. Then there exists a constant $c > 0$ such that, for any $M > 0$, $p \geq 1$, $r \geq 1$, $s > 1/p$ and n large enough, we have*

$$\inf_{\widetilde{f}} \sup_{f^{(d)} \in B^s_{p,r}(M)} \mathbb{E}\left(\int_0^1 \left(\widetilde{f}(x) - f^{(d)}(x)\right)^2 dx\right) \geq c\varphi_n^*,$$

where

$$\varphi_n^* = (\rho_n^*)^{-2s/(2s+2\delta+2d+1)}, \qquad \rho_n^* = \sum_{v=1}^{n} \sigma_v^{-2\delta}.$$

Theorem 5.2 shows that the rate of convergence φ_n achieved by $\widehat{f}$ is near optimal. Near is only due to the case $\pi \in [1,2)$ and $s > (1/p - 1/2)(2\delta + 2d + 1)$ where there is an extra logarithmic term.

Theorems 5.1 and 5.2 prove that $\widehat{f}$ is near optimal in the minimax sense.

5.2. Auxiliary results

In the three following result, we consider the framework of Theorem 5.1 and, for any integer $j \geq \tau$ and $k \in \{1, \ldots, 2^j - 1\}$, we set $\alpha_{j,k} = \int_0^1 f^{(d)}(t)\overline{\phi_{j,k}}(t)dt$ and $\beta_{j,k} = \int_0^1 f^{(d)}(t)\overline{\psi_{j,k}}(t)dt$, the wavelet coefficients (8) of $f^{(d)}$.

Proposition 5.1 (Gaussian distribution on the wavelet coefficient estimators)**.** *For any integer $j \geq \tau$ and $k \in \{0, \ldots, 2^j - 1\}$, we have*

$$\widehat{\alpha}_{j,k} \sim \mathcal{N}\left(\alpha_{j,k}, \epsilon^2 \frac{1}{\rho_n^2} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^{2\delta}} \sum_{\ell \in \mathcal{D}_j} (2\pi\ell)^{2d} \frac{|\mathcal{F}(\phi_{j,k})(\ell)|^2}{|\mathcal{F}(g_v)(\ell)|^2}\right)$$

and

$$\widehat{\beta}_{j,k} \sim \mathcal{N}\left(\beta_{j,k}, \epsilon^2 \frac{1}{\rho_n^2} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^{2\delta}} \sum_{\ell \in \mathcal{C}_j} (2\pi\ell)^{2d} \frac{|\mathcal{F}(\psi_{j,k})(\ell)|^2}{|\mathcal{F}(g_v)(\ell)|^2}\right).$$

Proposition 5.2 (Moment condition)**.**

- *There exists a constant $C > 0$ such that, for any integer $j \geq \tau$ and $k \in \{0, \ldots, 2^j - 1\}$,*

$$\mathbb{E}\left(|\widehat{\alpha}_{j_1,k} - \alpha_{j_1,k}|^2\right) \leq C\epsilon^2 2^{2(\delta+d)j_1} \rho_n^{-1},$$

- *There exists a constant $C > 0$ such that, for any integer $j \geq \tau$ and $k \in \{0, \ldots, 2^j - 1\}$,*

$$\mathbb{E}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}|^4\right) \leq C\epsilon^4 2^{4(\delta+d)j} \rho_n^{-2}.$$

Proposition 5.3 (Concentration condition)**.** *There exists a constant $\lambda > 0$ such that, for any $j \in \{j_1, \ldots, j_2\}$, any $K \in \mathcal{A}_j$ and n large enough,*

$$\mathbb{P}\left(\left(\sum_{k \in B_{j,K}} |\widehat{\beta}_{j,k} - \beta_{j,k}|^2\right)^{1/2} \geq \lambda 2^{(\delta+d)j} (\log \rho_n / \rho_n)^{1/2}\right) \leq \rho_n^{-2}.$$

6. Simulations results

In the following simulation study we consider the problem of estimating one of the derivatives of a function f from the heteroscedastic multichannel deconvolution model (1). Three test functions ("Wave", "Parabolas" and "TimeShiftedSine", initially introduced in [19]) representing different degrees of smoothness were used (see Fig 1). The "Wave" function was used to illustrate the performances of our estimator for smooth function. Note that the "Parabolas" function has big jumps in its second derivative.

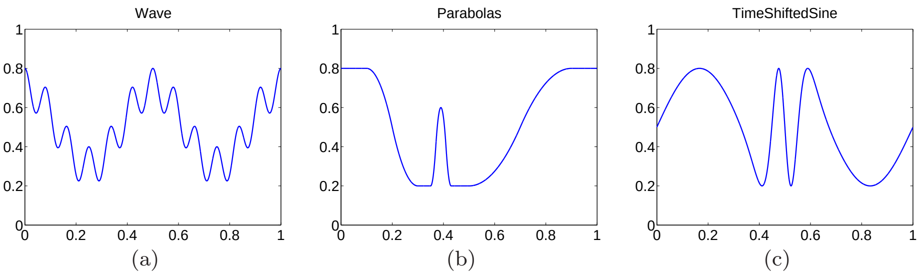


Fig 1. *Original Signals (a): Wave. (b): Parabolas. (c): TimeShiftedSine.*

We have compared the numerical performance of BlockJS to state-of-the-art classical thresholding methods of the litterature. In particular we consider the block estimator of [22] and two term-by-term thresholding methods. The first one is the classical hard thresholding and the other one corresponds to the non-negative garrote (introduced in wavelet estimation by [17]). In the sequel, we name the estimator of [22] by 'BlockH', the one of [17] by 'TermJS' and our estimator by 'BlockJS'. The performance of the estimators are measured in terms of peak signal-to-noise ratio (PSNR $= 20 \log_{10} \frac{n\|f\|_\infty}{\|\hat{f}-f\|_2}$) in decibels (dB). The known function g_v corresponds to a Laplace distribution and was used throughout all experiments.

6.1. Monochannel simulation

As an example of homoscedastic monochannel reconstruction (i.e. $n = 1$), we show in Fig 2 estimates obtained using the BlockJS method from $T = 4096$ equispaced values generated according to (1) with blurred signal-to-noise ratio equal to 25 (BSNR $= 10 \log_{10} \|f \star g_v\|/\epsilon^2$). For $d = 0$, the results are very effective for each test function. Note that the higher the index of the derivative increases, the harder it is to reconstruct fully the corresponding derivative. This is a consequence of computing the derivative of the noisy blurred signal in the frequency domain and the inverse Fourier transform that increases dramatically the high frequencies.

We then have compared the performance of our adaptive wavelet estimator with BlockH. The blurred signals were corrupted by a zero-mean white Gaussian noise such that the blurred signal-to-noise ratio ranged from 10 to 40 dB. The results are depicted in Fig 3 for $d = 0$, $d = 1$ and $d = 2$ respectively. One can see that our BlockJS thresholding estimator produce quite accurate estimates of f, f' and f'' for each test signals. These results clearly show that our approach compares favorably to BlockH and that BlockJS has good adaptive properties over a wide range of BSNR in the monochannel setting.

