



HAL
open science

Block thresholding for wavelet-based estimation of function derivatives from a heteroscedastic multichannel convolution model

Fabien Navarro, Christophe Chesneau, Jalal M. Fadili, Taoufik Sassi

► **To cite this version:**

Fabien Navarro, Christophe Chesneau, Jalal M. Fadili, Taoufik Sassi. Block thresholding for wavelet-based estimation of function derivatives from a heteroscedastic multichannel convolution model. *Electronic Journal of Statistics*, 2013, 7, pp.428-453. 10.1214/13-EJS776 . hal-00661544v3

HAL Id: hal-00661544

<https://hal.science/hal-00661544v3>

Submitted on 13 Mar 2013

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.

Block thresholding for wavelet-based estimation of function derivatives 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 $\{Y_v(t)\}_v$, where for any $v \in \{1, \dots, n\}$ and $t \in [0, 1]$, $Y_v(t)$ is the convolution product of an unknown function f and a known blurring function g_v corrupted by Gaussian noise. Under an ordinary smoothness assumption on g_1, \dots, g_n , our goal is to estimate the d -th derivatives (in weak sense) of f from the observations. We propose an adaptive estimator based on wavelet block thresholding, namely the "BlockJS estimator". Taking the mean integrated squared error (MISE), our main theoretical result investigates the minimax rates over Besov smoothness spaces, and shows that our block estimator can achieve the optimal minimax rate, or is at least nearly-minimax in the least favorable situation. We also report a comprehensive suite of numerical simulations to support our theoretical findings. The practical performance of our block estimator compares very favorably to existing methods of the literature on a large set of test functions.

AMS 2000 subject classifications: Primary 62G07, 62G20; secondary 62F12.

Keywords and phrases: deconvolution, multichannel observations, derivative estimation, wavelets, block thresholding, minimax.

Contents

1 Introduction 2

1.1 Problem statement 2

1.2 Overview of previous work 3

1.3 Contributions and relation to prior work 4

1.4 Paper organization 4

2 Wavelets and Besov balls 5

2.1 Periodized Meyer Wavelets 5

2.2 Besov balls 6

3 The deconvolution BlockJS estimator 6

3.1 The ordinary smoothness assumption 6

3.2 BlockJS estimator 7

4 Minimaxity results of BlockJS over Besov balls 8

4.1 Minimax upper-bound for the MISE 8

4.2 Minimax lower-bound for the MISE 9

5 Simulations results 10

5.1 Monochannel simulation 11

5.2 Multichannel simulation 11

6 Conclusion and perspectives 16

7 Proofs 16

7.1 Preparatory results 16

7.2 Proof of Theorem 4.1 21

7.3 Proof of Theorem 4.2 21

References 26

1. Introduction

1.1. Problem statement

Suppose that we observe n stochastic processes $Y_1(t), \dots, Y_n(t), t \in [0, 1]$ where, for any $v \in \{1, \dots, n\}$,

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

$\epsilon > 0$ is the noise level, $(f \star g_v)(t) = \int_0^1 f(t - u)g_v(u)du$ denotes the convolution product on $[0, 1]$, $W_1(t), \dots, W_n(t)$ are n unobserved independent standard Brownian motions, for any $v \in \{1, \dots, n\}$, $g_v : [0, 1] \rightarrow \mathbb{R}$ is a known blurring function and $f : [0, 1] \rightarrow \mathbb{R}$ is the unknown function that we target. We assume that f and g_1, \dots, g_n belong to $\mathbb{L}_{\text{per}}^2([0, 1]) = \{h; h \text{ is 1-periodic on } [0, 1] \text{ and } \int_0^1 h^2(t)dt < \infty\}$. The focus of this paper is to estimate f and its derivatives $f^{(d)}$ (to be understood in weak or distributional sense) from $Y_1(t), \dots, Y_n(t)$,

$t \in [0, 1]$. This is in general a severely ill-posed inverse problem. Application fields cover biomedical imaging, astronomical imaging, remote-sensing, seismology, etc. This list is by no means exhaustive.

In the following, any function $h \in \mathbb{L}_{\text{per}}^2([0, 1])$ can be represented by its Fourier series

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

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

$$\mathcal{F}(h)(\ell) = \int_0^1 h(t) e^{-2i\pi\ell t} dt, \quad \ell \in \mathbb{Z},$$

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

1.2. Overview of previous work

There is an extensive statistical literature on wavelet-based deconvolution problems. For obvious space limitations, we only focus on some of them.

In the special case where $g_1 = \dots = g_n$, (1) reduces to the form

$$d\tilde{Y}(t) = (f \star g_1)(t) dt + \epsilon n^{-1/2} d\tilde{W}(t), \quad t \in [0, 1], \quad (2)$$

where $\tilde{Y}(t) = (1/n) \sum_{v=1}^n Y_v(t)$, and $\tilde{W}(t) = (1/n^{1/2}) \sum_{v=1}^n W_v(t)$ is standard Brownian motion. In such a case, (2) becomes the standard deconvolution which attracted attention of a number of researchers spanning a wide range of fields including signal processing and statistics. For instance, wavelet-based estimators of f have been constructed and their asymptotic performance investigated in a number of papers, see e.g. [5–7, 10, 16, 19]. When g_1, \dots, g_n are not necessarily equal, estimators of f and their minimax rates under the mean integrated squared error (MISE) over Besov balls were proposed in [14, 22–24]. These authors develop wavelet thresholding estimators (hard thresholding in [14, 24] and block thresholding in [22, 23]) under various assumptions on g_1, \dots, g_n (typically, ordinary smooth and super-smooth case, or boxcar blurring functions).

Estimating the derivative of a function on the basis of noisy and blurred observations is of paramount importance in many fields such as signal processing, control or mathematical physics. For instance detecting the singularities of f or characterizing its concavity or convexity properties is a longstanding problem in signal processing. The estimation of the derivatives from noisy solutions of ordinary or partial differential equations is typical in many areas of mathematical physics such as astronomy or fluid dynamics. In the case where $d = 0$, several examples of recovering initial/boundary conditions from observations of a noisy and blurred solution of a PDE (e.g., parabolic, hyperbolic or elliptic PDE) are given in [22]. For higher-order derivatives (typically $d = 1$ or $d = 2$), there are

also physical applications where such a model occurs. We mention for instance frequency self-deconvolution encountered in the field of chemometrics. In this context, the first and the second derivatives can be used to detect important information or features in the raw spectra; see e.g. [1]). The wavelet estimator developed in the paper could be an interesting alternative to commonly used methods in this area.

The derivatives estimation have already been investigated from several standard nonparametric models. If we only restrict the review to wavelet methods, we refer to [3, 11] for model (2) and to [8, 9, 26] for density estimation problems.

1.3. Contributions and relation to prior work

In this paper, considering an appropriate ordinary smoothness assumption on g_1, \dots, g_n , we develop an adaptive wavelet-based block estimator $\widehat{f^{(d)}}$ of $f^{(d)}$ from (1), $d \in \mathbb{N}$. It is constructed using a periodized Meyer wavelet basis and a particular block thresholding rule which goes by the the name of BlockJS; see [2] for the original construction of BlockJS in the standard Gaussian noise model, and [3, 4, 12, 27] for further developments on BlockJS. Adopting the minimax approach under the MISE over Besov balls, we investigate the upper bounds of our estimator. This is featured in Theorem 4.1. We prove that the rates of our estimator are nearly optimal by establishing a lower bound as stated in Theorem 4.2.

Our work is related to some prior art in the literature. To the best of our knowledge, the closest ones are those of [22, 23]. For the case where $d = 0$ and the blurring function is ordinary-smooth or super-smooth, [22, 23, Theorems 1 and 2] provide the upper and lower bounds of the MISE over Besov balls for a block hard thresholding estimator from the functional deconvolution model¹. These bounds match ours but only for $d = 0$. In this respect, our work goes one step further as it tackles the estimation (with a different wavelet estimator) of f and its derivatives. As far as the methods of proof are concerned, we use similar tools (concentration and moment inequalities as well as the general result in [12]) as theirs for the upper bound, but the proof of the lower bounds are different. However unlike [22], we only cover the ordinary smooth convolution, while their results apply also to the super-smooth case. On the practical side, for $d = 0$, we will show in Section 5 that BlockJS behaves better than block hard thresholding [23, (2.9)] over several test functions that contain different degrees of irregularity.

1.4. Paper organization

The paper is organized as follows. Section 2 gives a brief account of periodized Meyer wavelets and Besov balls. Section 3 states ordinary smoothness assump-

¹This is a more general model which reduces to the multichannel deconvolution model when observed at a finite number of distinct points, see [22, Section 5] for further details.

tion on g_1, \dots, g_n , and then constructs the BlockJS-based estimator. The min-max upper and lower bounds of this estimator are investigated in Section 4. Section 5 describes and discusses the simulation results, before drawing some conclusions in Section 6. The proofs are deferred to Section 7 awaiting inspection by the interested reader.

2. Wavelets and Besov balls

2.1. 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 . These wavelets enjoy the following features.

- They are smooth and frequency band-limited, i.e. the Fourier transforms of ϕ and ψ have compact supports with

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

where supp denotes the support.

- The functions (ϕ, ψ) are C^∞ as their Fourier transforms have a compact support, and ψ has an infinite number of vanishing moments as its Fourier transform vanishes in a neighborhood of the origin:

$$\int_{-\infty}^{+\infty} t^u \psi(t) dt = 0, \quad \forall u \in \mathbb{N}. \quad (4)$$

- If the Fourier transforms of ϕ and ψ are also in C^m for a chosen $m \in \mathbb{N}$, then it can be easily shown that ϕ and ψ obey

$$|\phi(t)| = O((1 + |t|)^{-m-1}), \quad |\psi(t)| = O((1 + |t|)^{-m-1}) \quad (5)$$

for every $t \in \mathbb{R}$.

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

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

be the elements of the wavelet basis, and

$$\phi_{j,k}^{\text{per}}(t) = \sum_{\ell \in \mathbb{Z}} \phi_{j,k}(t - \ell), \quad \psi_{j,k}^{\text{per}}(t) = \sum_{\ell \in \mathbb{Z}} \psi_{j,k}(t - \ell),$$

their periodized versions. There exists an integer j_* such that the collection $\{\phi_{j_*,k}^{\text{per}}(\cdot), k \in \{0, \dots, 2^{j_*} - 1\}; \psi_{j,k}^{\text{per}}(\cdot), j \geq j_*, k \in \{0, \dots, 2^j - 1\}\}$ forms an orthonormal basis of $\mathbb{L}_{\text{per}}^2([0, 1])$. In what follows, the superscript "per" will be

dropped from ϕ^{per} and ψ^{per} to lighten the notation.

Let $l \geq j_*$, any function $h \in \mathbb{L}_{\text{per}}^2([0, 1])$ can be expanded into a wavelet series as

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

where

$$\alpha_{l,k} = \int_0^1 h(t) \bar{\phi}_{l,k}(t) dt, \quad \beta_{j,k} = \int_0^1 h(t) \bar{\psi}_{j,k}(t) dt. \quad (6)$$

See [21, Vol. 1 Chapter III.11] for a detailed account on periodized orthonormal wavelet bases.