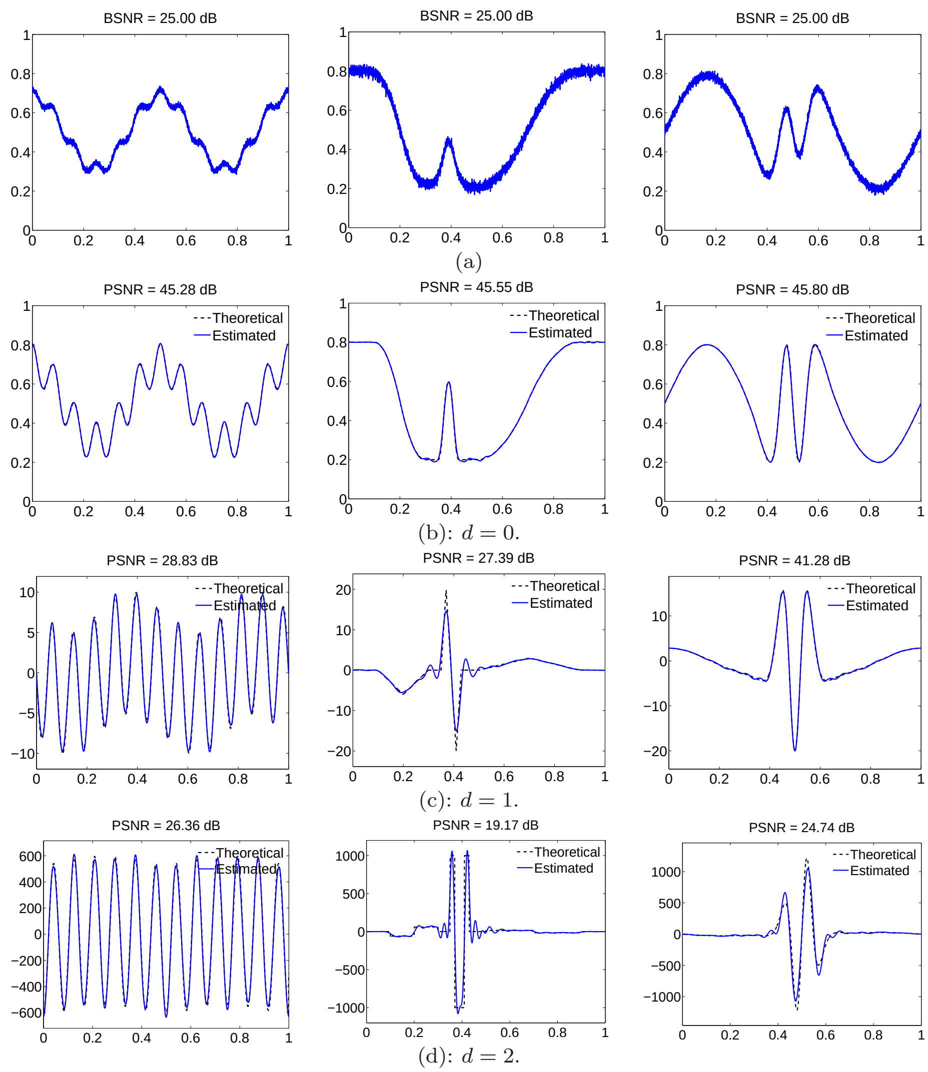
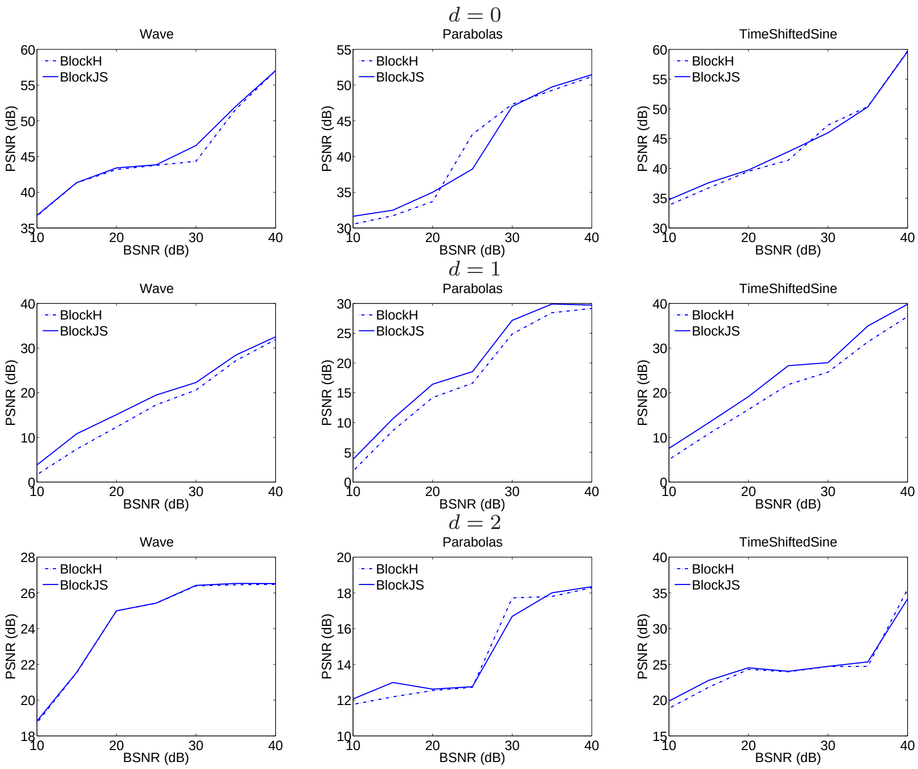


FIG 2. *Original (dashed) and estimated signals (solid) using the BlockJS thresholding estimator based on the (a) noisy blurred observations for (b):* $d = 0$*. (c):* $d = 1$ *(d):* $d = 2$*. From left to right Wave, Parabolas and TimeShiftedSine.*



(a) (b) (c)

FIG 3. *PSNR values as a function of the initial BSNR from 10 replications for (a): Wave. (b): Parabolas. (c): TimeShiftedSine from top to bottom* $d = 0, 1, 2$.