2.2. Besov balls

Let $0 < M < \infty$, $s > 0$, $1 \leq p, r \leq \infty$. Among the several characterizations of Besov spaces for periodic functions on $\mathbb{L}^p([0, 1])$, we will focus on the usual one based on the corresponding coefficients in a sufficiently q -regular (periodized) wavelet basis ($q = \infty$ for Meyer wavelets). More precisely, we say that a function h belongs to the Besov ball $\mathbf{B}_{p,r}^s(M)$ if and only if $\int_0^1 |h(t)|^p dt \leq M$, and there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients (6) satisfy

$$2^{j_*(1/2-1/p)} \left(\sum_{k=0}^{2^{j_*}-1} |\alpha_{j_*,k}|^p \right)^{1/p} + \left(\sum_{j=j_*}^{\infty} \left(2^{j(s+1/2-1/p)} \left(\sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \right)^{1/p} \right)^r \right)^{1/r} \leq M^*, \quad (7)$$

with a smoothness parameter $0 < s < q$, and the norm parameters p and r . Besov spaces capture a variety of smoothness features in a function including spatially inhomogeneous behavior, see [21].

3. The deconvolution BlockJS estimator

3.1. The ordinary smoothness assumption

In this study, we focus on the following particular ordinary smoothness assumption on g_1, \dots, g_n . We assume that there exist three constants, $c_g > 0$, $C_g > 0$ and $\delta > 1$, and n positive real numbers $\sigma_1, \dots, \sigma_n$ such that, for any $\ell \in \mathbb{Z}$ and any $v \in \{1, \dots, n\}$,

$$c_g \frac{1}{(1 + \sigma_v^2 \ell^2)^{\delta/2}} \leq |\mathcal{F}(g_v)(\ell)| \leq C_g \frac{1}{(1 + \sigma_v^2 \ell^2)^{\delta/2}}. \quad (8)$$

This assumption controls the decay of the Fourier coefficients of g_1, \dots, g_n , and thus the smoothness of g_1, \dots, g_n . It is a standard hypothesis usually adopted in the field of nonparametric estimation for deconvolution problems. See e.g. [25], [17] and [19].

Example 3.1. Let τ_1, \dots, τ_n be n positive real numbers. For any $v \in \{1, \dots, n\}$, consider the square-integrable 1-periodic function g_v defined by

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

Then, for any $\ell \in \mathbb{Z}$, $\mathcal{F}(g_v)(\ell) = 2(1 + 4\pi^2 \ell^2 \tau_v^2)^{-1}$ and (8) is satisfied with $\delta = 2$ and $\sigma_v = 2\pi\tau_v$.

In the sequel, we set

$$\rho_n = \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^\delta}. \quad (9)$$

For a technical reason that is not restrictive at all (see Section 7), we suppose that $\rho_n \geq e$ and $\lim_{n \rightarrow \infty} (\ln \rho_n)^v \rho_n^{-1} = 0$ for any $v > 0$.

3.2. BlockJS estimator

We suppose that $f^{(d)} \in \mathbb{L}_{\text{per}}^2([0, 1])$ and that the ordinary smoothness assumption (8) holds, where δ refers to the exponent in the assumption. We are ready to construct our adaptive procedure for the estimation of $f^{(d)}$.

Let $j_1 = \lfloor \log_2(\log \rho_n) \rfloor$ be the coarsest resolution level, and $j_2 = \lfloor (1/(2\delta + 2d + 1)) \log_2(\rho_n / \log \rho_n) \rfloor$, where, for any $a \in \mathbb{R}$, $\lfloor a \rfloor$ denotes the integer part of a . For any $j \in \{j_1, \dots, j_2\}$, let $L = \lfloor \log \rho_n \rfloor$ be the block size.

Let $A_j = \{1, \dots, \lfloor 2^j L^{-1} \rfloor\}$ be the set indexing the blocks at resolution j . For each j , let $\{B_{j,K}\}_{K \in A_j}$ be a uniform and disjoint open covering of $\{0, \dots, 2^j - 1\}$, i.e. $\bigcup_{K \in A_j} B_{j,K} = \{0, \dots, 2^j - 1\}$ and for any $(K, K') \in A_j^2$ with $K \neq K'$, $B_{j,K} \cap B_{j,K'} = \emptyset$ and $\text{Card}(B_{j,K}) = L$, where $B_{j,K} = \{k \in \{0, \dots, 2^j - 1\}; (K-1)L \leq k \leq KL - 1\}$ is the K th block.

We define the Block James-Stein estimator (BlockJS) of $f^{(d)}$ by

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

where for any resolution j and position $k \in B_{j,K}$ within the K th block, the wavelet coefficients of $f^{(d)}$ are estimated via the rule

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

with, for any $a \in \mathbb{R}$, $(a)_+ = \max(a, 0)$, $\lambda > 0$, and $\widehat{\alpha}_{j_1, k}$ and $\widehat{\beta}_{j, k}$ are respectively the empirical scaling and detail coefficients, defined as

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

and

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

Notice that thanks to (3), for any $j \in \{j_1, \dots, j_2\}$ and $k \in \{0, \dots, 2^j - 1\}$

$$\begin{cases} \mathcal{D}_{j_1} = \text{supp}(\mathcal{F}(\phi_{j_1, k})) \subset [-4\pi 3^{-1} 2^{j_1}, 4\pi 3^{-1} 2^{j_1}], \\ \mathcal{C}_j = \text{supp}(\mathcal{F}(\psi_{j, k})) \subset [-8\pi 3^{-1} 2^j, -2\pi 3^{-1} 2^j] \cup [2\pi 3^{-1} 2^j, 8\pi 3^{-1} 2^j]. \end{cases} \quad (11)$$

4. Minimaxity results of BlockJS over Besov balls

4.1. Minimax upper-bound for the MISE

Theorem 4.1 below determines the rates of convergence achieved by $\widehat{f^{(d)}}$ under the MISE over Besov balls.

Theorem 4.1. *Consider the model (1) and recall that we want to estimate $f^{(d)}$ with $d \in \mathbb{N}$. Assume that (ϕ, ψ) satisfy (5) for some $m \geq d$ and (8) is satisfied. Let $\widehat{f^{(d)}}$ 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*

$$\sup_{f^{(d)} \in \mathbf{B}_{p, r}^s(M)} \mathbb{E} \left(\int_0^1 \left(\widehat{f^{(d)}}(t) - f^{(d)}(t) \right)^2 dt \right) \leq C \varphi_n,$$

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 4.1 will be proved using the more general theorem [12, Theorem 3.1]. To apply this result, two conditions on the wavelet coefficients estimator are required: a moment condition and a concentration condition. They are established in Propositions 7.2 and 7.3, see Section 7.