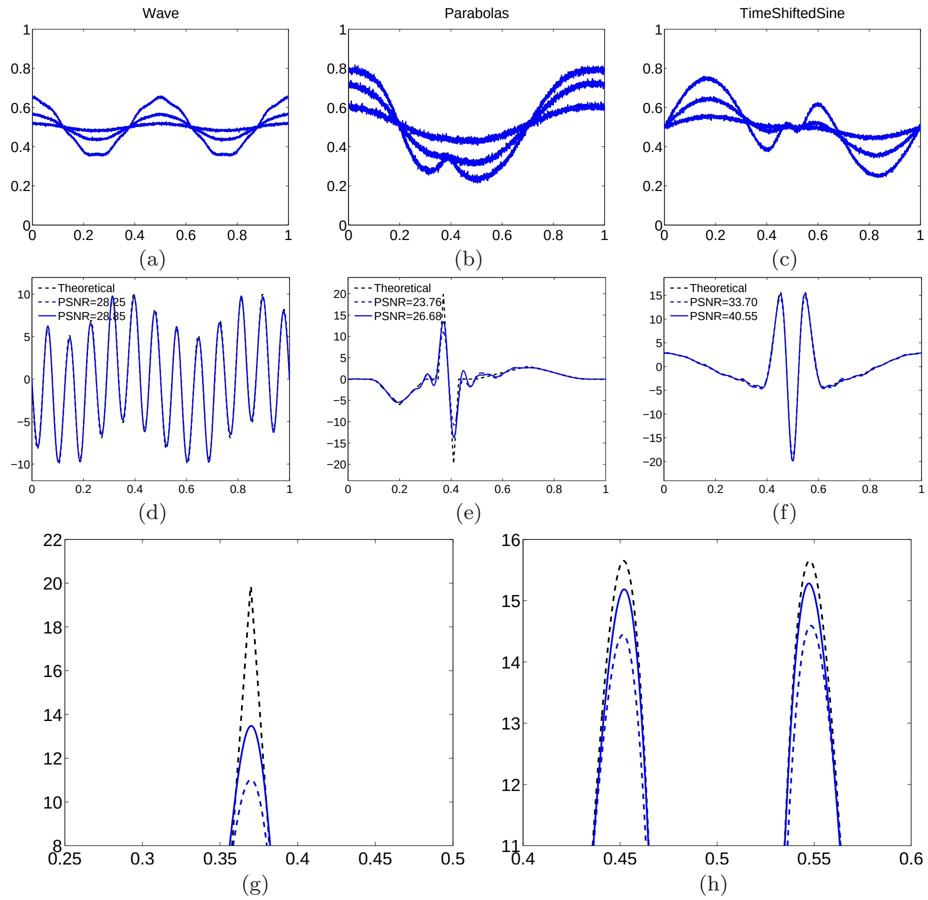


Fig 4. *Original Signals (dashed black) and estimate for $\sigma_v = v$ (dashed blue) and σ_v randomly generated (solid blue) from $n = 10$-channels. (a)-(c): noisy blurred observations (sample of 3 curves out of 10). (g): zoom on Parabolas. (h): zoom on TimeShiftedSine.*

6.2. Multichannel simulation

We would like to stress the fact that some choices of $\sigma_1, \ldots, \sigma_n$ can severely deteriorate the performance of the estimators. To illustrate this, we showed an example of multichannel estimates (in Fig 4) obtained using the BlockJS method from $T = 4096$ equispaced values with BSNR= 25, $\sigma_v = v$ (dashed blue) and σ_v randomly generated in $(0, +\infty)$ (solid blue). We can see a significant PSNR improvement up to 6.85 dB for the first derivative of TimeShiftedSine. Note that this improvement is marginal (about 0.60 dB) for the most regular test signal (i.e. Wave).

We concluded this simulation study by a comparison to the different methods used. For each test function, $T = 4096$ equally spaced samples on $[0, 1]$ were generated according to (1).

$\mathrm{BSNR}_{in} = 40$

	$d = 0$				$d = 1$				$d = 2$			
n	10	20	50	100	10	20	50	100	10	20	50	100
Wave												
BlockJS	72.16	74.92	76.94	77.91	43.02	35.00	37.52	35.66	32.33	29.14	30.41	23.06
BlockH	67.90	70.63	73.47	72.14	31.90	27.15	30.16	30.29	28.95	26.28	22.40	16.38
TermJS	70.27	73.03	74.28	73.53	35.24	30.42	33.39	32.23	31.35	27.84	23.75	18.77
TermH	67.49	69.33	67.08	66.84	26.00	24.55	28.07	27.48	28.59	22.56	16.83	15.23
TimeShiftedSine												
BlockJS	71.55	70.83	74.41	74.63	41.74	40.10	38.22	38.20	28.02	32.60	30.55	28.00
BlockH	65.08	68.07	70.89	67.82	33.59	31.45	31.29	33.64	24.77	29.19	20.15	17.57
TermJS	68.12	69.43	72.04	70.65	35.80	34.79	34.99	36.83	26.73	28.72	23.75	22.02
TermH	62.30	66.04	65.67	64.50	26.62	28.01	30.04	33.62	24.00	19.76	17.69	16.01
Parabolas												
BlockJS	64.10	64.13	68.62	69.31	39.61	39.10	34.34	35.76	22.75	22.78	22.46	22.15
BlockH	62.07	63.47	67.84	67.14	35.34	33.21	28.04	31.68	22.23	22.88	20.46	18.56
TermJS	63.45	63.90	68.33	68.04	37.62	34.74	31.78	34.29	22.65	22.78	21.26	19.32
TermH	61.70	63.07	63.52	63.00	29.92	28.30	27.59	31.17	22.33	19.74	15.21	14.78

$\mathrm{BSNR}_{in} = 25$

	$d = 0$				$d = 1$				$d = 2$			
n	10	20	50	100	10	20	50	100	10	20	50	100
Wave												
BlockJS	56.97	58.19	62.01	63.11	28.62	23.69	21.46	21.91	20.87	18.18	17.97	15.67
BlockH	52.43	54.03	57.91	56.87	21.07	17.41	16.74	17.29	17.03	15.96	15.45	13.39
TermJS	55.14	56.10	59.70	58.55	23.13	20.29	19.01	18.73	19.30	21.76	14.18	13.24
TermH	52.39	52.25	52.61	52.36	18.14	16.94	15.95	19.38	16.21	18.99	11.94	11.77
TimeShiftedSine												
BlockJS	58.36	56.59	60.13	61.29	23.61	25.00	23.64	22.60	21.11	19.78	19.07	13.70
BlockH	51.73	53.21	56.10	54.81	18.83	16.98	17.41	18.56	16.28	15.88	15.81	13.02
TermJS	54.51	55.05	56.88	55.87	18.24	19.80	22.01	22.30	19.28	17.86	18.19	13.80
TermH	50.43	51.60	50.48	49.15	17.36	14.35	18.56	23.12	15.89	19.41	16.84	13.24
Parabolas												
BlockJS	51.40	54.44	56.82	57.63	26.37	29.35	20.81	21.88	15.53	19.63	15.79	15.29
BlockH	47.72	50.73	54.27	52.82	18.56	20.67	14.29	17.52	13.26	17.47	14.65	12.72
TermJS	49.71	52.77	54.72	54.45	21.60	21.99	17.89	20.11	14.21	21.11	12.72	13.87
TermH	47.08	50.62	48.80	48.77	16.18	14.97	13.70	20.24	22.10	20.87	12.54	10.72

$\mathrm{BSNR}_{in} = 10$

	$d = 0$				$d = 1$				$d = 2$			
n	10	20	50	100	10	20	50	100	10	20	50	100
Wave												
BlockJS	42.04	43.29	47.04	48.16	18.10	16.70	15.63	15.63	15.78	16.64	14.63	12.65
BlockH	37.48	39.16	42.94	38.57	14.91	13.46	13.03	15.22	14.23	15.27	12.20	11.79
TermJS	40.23	41.22	44.74	43.63	15.53	14.42	14.59	15.43	15.59	15.78	12.47	11.93
TermH	37.46	37.35	37.70	37.59	13.37	12.30	13.29	15.05	16.78	12.88	11.41	10.85
TimeShiftedSine												
BlockJS	43.44	41.75	45.20	46.39	18.61	17.83	16.62	16.61	15.64	17.66	14.12	12.02
BlockH	36.97	38.40	41.17	39.93	15.17	12.76	12.92	14.16	13.61	15.86	12.04	11.07
TermJS	39.60	40.35	42.00	41.00	15.53	15.03	15.22	15.44	15.80	14.62	12.54	11.21
TermH	35.68	37.17	35.78	34.42	12.61	12.35	13.85	14.58	17.07	12.80	11.35	10.76
Parabolas												
BlockJS	37.00	40.27	42.25	43.13	15.99	15.49	13.73	14.45	12.02	13.83	12.30	11.25
BlockH	33.16	36.20	39.60	38.19	13.49	12.31	11.29	12.91	12.30	14.69	12.09	11.04
TermJS	35.00	38.22	39.97	39.76	13.85	13.14	12.48	14.03	13.58	14.81	11.95	11.16
TermH	32.46	35.95	34.18	34.26	11.57	11.56	12.93	14.01	14.67	13.54	10.77	11.04

Table 1

Comparison of average PSNR in decibels (dB) over ten realizations for $d = 0$ (left), $d = 1$ (middle) and $d = 2$ (right). From top to bottom $BSNR_{in} = 40, 25, 10$.