Remark 4.1. *Theorem 4.1 can be generalized by allowing the decay exponent δ of g_1, \dots, g_n to vary across the channels. More precisely, our ordinary assumption (8) is uniform over the channels, in the sense that we assumed that the decay exponent δ is the same for all v . Assume now that there exist three*

constants, $c_g > 0$, $C_g > 0$, and $2n$ positive real numbers $\sigma_1, \dots, \sigma_n, \delta_1, \dots, \delta_n$ with $\min(\delta_1, \dots, \delta_n) > 1/2$ such that, for any $\ell \in \mathbb{Z}$ and any $v \in \{1, \dots, n\}$,

$$|\mathcal{F}(g_v)(\ell)| \leq C_g \frac{1}{(1 + \sigma_v^2 \ell^2)^{\delta_v/2}}.$$

Set

$$\omega = \min(\delta_1, \dots, \delta_n).$$

Let us now define the BlockJS estimator of $f^{(d)}$ as in Section 3.2 but with the weights

$$\rho_n = \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^{\delta_v}},$$

resolution levels $j_1 = \lfloor \log_2(\log \rho_n) \rfloor$, $j_2 = \lfloor (1/(2\omega + 2d + 1)) \log_2(\rho_n / \log \rho_n) \rfloor$, and

$$\begin{aligned} \widehat{\beta}_{j,k}^* &= \widehat{\beta}_{j,k} \left(1 - \frac{\lambda \epsilon^2 \rho_n^{-1} 2^{2j(\omega+d)}}{\frac{1}{L} \sum_{k \in B_{j,K}} |\widehat{\beta}_{j,k}|^2} \right)_+, \\ \widehat{\alpha}_{j_1,k} &= \frac{1}{\rho_n} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^{\delta_v}} \sum_{\ell \in \mathcal{D}_{j_1}} (2\pi i \ell)^d \frac{\overline{\mathcal{F}(\phi_{j_1,k})(\ell)}}{\mathcal{F}(g_v)(\ell)} \int_0^1 e^{-2\pi i \ell t} dY_v(t) \end{aligned}$$

and

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

Then Theorem 4.1 holds with the convergence rate

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

Of course, this generalization encompasses the statement of Theorem 4.1. This generalized result also tells us that the adaptivity of the estimator makes its performance mainly influenced by the best g_i , that is the one with the smallest δ_i , which is a nice feature.

4.2. Minimax lower-bound for the MISE

We now turn to the lower bound of the MISE to formally answer the question whether φ_n is indeed the optimal rate of convergence or not. This is the goal of Theorem 4.2 which gives a positive answer.

Theorem 4.2. Consider the model (1) and recall that we want to estimate $f^{(d)}$ with $d \in \mathbb{N}$. Assume that (8) 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

$$\inf_{f^{(d)}} \sup_{f^{(d)} \in \mathbf{B}_{p,r}^s(M)} \mathbb{E} \left(\int_0^1 \left(\widetilde{f^{(d)}}(t) - f^{(d)}(t) \right)^2 dt \right) \geq c \varphi_n^*,$$

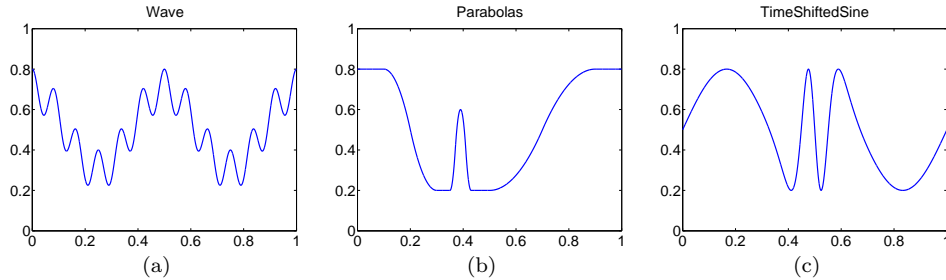


FIG 1. Original Signals (a): Wave. (b): Parabolas. (c): TimeShiftedSine.

where

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

and the infimum is taken over all the estimators $\widetilde{f}^{(d)}$ of $f^{(d)}$.

It can then be concluded from Theorem 4.1 and Theorem 4.2 that the rate of convergence φ_n achieved by $\widetilde{f}^{(d)}$ is near optimal in the minimax sense. Near minimaxity is only due to the case $p \in [1, 2)$ and $s > (1/p - 1/2)(2\delta + 2d + 1)$ where there is an extra logarithmic term.

5. 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 [20]) representing different degrees of smoothness were used (see FIG 1). The “Wave” function was used to illustrate the performance of our estimator on a smooth function. Note that the “Parabolas” function has big jumps in its second derivative.

We have compared the numerical performance of BlockJS to state-of-the-art classical thresholding methods of the literature. In particular we consider the block estimator of [23] 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 [18]). In the sequel, we name the estimator of [23] by ‘BlockH’, the one of [18] by ‘TermJS’ and our estimator by ‘BlockJS’. For numerical implementation, the test functions were finely discretized by taking T equispaced samples $t_i = i/T \in [0, 1]$, $i = 0, \dots, T - 1$. The deconvolution estimator was efficiently implemented in the Fourier domain given that Meyer wavelets are frequency band-limited. The performance of the estimators are measured in terms of peak signal-to-noise ratio (PSNR = $10 \log_{10} \frac{\max_{t_i \in [0,1]} |f^{(d)}(t_i)|^2}{\sum_{i=0}^{T-1} (f^{(d)}(t_i) - \widehat{f}^{(d)}(t_i))^2 / T}$) in decibels (dB)). For any

$v \in \{1, \dots, n\}$, the blurring function g_v is that of Example 3.1 and was used throughout all experiments.

5.1. 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 samples generated according to (1) with blurred signal-to-noise ratio (BSNR) of 25 dB ($\text{BSNR} = 10 \log_{10} \frac{\sum_{i=0}^{T-1} (f \star g_v)(t_i)^2}{T \epsilon^2}$ dB). For $d = 0$, the results are very effective for each test function where the singularities are well estimated. The estimator does also a good job in estimating the first and second-order derivatives, although the estimation quality decreases as the order of the derivative increases. This is in agreement with the predictions of the minimaxity results. We then have compared the performance of BlockJS with BlockH. The blurred signals were corrupted by a zero-mean white Gaussian noise such that the BSNR ranged from 10 to 40 dB. The PSNR values averaged over 10 noise realizations are depicted in FIG 3 for $d = 0$, $d = 1$ and $d = 2$ respectively. One can see that our BlockJS thresholding estimator produces quite accurate estimates of f , f' and f'' for each test function. These results clearly show that our approach compares favorably to BlockH and that BlockJS has good adaptive properties over a wide range of noise levels in the monochannel setting.

5.2. Multichannel simulation

A first point we would like to highlight is the fact that some choices of $\sigma_1, \dots, \sigma_n$ can severely impact the performance of the estimators. To illustrate this, we show in FIG 4 an example of first derivative estimates obtained using BlockJS from $n = 10$ channels with $T = 4096$ samples and noise level corresponding to $\text{BSNR} = 25$ dB, for $\sigma_v = v$ (dashed blue) and σ_v randomly generated in $(0, +\infty)$ (solid blue). With σ_v randomly generated, we can observe 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 finally report a simulation study by quantitatively comparing BlockJS to the other thresholding estimators described above. For each test function, we generated $T = 4096$ equispaced samples on $[0, 1]$ according to (1) with varying number of channels ranging from $n = 10$ to 100.

TABLE 1 summarizes the results. It shows in particular that BlockJS consistently outperforms the other methods in almost all cases in terms of PSNR. As expected and predicted by our theoretical findings, on the one hand, the performance gets better as the number of channels increases. On the other hand, it degrades with increasing noise level and/or d . Indeed, the derivatives estimation for $\text{BSNR} = 10$ dB is rather difficult to estimate, especially for functions having highly irregular derivatives such as “Parabolas” (which has big jumps in its second derivative, see FIG. 2(d)).

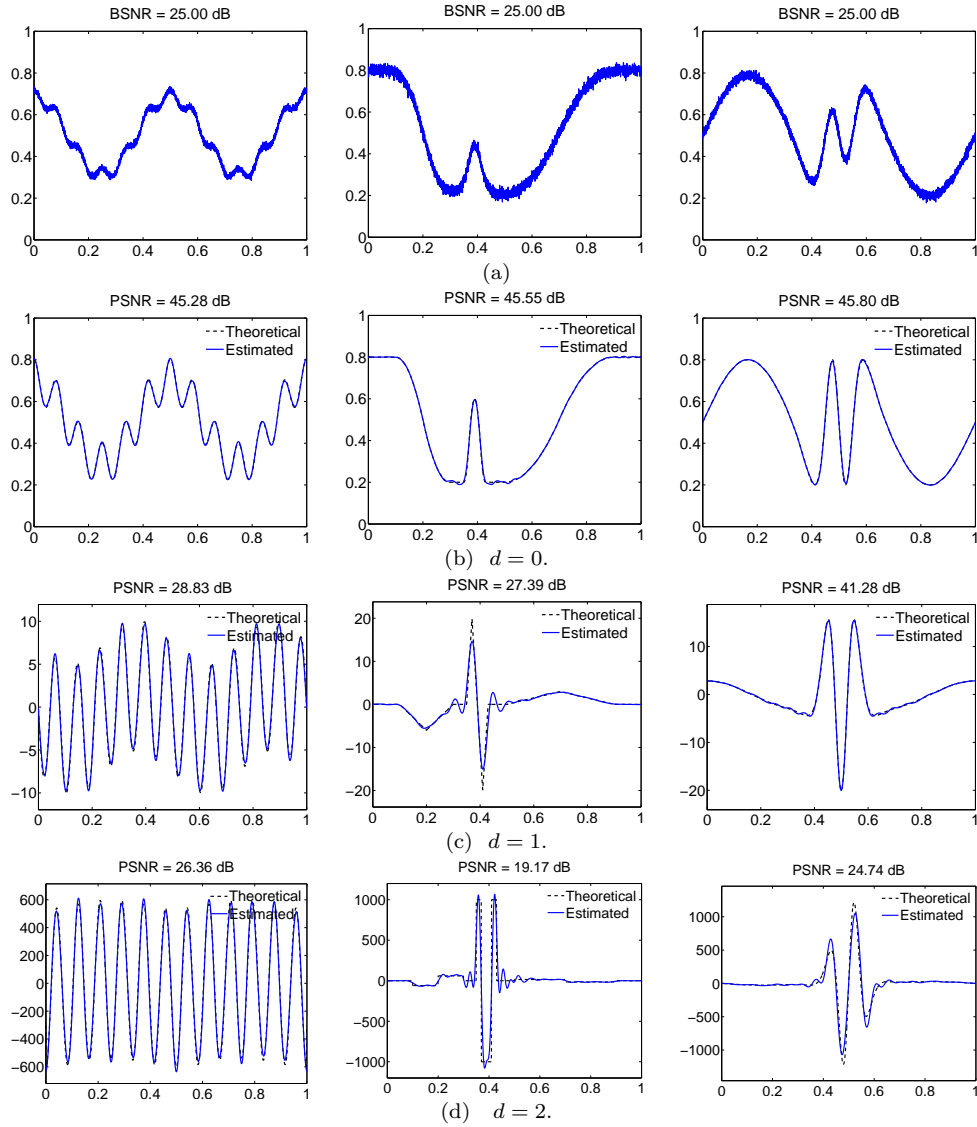


FIG 2. Original (dashed) and estimated function/derivatives (solid) using the BlockJS estimator applied to noisy blurred observations shown in (a). (b): $d = 0$. (c): $d = 1$ (d): $d = 2$. From left to right Wave, Paraboloids and TimeShiftedSine.