TABLE 1 shows that BlockJS uniformly outperforms the others methods in almost all cases in terms of peak signal-to-noise ratio (PSNR). Not surprisingly, the derivative estimation with BSNR_{in} equal to 10 seems to be pretty hard, especially for "Parabolas" (which has big jumps in its second derivative, see FIG. 3: (d)). Indeed, computing the derivative in the frequency domain degrades the PSNR of the estimations, since the high frequency components are dominated by the noise.

This numerical study confirms that under the heteroscedastic multichannel deconvolution model BlockJS thresholding wavelet estimator is very efficient.

Conclusion and perspectives

In this work, an adaptive wavelet block thresholding estimator was constructed to estimate one of the derivative of a function f from the heteroscedastic multichannel deconvolution model. Under ordinary smooth assumption on $g_1, \ldots, g_n$, it was proved that it is nearly optimal in the minimax sense. The practical comparisons to state-of-the art methods have demonstrated the usefulness and the efficiency of adaptive block thresholding methods in estimating a function f and its first derivatives in the functional deconvolution setting.

It would be interesting to consider the case where g_v are unknown, which is the case in many practical situations. Another interesting perspective would be to extend our results to a multidimensional setting. These aspects need further investigations that we leave for a future work.

7. Proofs

In the following proofs, c and C denote positive constants which can take different values for each mathematical term.

7.1. Proofs of the auxiliary results

Proof of Proposition 5.1. Let us prove the second point, the first one can be proved in a similar way. For any $\ell \in \mathbb{Z}$ and any $v \in \{1, \ldots, n\}$, $\mathcal{F}(f \star g_v)(\ell) = \mathcal{F}(f)(\ell)\mathcal{F}(g_v)(\ell)$. Therefore, if we set

$$y_{\ell,v} = \int_0^1 e^{-2\pi i \ell t} dY_v(t), \qquad e_{\ell,v} = \int_0^1 e^{-2\pi i \ell t} dW_v(t),$$

It follows from (1) that

$$y_{\ell,v} = \mathcal{F}(f)(\ell)\mathcal{F}(g_v)(\ell) + \epsilon e_{\ell,v}. \tag{12}$$

Note that, since f is 1-periodic, for any $u \in \{0, \ldots, d\}$, $f^{(u)}$ is 1-periodic and $f^{(u)}(0) = f^{(u)}(1)$. By d integrations by parts, for any $\ell \in \mathbb{Z}$, we have

$\mathcal{F}(f^{(d)})(\ell) = (2\pi i\ell)^d\mathcal{F}(f)(\ell)$. The Plancherel-Parseval theorem gives

$$\begin{aligned}\beta_{j,k} &= \int_0^1 f^{(d)}(t)\overline{\psi_{j,k}}(t)dt = \sum_{\ell\in\mathcal{C}_j}\mathcal{F}(f^{(d)})(\ell)\overline{\mathcal{F}(\psi_{j,k})}(\ell)\\ &= \sum_{\ell\in\mathcal{C}_j}(2\pi i\ell)^d\mathcal{F}(f)(\ell)\overline{\mathcal{F}(\psi_{j,k})}(\ell).\end{aligned}$$

Using (12), we have

$$\begin{aligned}\widehat{\beta}_{j,k} &= \frac{1}{\rho_n}\sum_{v=1}^n\frac{1}{(1+\sigma_v^2)^\delta}\sum_{\ell\in\mathcal{C}_j}(2\pi i\ell)^d\frac{\overline{\mathcal{F}(\psi_{j,k})(\ell)}}{\mathcal{F}(g_v)(\ell)}\mathcal{F}(f)(\ell)\mathcal{F}(g_v)(\ell)\\ &+ \epsilon\frac{1}{\rho_n}\sum_{v=1}^n\frac{1}{(1+\sigma_v^2)^\delta}\sum_{\ell\in\mathcal{C}_j}(2\pi i\ell)^d\frac{\overline{\mathcal{F}(\psi_{j,k})(\ell)}}{\mathcal{F}(g_v)(\ell)}e_{\ell,v}\\ &= \sum_{\ell\in\mathcal{C}_j}(2\pi i\ell)^d\mathcal{F}(f)(\ell)\overline{\mathcal{F}(\psi_{j,k})}(\ell)\\ &+ \epsilon\frac{1}{\rho_n}\sum_{v=1}^n\frac{1}{(1+\sigma_v^2)^\delta}\sum_{\ell\in\mathcal{C}_j}(2\pi i\ell)^d\frac{\overline{\mathcal{F}(\psi_{j,k})(\ell)}}{\mathcal{F}(g_v)(\ell)}e_{\ell,v}\\ &= \beta_{j,k}+\epsilon\frac{1}{\rho_n}\sum_{v=1}^n\frac{1}{(1+\sigma_v^2)^\delta}\sum_{\ell\in\mathcal{C}_j}(2\pi i\ell)^d\frac{\overline{\mathcal{F}(\psi_{j,k})(\ell)}}{\mathcal{F}(g_v)(\ell)}e_{\ell,v}.\end{aligned}$$

Since $(e^{-2\pi i\ell\cdot})_{\ell\in\mathbb{Z}}$ is an orthnormal basis of $\mathbb{L}^2_{per}([0,1])$ and $W_1(t),\dots,W_n(t)$ are i.i.d. standard Brownian motions, $\left(\int_0^1 e^{-2\pi i\ell t}dW_v(t)\right)_{(\ell,v)\in\mathbb{Z}\times\{1,\dots,n\}}$ is a sequence of i.i.d. random variables with the common distribution $\mathcal{N}(0,1)$. Therefore

$$\widehat{\beta}_{j,k}\sim\mathcal{N}\left(\beta_{j,k},\epsilon^2\frac{1}{\rho_n^2}\sum_{v=1}^n\frac{1}{(1+\sigma_v^2)^{2\delta}}\sum_{\ell\in\mathcal{C}_j}(2\pi\ell)^{2d}\frac{|\mathcal{F}(\psi_{j,k})(\ell)|^2}{|\mathcal{F}(g_v)(\ell)|^2}\right).$$

Proposition 5.1 is proved.