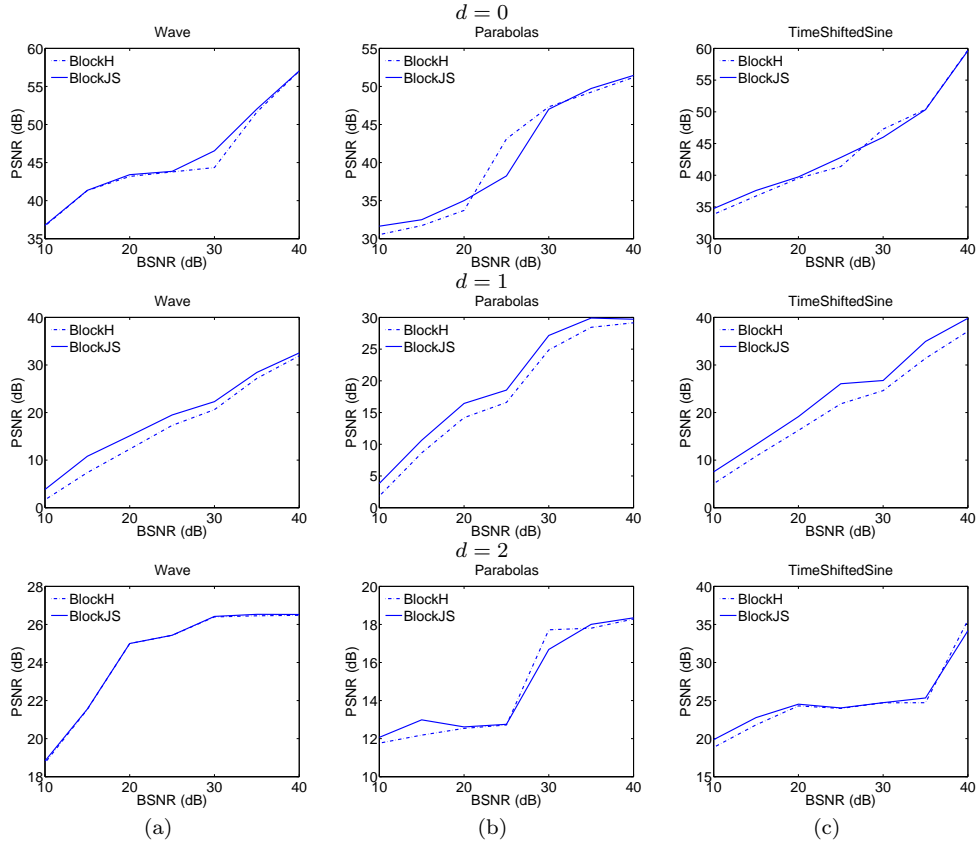


FIG 3. Averaged PSNR values as a function of the input BSNR from 10 replications of the noise. (a): Wave. (b): Parabolos. (c): TimeShiftedSine. From top to bottom $d = 0, 1, 2$.

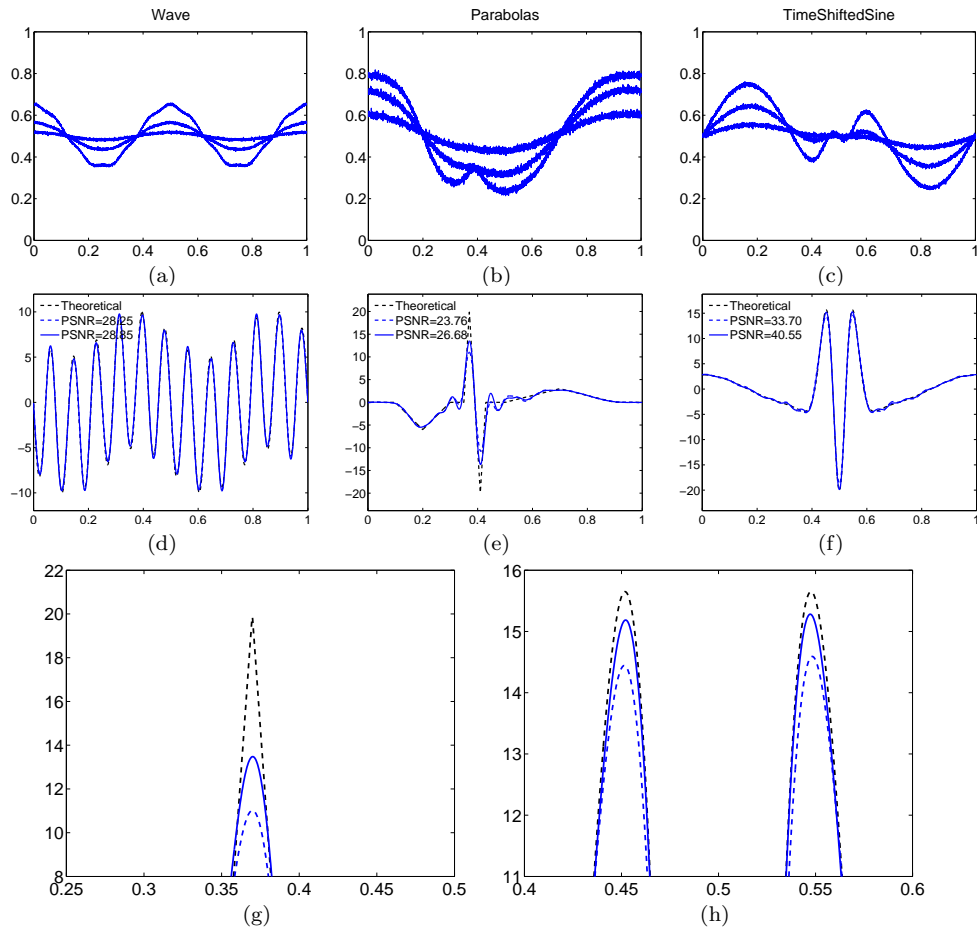


FIG 4. Original functions (dashed black) and the estimate for $\sigma_v = v$ (dashed blue) and σ_v randomly generated (solid blue) with $n = 10$ channels. (a)-(c): noisy blurred observations (3 channels out of 10 shown). (d)-(f) BlockJS estimates of the first derivative. Zoom on the estimates (g): Parabolas, (h): TimeShiftedSine.