□

Proof of Proposition 5.2. Let us prove the second point, the first one can be proved in a similar way. Let us recall that, by Proposition 5.1, for any $j\in\{j_1,\dots,j_2\}$ and any $k\in\{0,\dots,2^j-1\}$, we have

$$\widehat{\beta}_{j,k}-\beta_{j,k}\sim\mathcal{N}\left(0,\rho_n^{-2}\sigma_{j,k}^2\right), \tag{13}$$

where

$$\sigma_{j,k}^2=\epsilon^2\sum_{v=1}^n\frac{1}{(1+\sigma_v^2)^{2\delta}}\sum_{\ell\in\mathbb{Z}}(2\pi\ell)^{2d}\frac{|\mathcal{F}(\psi_{j,k})(\ell)|^2}{|\mathcal{F}(g_v)(\ell)|^2}. \tag{14}$$

Due to (3) and (11), for any $v \in \{1, \ldots, n\}$, we have

$$\begin{aligned} \sup_{\ell \in \mathcal{C}_j} \left(\frac{(2\pi\ell)^{2d}}{|\mathcal{F}(g_v)(\ell)|^2} \right) &\leq C \sup_{\ell \in \mathcal{C}_j} \left((2\pi\ell)^{2d} \left(1 + \sigma_v^2 \ell^2\right)^{\delta} \right) \\ &\leq C(1+\sigma_v^2) \sup_{\ell \in \mathcal{C}_j} \left((2\pi\ell)^{2d} \left(1 + \ell^2\right)^{\delta} \right) \\ &\leq C(1+\sigma_v^2) 2^{2(\delta+d)j}. \end{aligned} \tag{15}$$

It follows from (15) and the Plancherel-Parseval theorem that

$$\begin{aligned} \sigma_{j,k}^2 &\leq \epsilon^2 \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^{2\delta}} \sup_{\ell \in \mathcal{C}_j} \left(\frac{(2\pi\ell)^{2d}}{|\mathcal{F}(g_v)(\ell)|^2} \right) \sum_{\ell \in \mathcal{C}_j} |\mathcal{F}(\psi_{j,k})(\ell)|^2 \\ &\leq C\epsilon^2 2^{2(\delta+d)j} \left(\sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^{\delta}} \right) \left(\sum_{\ell \in \mathcal{C}_j} |\mathcal{F}(\psi_{j,k})(\ell)|^2 \right) \\ &= C\epsilon^2 2^{2(\delta+d)j} \rho_n \int_{-\infty}^{\infty} |\mathcal{F}(\psi_{j,k})(y)|^2 dy \\ &= C\epsilon^2 2^{2(\delta+d)j} \rho_n \int_0^1 |\psi_{j,k}(x)|^2 dx = C\epsilon^2 \rho_n 2^{2(\delta+d)j}. \end{aligned} \tag{16}$$

Putting (13), (14) and (16) together, we obtain

$$\mathbb{E}\left(|\widehat{\beta}_{j,k} - \beta_{j,k}|^4 \right) \leq C(\epsilon^2 2^{2(\delta+d)j} \rho_n \rho_n^{-2})^2 = C\epsilon^4 2^{4(\delta+d)j} \rho_n^{-2}.$$

Proposition 5.2 is proved.

□

Proof of Proposition 5.3. We need the Cirelson inequality presented in Lemma 7.1 below.

Lemma 7.1 ([12]). *Let $(\vartheta_t)_{t \in D}$ be a centered Gaussian process. If*

$$\mathbb{E}\left(\sup_{t \in D} \vartheta_t \right) \leq N, \qquad \sup_{t \in D} \mathbb{V}(\vartheta_t) \leq V$$

then, for any $x > 0$, we have

$$\mathbb{P}\left(\sup_{t \in D} \vartheta_t \geq x + N \right) \leq \exp\left(-\frac{x^2}{2V} \right).$$

For the sake of simplicity, set

$$V_{j,k} = \widehat{\beta}_{j,k} - \beta_{j,k}.$$

Recall that, by Proposition 5.1, we have $V_{j,k} \sim \mathcal{N}\left(0, \rho_n^{-2} \sigma_{j,k}^2\right)$, where $\sigma_{j,k}^2$ is defined by (14). Consider the set Ω defined by $\Omega = \{a = (a_k) \in \mathbb{C};\ \sum_{k \in B_{j,K}} |a_k|^2 \leq$

1}. For any $a \in \Omega$, let $Z(a)$ be the centered Gaussian process defined by

$$\begin{aligned} Z(a) &= \sum_{k \in B_{j,K}} a_k V_{j,k} \\ &= \epsilon \frac{1}{\rho_n} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^\delta} \sum_{\ell \in \mathcal{C}_j} (2\pi i \ell)^d \frac{e_{\ell,v}}{\mathcal{F}(g_v)(\ell)} \sum_{k \in B_{j,K}} a_k \overline{\mathcal{F}\left(\psi_{j,k}\right)(\ell)}. \end{aligned}$$

By an argument of duality, we have

$$\sup_{a \in \Omega} Z(a) = \left(\sum_{k \in B_{j,K}} |V_{j,k}|^2 \right)^{1/2} = \left(\sum_{k \in B_{j,K}} |\widehat{\beta}_{j,k} - \beta_{j,k}|^2 \right)^{1/2}.$$

Now, let us determine the values of N and V which appeared in the Cirelson inequality.

Value of N. Using the Hölder inequality and (16), we obtain

$$\begin{aligned} \mathbb{E}\left(\sup_{a \in \Omega} Z(a)\right) &= \mathbb{E}\left(\left(\sum_{k \in B_{j,K}} |V_{j,k}|^2 \right)^{1/2} \right) \leq \left(\sum_{k \in B_{j,K}} \mathbb{E}\left(|V_{j,k}|^2\right) \right)^{1/2} \\ &\leq C \left(\rho_n^{-2} \sum_{k \in B_{j,K}} \sigma_{j,k}^2 \right)^{1/2} \leq C \left(\rho_n^{-2} \epsilon^2 \rho_n 2^{2(\delta+d)j} \operatorname{Card}(B_{j,K}) \right)^{1/2} \\ &= C \epsilon 2^{(\delta+d)j} \rho_n^{-1/2} \end{aligned}$$

Hence $N = C\epsilon 2^{(\delta+d)j} (\log \rho_n / \rho_n)^{1/2}$.

Value of V. Note that, for any $(\ell, \ell') \in \mathbb{Z}^2$ and any $(v, v') \in \{1, \dots, n\}^2$,

$\mathbb{E}\left(e_\ell \overline{e_{\ell'}}\right) = \begin{cases} 1 & \text{if } \ell = \ell' \text{ and } v = v', \\ 0 & \text{otherwise,} \end{cases}$ it comes

$$\sup_{a\in\Omega} \mathbb{V}(Z(a)) = \sup_{a\in\Omega} \mathbb{E}\left(\left|\sum_{k\in B_{j,K}} a_k V_{j,k}\right|^2\right)$$
$$= \sup_{a\in\Omega} \mathbb{E}\left(\sum_{k\in B_{j,K}} \sum_{k'\in B_{j,K}} a_k \overline{a_{k'}} V_{j,k} \overline{V_{j,k'}}\right)$$
$$= \epsilon^2 \rho_n^{-2} \sup_{a\in\Omega} \sum_{k\in B_{j,K}} \sum_{k'\in B_{j,K}} a_k \overline{a_{k'}} \sum_{\ell\in\mathcal{C}_j} \sum_{\ell'\in\mathcal{C}_j} \sum_{v'=1}^{n} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^\delta} \frac{1}{(1+\sigma_{v'}^2)^\delta} \times$$
$$\frac{(2\pi i\ell)^d}{\mathcal{F}(g_v)(\ell)} \overline{\mathcal{F}(\psi_{j,k})}(\ell) \frac{\overline{(2\pi i\ell')^d}}{\overline{\mathcal{F}(g_{v'})(\ell')}} \mathcal{F}(\psi_{j,k'})(\ell') \mathbb{E}\left(e_{\ell,v} \overline{e_{\ell',v'}}\right)$$
$$= \epsilon^2 \rho_n^{-2} \sup_{a\in\Omega} \sum_{k\in B_{j,K}} \sum_{k'\in B_{j,K}} a_k \overline{a_{k'}} \sum_{\ell\in\mathcal{C}_j} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^{2\delta}} \frac{(2\pi\ell)^{2d}}{|\mathcal{F}(g_v)(\ell)|^2} \overline{\mathcal{F}\left(\psi_{j,k}\right)}(\ell) \mathcal{F}(\psi_{j,k'})(\ell)$$
$$= \epsilon^2 \rho_n^{-2} \sup_{a\in\Omega} \sum_{\ell\in\mathcal{C}_j} \sum_{v=1}^{n} \frac{1}{(1+\sigma_v^2)^{2\delta}} \frac{(2\pi\ell)^{2d}}{|\mathcal{F}(g_v)(\ell)|^2} \left|\sum_{k\in B_{j,K}} \overline{a_k} \mathcal{F}\left(\psi_{j,k}\right)(\ell)\right|^2. \tag{17}$$

For any $a \in \Omega$, the Plancherel-Parseval theorem gives

$$\sum_{\ell\in\mathcal{C}_j} \left|\sum_{k\in B_{j,K}} \overline{a_k} \mathcal{F}(\psi_{j,k})(\ell)\right|^2 = \sum_{\ell\in\mathcal{C}_j} \left|\mathcal{F}\left(\sum_{k\in B_{j,K}} \overline{a_k} \psi_{j,k}\right)(\ell)\right|^2$$
$$= \int_{-\infty}^{\infty} \left|\mathcal{F}\left(\sum_{k\in B_{j,K}} \overline{a_k} \psi_{j,k}\right)(y)\right|^2 dy = \int_0^1 \left|\sum_{k\in B_{j,K}} \overline{a_k} \psi_{j,k}(x)\right|^2 dx$$
$$= \sum_{k\in B_{j,K}} |a_k|^2 \leq 1. \tag{18}$$

Putting (17), (15) and (18) together, we have

$$\sup_{a\in\Omega} \mathbb{V}(Z(a)) \leq C\epsilon^2 \rho_n^{-1} 2^{2(\delta+d)j} \sup_{a\in\Omega} \sum_{\ell\in\mathcal{C}_j} \left|\sum_{k\in B_{j,K}} \overline{a_k} \mathcal{F}\left(\psi_{j,k}\right)(\ell)\right|^2$$
$$\leq C\epsilon^2 \rho_n^{-1} 2^{2(\delta+d)j}.$$

Hence $V = C\epsilon^2 \rho_n^{-1} 2^{2(\delta+d)j}$.

Taking λ large enough and $x = 2^{-1}\lambda\epsilon 2^{(\delta+d)j}(\log\rho_n/\rho_n)^{1/2}$, the Cirelson inequality described in Lemma 7.1 yields

$$\begin{aligned}
&\mathbb{P}\left(\left(\sum_{k\in B_{j,K}}|V_{j,k}|^2\right)^{1/2} \geq \lambda\epsilon 2^{(\delta+d)j}(\log\rho_n/\rho_n)^{1/2}\right)\\
\leq\ & \mathbb{P}\left(\left(\sum_{k\in B_{j,K}}|V_{j,k}|^2\right)^{1/2} \geq 2^{-1}\lambda\epsilon 2^{(\delta+d)j}(\log\rho_n/\rho_n)^{1/2}+N\right)\\
=\ & \mathbb{P}\left(\sup_{a\in\Omega} Z(a)\geq x+N\right)\leq \exp\left(-x^2/(2V)\right)\leq \exp\left(-C\lambda^2\log\rho_n\right)\\
\leq\ & \rho_n^{-2}.
\end{aligned}$$

Proposition 5.3 is proved.

□

Putting Propositions 5.2 and 5.3 in [11, Theorem 3.1], we end the proof of Theorem 5.1.

□

Proof of Theorem 5.2. Let us now present a consequence of the Fano lemma.

Lemma 7.2. *Let $m \in \mathbb{N}^*$ and A be a sigma algebra on the space Ω. For any $i \in \{0,\dots,m\}$, let $A_i \in A$ such that, for any $(i,j)\in\{0,\dots,m\}^2$ with $i \neq j$,*

$$A_i \cap A_j = \varnothing.$$

Let $(\mathbb{P}_i)_{i\in\{0,\dots,m\}}$ be $m+1$ probability measures on (Ω, A). Then

$$\sup_{i\in\{0,\dots,m\}} \mathbb{P}_i\left(A_i^c\right) \geq \min\left(2^{-1}, \exp(-3e^{-1})\sqrt{m}\exp(-\chi_m)\right),$$

where

$$\chi_m = \inf_{v\in\{0,\dots,m\}} \frac{1}{m}\sum_{\substack{k\in\{0,\dots,m\}\\ k\neq v}} K(\mathbb{P}_k,\mathbb{P}_v),$$

and K is the Kullbak-Leibler divergence defined by

$$K(\mathbb{P},\mathbb{Q}) = \begin{cases}\int \ln\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)d\mathbb{P} & \text{if } \mathbb{P} << \mathbb{Q},\\ \infty & \text{otherwise.}\end{cases}$$

The proof of Lemma 7.2 can be found in [15, Lemma 3.3]. For further details and applications of the Fano lemma, see [25].

Consider the Besov balls $B^s_{\pi,r}(M)$ (see (9)). Let j_0 be an integer suitably chosen below. For any $\varepsilon = (\varepsilon_k)_{k\in\{0,\dots,2^{j_0}-1\}} \in \{0,1\}^{2^{j_0}}$ and $d \in \mathbb{N}^*$, set

$$\begin{aligned} h_\varepsilon(x) &= M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \varepsilon_k \frac{1}{(d-1)!} \int_0^x (x-y)^{d-1} \psi_{j_0,k}(y) dy, \\ & x \in [0,1], \end{aligned}$$

(and, if $d = 0$, set $h_\varepsilon(x) = M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \varepsilon_k \psi_{j_0,k}(x), x \in [0,1]$). Notice that, due to (7), h_ε exists and, since $\psi_{j_0,k}$ is 1-periodic, h_ε is also 1-periodic. Using the Cauchy formula for repeated integration, we have

$$h^{(d)}_\varepsilon(x) = M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \varepsilon_k \psi_{j_0,k}(x), \qquad x \in [0,1].$$

So, for any $j \geq \tau$ and any $k \in \{0,\dots,2^j-1\}$, the (mother) wavelet coefficient of $h^{(d)}_\varepsilon$ is

$$\beta_{j,k} = \int_0^1 h^{(d)}_\varepsilon(x) \overline{\psi_{j,k}}(x) dx = \begin{cases} M_* \varepsilon_k 2^{-j_0(s+1/2)}, & \text{if } j = j_0, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore $h^{(d)}_\varepsilon \in B^s_{p,r}(M)$. The Varshamov-Gilbert theorem (see [25, Lemma 2.7]) asserts that there exist a set $E_{j_0} = \left\{\varepsilon^{(0)},\dots,\varepsilon^{(T_{j_0})}\right\}$ and two constants, $c \in]0,1[$ and $\alpha \in]0,1[$, such that, for any $u \in \{0,\dots,T_{j_0}\}$, $\varepsilon^{(u)} = (\varepsilon^{(u)}_k)_{k\in\{0,\dots,2^{j_0}-1\}} \in \{0,1\}^{2^{j_0}}$ and any $(u,v) \in \{0,\dots,T_{j_0}\}^2$ with $u < v$, the following hold:

$$\sum_{k=0}^{2^{j_0}-1} |\varepsilon^{(u)}_k - \varepsilon^{(v)}_k| \geq c 2^{j_0}, \qquad T_{j_0} \geq e^{\alpha 2^{j_0}}.$$