BSNR= 40 dB												
n	d = 0				d = 1				d = 2			
	10	20	50	100	10	20	50	100	10	20	50	100
<i>Wave</i>												
BlockJS	57.42	66.40	66.83	74.62	42.64	43.58	43.94	50.10	22.04	30.34	33.88	36.69
BlockH	57.34	66.31	66.78	74.72	42.43	43.57	43.07	50.07	22.66	29.14	33.78	36.65
TermJS	52.51	61.94	64.86	73.57	40.64	41.56	33.90	49.30	19.17	29.57	30.87	33.20
TermH	50.97	52.72	55.69	74.73	31.39	33.92	37.20	39.48	17.17	28.82	33.89	35.61
<i>TimeShiftedSine</i>												
BlockJS	65.11	65.58	68.47	71.17	42.16	46.25	46.85	49.53	41.09	42.67	43.07	46.35
BlockH	62.11	62.29	62.71	70.43	41.97	45.57	45.75	49.35	39.20	40.55	41.97	42.00
TermJS	64.01	64.73	66.13	68.21	40.96	42.13	43.63	43.98	38.57	41.17	41.73	45.15
TermH	61.84	62.05	67.12	68.80	41.22	42.12	43.32	44.69	39.39	39.81	39.24	43.04
<i>Parabolas</i>												
BlockJS	56.18	57.42	58.10	58.40	29.66	29.76	31.04	31.40	20.63	21.88	21.93	21.90
BlockH	55.92	57.25	57.70	58.10	29.69	29.67	29.91	29.94	20.74	21.09	21.99	21.73
TermJS	54.33	57.05	57.96	58.24	29.21	29.38	29.65	29.79	20.57	20.93	20.91	21.51
TermH	54.59	56.96	57.88	58.10	29.60	29.60	29.90	29.94	20.51	20.96	20.84	21.49

BSNR= 25 dB												
n	d = 0				d = 1				d = 2			
	10	20	50	100	10	20	50	100	10	20	50	100
<i>Wave</i>												
BlockJS	44.04	51.93	52.47	59.76	30.14	30.90	30.90	35.93	26.83	26.83	26.98	27.00
BlockH	42.12	51.35	51.82	59.73	28.69	28.69	28.72	35.26	26.89	26.85	26.98	26.48
TermJS	41.67	48.10	49.44	52.59	28.49	28.69	28.72	28.71	25.25	26.62	26.78	26.94
TermH	40.95	49.03	50.22	55.04	28.69	28.69	28.72	28.72	25.28	26.85	26.98	26.86
<i>TimeShiftedSine</i>												
BlockJS	51.85	52.33	55.66	60.49	39.48	41.46	41.88	41.92	27.39	28.72	29.05	35.95
BlockH	52.93	51.51	55.82	60.35	38.68	41.24	41.87	41.84	26.68	29.36	29.54	35.15
TermJS	47.19	47.83	54.45	56.63	29.46	41.34	41.85	41.79	23.66	25.84	25.95	27.58
TermH	47.54	47.47	54.44	59.63	31.42	41.03	40.75	41.79	23.66	25.69	25.91	30.67
<i>Parabolas</i>												
BlockJS	47.81	49.88	52.90	54.40	25.52	26.11	28.74	29.57	17.48	18.24	18.96	20.62
BlockH	47.74	49.23	52.00	53.62	25.43	25.16	29.55	29.69	17.08	18.92	19.00	20.63
TermJS	44.52	49.80	50.80	52.79	24.84	25.78	25.71	25.73	16.39	18.58	19.00	20.57
TermH	43.74	48.84	51.35	53.24	24.86	25.17	25.10	27.98	16.49	18.64	18.95	20.34

BSNR= 10 dB												
n	d = 0				d = 1				d = 2			
	10	20	50	100	10	20	50	100	10	20	50	100
<i>Wave</i>												
BlockJS	35.82	43.46	43.91	45.95	26.96	27.34	28.22	28.45	19.02	25.12	26.16	26.78
BlockH	34.08	43.41	43.83	44.60	26.90	27.34	28.19	28.41	19.21	25.16	26.06	26.68
TermJS	33.30	43.16	43.24	44.48	26.67	27.25	28.14	28.36	18.32	18.89	22.87	26.65
TermH	33.22	40.16	39.32	44.27	26.77	27.35	28.15	28.39	18.61	19.07	18.95	26.68
<i>TimeShiftedSine</i>												
BlockJS	38.69	39.28	41.24	48.38	30.12	38.85	39.26	40.19	22.58	22.84	23.37	23.53
BlockH	38.66	38.97	40.12	45.34	26.66	38.22	39.22	39.08	22.58	22.73	23.33	23.48
TermJS	38.41	38.68	40.31	41.45	25.68	29.24	36.79	36.68	22.52	22.82	15.27	23.26
TermH	38.23	38.44	39.07	41.46	26.98	28.10	35.18	36.98	22.49	22.72	15.47	17.56
<i>Parabolas</i>												
BlockJS	35.26	37.04	40.64	44.74	22.03	24.56	24.77	25.56	12.76	12.79	13.09	13.29
BlockH	34.12	35.53	39.29	43.41	22.22	24.54	24.47	25.00	12.64	12.76	12.78	12.77
TermJS	34.29	35.18	39.38	41.37	21.70	24.51	24.63	24.96	12.56	12.76	12.79	12.79
TermH	33.27	34.53	39.45	42.21	21.88	23.98	24.21	24.86	12.70	12.76	12.79	12.79

TABLE 1
 Comparison of average PSNR in decibels (dB) over 10 realizations of the noise for $d = 0$, $d = 1$ and $d = 2$. From top to bottom BSNR= 40, 25, 10 dB.