Considering such a E_{j_0}, for any $(u,v) \in \{0,\dots,T_{j_0}\}^2$ with $u \neq v$, we have

$$\begin{aligned} & \left(\int_0^1 \left(h^{(d)}_{\varepsilon^{(u)}}(x) - h^{(d)}_{\varepsilon^{(v)}}(x) \right)^2 dx \right)^{1/2} \\ &= c 2^{-j_0(s+1/2)} \left(\sum_{k=0}^{2^{j_0}-1} \left| \varepsilon^{(u)}_k - \varepsilon^{(v)}_k \right|^2 \right)^{1/2} \\ &= c 2^{-j_0(s+1/2)} \left(\sum_{k=0}^{2^{j_0}-1} \left| \varepsilon^{(u)}_k - \varepsilon^{(v)}_k \right| \right)^{1/2} \\ &\geq 2\delta_{j_0}, \end{aligned}$$

where

$$\delta_{j_0} = c 2^{j_0/2} 2^{-j_0(s+1/2)} = c 2^{-j_0 s}.$$

Using the Markov inequality, for any estimator $\widetilde{f}$ of $f^{(d)}$, we have

$$\delta_{j_0}^{-2} \sup_{f^{(d)} \in B^s_{\pi,r}(M)} \mathbb{E}\left(\int_0^1 \left(\widetilde{f}(x) - f^{(d)}(x) \right)^2 dx \right) \geq \sup_{u \in \{0,\dots,T_{j_0}\}} \mathbb{P}_{h_{\varepsilon^{(u)}}}(A_u^c) = p,$$

where

$$A_u = \left\{ \left(\int_0^1 \left(\widetilde{f}(x) - h^{(d)}_{\varepsilon^{(u)}}(x) \right)^2 dx \right)^{1/2} < \delta_{j_0} \right\}$$

and $\mathbb{P}_f$ is the distribution of (1). Notice that, for any $(u,v) \in \{0,\dots,T_{j_0}\}^2$ with $u \neq v$, $A_u \cap A_v = \varnothing$. Lemma 7.2 applied to the probability measures $\left(\mathbb{P}_{h_{\varepsilon^{(u)}}}\right)_{u \in \{0,\dots,T_{j_0}\}}$ gives

$$p \geq \min\left(2^{-1}, \exp(-3e^{-1}) \sqrt{T_{j_0}} \exp(-\chi_{T_{j_0}}) \right), \tag{19}$$

where

$$\chi_{T_{j_0}} = \inf_{v \in \{0,\dots,T_{j_0}\}} \frac{1}{T_{j_0}} \sum_{\substack{u \in \{0,\dots,T_{j_0}\} \\ u \neq v}} K\left(\mathbb{P}_{h_{\varepsilon^{(u)}}}, \mathbb{P}_{h_{\varepsilon^{(v)}}} \right).$$

Let us now bound $\chi_{T_{j_0}}$. For any functions f_1 and f_2 in $\mathbb{L}^2_{per}([0,1])$, we have

$$\begin{aligned} K\left(\mathbb{P}_{f_1}, \mathbb{P}_{f_2}\right) &= \frac{1}{2\epsilon^2} \sum_{v=1}^n \int_0^1 \left((f_1 \star g_v)(x) - (f_2 \star g_v)(x) \right)^2 dx \\ &= \frac{1}{2\epsilon^2} \sum_{v=1}^n \int_0^1 \left(((f_1 - f_2) \star g_v)(x) \right)^2 dx. \end{aligned}$$

The Plancherel-Parseval theorem yields

$$\begin{aligned} K\left(\mathbb{P}_{f_1}, \mathbb{P}_{f_2}\right) &= \frac{1}{2\epsilon^2} \sum_{v=1}^n \int_{-\infty}^{\infty} |\mathcal{F}((f_1 - f_2) \star g_v)(x)|^2 dx \\ &= \frac{1}{2\epsilon^2} \sum_{v=1}^n \int_{-\infty}^{\infty} |\mathcal{F}(f_1 - f_2)(x)|^2 \, |\mathcal{F}(g_v)(x)|^2 dx. \end{aligned}$$

So, for any $(u,v) \in \{0,\dots,T_{j_0}\}^2$ with $u \neq v$, we have

$$K\left(\mathbb{P}_{h_{\varepsilon^{(u)}}}, \mathbb{P}_{h_{\varepsilon^{(v)}}} \right) = \frac{1}{2\epsilon^2} \sum_{v=1}^n \int_{-\infty}^{\infty} \left| \mathcal{F}\left(h_{\varepsilon^{(u)}} - h_{\varepsilon^{(v)}} \right)(x) \right|^2 |\mathcal{F}(g_v)(x)|^2 dx. \tag{20}$$

By definition, for any $(u,v) \in \{0,\dots,T_{j_0}\}^2$ with $u \neq v$ and $x \in \mathbb{R}$, we have

$$\begin{aligned} &\mathcal{F}\left(h_{\varepsilon^{(u)}} - h_{\varepsilon^{(v)}} \right)(x) \\ &= M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)} \right) \times \\ &\quad \frac{1}{(d-1)!} \mathcal{F}\left(\int_{-\infty}^{\cdot} (. - y)^{d-1} \psi_{j_0,k}(y) dy \right)(x). \end{aligned} \tag{21}$$

Let us set, for any $k \in \{0, \dots, 2^{j_0} - 1\}$,

$$\theta_k(x) = \int_{-\infty}^{x} (x-y)^{d-1} \psi_{j_0,k}(y) dy, \qquad x \in [0,1].$$

Then, for any $u \in \{0, \dots, d\}$, $\theta_k^{(u)}$ is 1-periodic and $\theta_k^{(u)}(0) = \theta_k^{(u)}(1)$. Therefore, by d integrations by parts, for any $x \in \mathbb{R}$, we have

$$\mathcal{F}\left(\theta_k^{(d)}\right)(x) = (2\pi i x)^d \mathcal{F}(\theta_k)(x).$$

Using again the Cauchy formula for repeated integration, we have $\theta_k^{(d)}(x) = \psi_{j_0,k}(x)$, $x \in [0,1]$. So, for any $x \in \mathcal{C}_{j_0}$ (excluding 0), (21) implies that

$$\begin{aligned} &\mathcal{F}(h_{\varepsilon^{(u)}} - h_{\varepsilon^{(v)}})(x) \\ &= \frac{M_*}{(d-1)!} 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)}\right) \frac{1}{(2\pi i x)^d} \mathcal{F}\left(\psi_{j_0,k}\right)(x). \end{aligned} \tag{22}$$

The equalities (20) and (22) imply that

$$\begin{aligned} &K\left(\mathbb{P}_{h_{\varepsilon^{(u)}}}, \mathbb{P}_{h_{\varepsilon^{(v)}}}\right) \\ &= C 2^{-2j_0(s+1/2)} \sum_{v=1}^{n} \int_{\mathcal{C}_{j_0}} \left| \sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)}\right) \mathcal{F}\left(\psi_{j_0,k}\right)(x) \right|^2 \frac{1}{(2\pi x)^{2d}} |\mathcal{F}(g_v)(x)|^2 dx. \end{aligned} \tag{23}$$

By (3) and (11), for any $v \in \{1, \dots, n\}$,

$$\begin{aligned} \sup_{x \in \mathcal{C}_{j_0}} \left(\frac{1}{(2\pi x)^{2d}} |\mathcal{F}(g_v)(x)|^2 \right) &\leq C \sup_{x \in \mathcal{C}_{j_0}} \left(\frac{1}{(2\pi x)^{2d}} \left(1 + \sigma_v^2 x^2\right)^{-\delta} \right) \\ &\leq C \sigma_v^{-2\delta} \sup_{x \in \mathcal{C}_{j_0}} \left(x^{-2(\delta+d)} \right) \leq C \sigma_v^{-2\delta} 2^{-2j_0(\delta+d)}. \end{aligned} \tag{24}$$

Moreover, the Plancherel-Parseval theorem implies that

$$\begin{aligned} &\int_{\mathcal{C}_{j_0}} \left| \sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)}\right) \mathcal{F}\left(\psi_{j_0,k}\right)(x) \right|^2 dx \\ &= \int_{-\infty}^{\infty} \left| \mathcal{F}\left(\sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)}\right) \psi_{j_0,k} \right)(x) \right|^2 dx \\ &= \int_0^1 \left| \sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)}\right) \psi_{j_0,k}(x) \right|^2 dx = \sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)}\right)^2 \leq C 2^{j_0}. \end{aligned} \tag{25}$$

It follows from (23), (24) and (25) that

$$K\left(\mathbb{P}_{h_{\varepsilon(u)}}, \mathbb{P}_{h_{\varepsilon(v)}}\right) \leq C2^{-2j_0(s+1/2)}2^{-2j_0(\delta+d)}2^{j_0}\sum_{v=1}^{n}\sigma_v^{-2\delta} = C\rho_n^* 2^{-2j_0(s+1/2+\delta+d)}2^{j_0}.$$

Hence

$$\begin{aligned}\chi_{T_{j_0}} &= \inf_{v\in\{0,\ldots,T_{j_0}\}}\frac{1}{T_{j_0}}\sum_{\substack{u\in\{0,\ldots,T_{j_0}\}\\ u\neq v}} K\left(\mathbb{P}_{h_{\varepsilon(u)}}, \mathbb{P}_{h_{\varepsilon(v)}}\right)\\ &\leq C\rho_n^* 2^{-2j_0(s+1/2+\delta+d)}2^{j_0}. \qquad (26)\end{aligned}$$

Putting (19) and (26) together and choosing j_0 such that

$$2^{-j_0(s+1/2+\delta+d)} = c_0(\rho_n^*)^{-1/2},$$

where c_0 denotes a well chosen constant, for any estimator $\widetilde{f}$ of $f^{(d)}$, we have

$$\begin{aligned}\delta_{j_0}^{-2}\sup_{f^{(d)}\in B^s_{\pi,r}(M)}\mathbb{E}\left(\int_0^1\left(\widetilde{f}(x) - f^{(d)}(x)\right)^2 dx\right) &\geq c\exp\left((\alpha/2)2^{j_0} - Cc_0^2 2^{j_0}\right)\\ &\geq c,\end{aligned}$$

where

$$\delta_{j_0} = c2^{-j_0 s} = c(\rho_n^*)^{-s/(2s+2\delta+2d+1)}.$$

This complete the proof of Theorem 5.2.

□

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

References

[1] T. Cai. Adaptive wavelet estimation: A block thresholding and oracle inequality approach. *The Annals of Statistics*, 27:898–924, 1999.

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

[3] L. Cavalier. Penalized blockwise stein's method, monotone oracles and sharp adaptive estimation. *Math. Methods of Stat.*, 10:247–282, 2001.

[4] L. Cavalier. Nonparametric statistical inverse problems. *Inverse Problems*, 24, 2008.

[5] L. Cavalier, Y. Golubev, and A. B. Tsybakov. Block thresholding and sharp adaptive estimation in severely ill-posed inverse problems. *Probability Theory and Related Fields*, 48(3):426–446, 2004.

[6] L. Cavalier and A. B. Tsybakov. Sharp adaptation for inverse problems with random noise. *Probability Theory and Related Fields*, 123:323–354, 2002.

[7] Y.P. Chaubey and H. Doosti. Wavelet based estimation of the derivatives of a density for m-dependent random variables. *Journal of the Iranian Statistical Society*, 4(2):97–105, 2005.

[8] Y.P. Chaubey, H. Doosti, and B.L.S. Prakasa Rao. Wavelet based estimation of the derivatives of a density with associated variables. *International Journal of Pure and Applied Mathematics*, 27(1):97–106, 2006.

[9] C. Chesneau. Wavelet estimation via block thresholding: A minimax study under the $\mathbb{L}^p$ risk. *Statistica Sinica*, 18(3):1007–1024, 2008.

[10] C. Chesneau. Wavelet estimation of the derivatives of an unknown function from a convolution model. *Current Developments in Theory and Applications of Wavelets*, 4(2):131–151, 2010.

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

[12] C. Cirelson, I. Ibragimov, and V. Sudakov. *Norm of Gaussian sample functions*. Springer Verlag, 1976.

[13] A. Cohen, I. Daubechies, B. Jawerth, and P. Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 24(1):54–81, 1993.

[14] D. De Canditiis. Simultaneous wavelet deconvolution in periodic setting. *Scandinavian Journal of Statistics*, 33:293–306, 2006.

[15] R. DeVore, G. Kerkyacharian, D. Picard, and V. Temlyakov. On mathematical methods of learning. *Foundations of Computational Mathematics*, special issue for D. Smale 6(1):3–58, 2006.

[16] J. Fan and J.Y. Koo. Wavelet deconvolution. *IEEE transactions on information theory*, 48:734–747, 2002.

[17] H.Y. Gao. Wavelet shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7:469–488, 1998.

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

[19] M.H. Neumann J.S. Marron, S. Adak. I.M. Johnstone and P. Patil. Exact risk analysis of wavelet regression. *J. Comput. Graph. Statist.*, 7:278–309, 2004.

[20] Y. Meyer. *Wavelets and Operators*. Cambridge University Press, Cambridge, 1992.

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

[22] M. Pensky and T. Sapatinas. On convergence rates equivalency and sampling strategies in a functional deconvolution model. *The Annals of Statistics*, 38(3):1793–1844, 2010.

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

[24] B.L.S. Prakasa Rao. Nonparametric estimation of the derivatives of a density by the method of wavelets. *Bull. Inform. Cyb.*, 28:91–100, 1996.
[25] A. B. Tsybakov. *Introduction à l'estimation nonparametrique*. Springer Verlag, Berlin, 2004.
[26] G. Walter. *Wavelets and other orthogonal systems in applications*. CRC Press, Boca Raton, 1994.
[27] A. Zayed. Characterization of analytic functions in term of their wavelet coefficients. *Complex Variables*, 29:265–276, 1996.