6. 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, \dots, 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. Preparatory results

In the three following results, we consider the framework of Theorem 4.1 and, for any integer $j \geq j_*$ and $k \in \{1, \dots, 2^j - 1\}$, we set $\alpha_{j,k} = \int_0^1 f^{(d)}(t) \bar{\phi}_{j,k}(t) dt$ and $\beta_{j,k} = \int_0^1 f^{(d)}(t) \bar{\psi}_{j,k}(t) dt$, the wavelet coefficients (6) of $f^{(d)}$.

Proposition 7.1 (Gaussian distribution on the wavelet coefficient estimators). *For any integer $j \geq j_*$ and $k \in \{0, \dots, 2^j - 1\}$, we have*

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

and

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

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

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

It follows from (1) that

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

Note that, since f is 1-periodic, for any $u \in \{0, \dots, d\}$, $f^{(u)}$ is 1-periodic and $f^{(u)}(0) = f^{(u)}(1)$. By classical properties of the Fourier series, for any $\ell \in \mathbb{Z}$, we have $\mathcal{F}(f^{(d)})(\ell) = (2\pi i\ell)^d \mathcal{F}(f)(\ell)$. The 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 orthonormal basis of $\mathbb{L}_{\text{per}}^2([0, 1])$ 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 7.1 is proved. \square

Proposition 7.2 (Moment inequalities).

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

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

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

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

Proof. Let us prove the second point, the first one can be proved in a similar way. Let us recall that, by Proposition 7.1, for any $j \in \{j_1, \dots, j_2\}$ and any $k \in \{0, \dots, 2^j - 1\}$, we have

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

where

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

Due to (8) and (11), for any $v \in \{1, \dots, n\}$, we have

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

It follows from (15) and the Parseval identity that

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

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 7.2 is proved. \square

Proposition 7.3 (Concentration inequality). *There exists a constant $\lambda > 0$ such that, for any $j \in \{j_1, \dots, j_2\}$, any $K \in \mathcal{A}_j$ and n large enough,*

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

Proof. We need the Tsirelson inequality stated in Lemma 7.1 below.

Lemma 7.1 ([13]). *Let $(\vartheta_t)_{t \in D}$ be a centered Gaussian process. If*

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

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 notational clarity, let

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

Recall that, by Proposition 7.1, we have $V_{j,k} \sim \mathcal{N} \left(0, \rho_n^{-2} \sigma_{j,k}^2 \right)$, where $\sigma_{j,k}^2$ is given in (14). Let $\mathbb{B}(1)$ the unit 2-norm ball in $\mathbb{C}^{\text{Card}(B_{j,K})}$, i.e. $\mathbb{B}(1) = \{a \in \mathbb{C}^{\text{Card}(B_{j,K})}; \sum_{k \in B_{j,K}} |a_k|^2 \leq 1\}$. For any $a \in \mathbb{B}(1)$, let $Z(a)$ be the centered Gaussian process defined by

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

By a simple Legendre-Fenchel conjugacy argument, we have

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

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

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

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

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 $(v, v') \in \{1, \dots, n\}^2$,

$$\mathbb{E} (e_{\ell,n} \bar{e}_{\ell',v'}) = \begin{cases} 1 & \text{if } \ell = \ell' \text{ and } v = v', \\ 0 & \text{otherwise.} \end{cases}$$

It then follows that

$$\begin{aligned}
\sup_{a \in \mathbb{B}(1)} \mathbb{V}(Z(a)) &= \sup_{a \in \mathbb{B}(1)} \mathbb{E} \left(\left| \sum_{k \in B_{j,\kappa}} a_k V_{j,k} \right|^2 \right) \\
&= \sup_{a \in \mathbb{B}(1)} \mathbb{E} \left(\sum_{k \in B_{j,\kappa}} \sum_{k' \in B_{j,\kappa}} a_k \bar{a}_{k'} V_{j,k} \bar{V}_{j,k'} \right) \\
&= \epsilon^2 \rho_n^{-2} \sup_{a \in \mathbb{B}(1)} \sum_{k \in B_{j,\kappa}} \sum_{k' \in B_{j,\kappa}} a_k \bar{a}_{k'} \sum_{\ell \in \mathcal{C}_j} \sum_{\ell' \in \mathcal{C}_j} \sum_{v=1}^n \sum_{v'=1}^n \frac{1}{(1 + \sigma_v^2)^\delta} \frac{1}{(1 + \sigma_{v'}^2)^\delta} \times \\
&\quad \frac{(2\pi i \ell)^d}{\mathcal{F}(g_v)(\ell)} \overline{\mathcal{F}(\psi_{j,k})(\ell)} \frac{(2\pi i \ell')^d}{\overline{\mathcal{F}(g_{v'})(\ell')}} \mathcal{F}(\psi_{j,k'})(\ell') \mathbb{E}(e_{\ell,v} \bar{e}_{\ell',v'}) \\
&= \epsilon^2 \rho_n^{-2} \sup_{a \in \mathbb{B}(1)} \sum_{k \in B_{j,\kappa}} \sum_{k' \in B_{j,\kappa}} a_k \bar{a}_{k'} \sum_{\ell \in \mathcal{C}_j} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^{2\delta}} \frac{(2\pi \ell)^{2d}}{|\mathcal{F}(g_v)(\ell)|^2} \overline{\mathcal{F}(\psi_{j,k})(\ell)} \mathcal{F}(\psi_{j,k'})(\ell) \\
&= \epsilon^2 \rho_n^{-2} \sup_{a \in \mathbb{B}(1)} \sum_{\ell \in \mathcal{C}_j} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^{2\delta}} \frac{(2\pi \ell)^{2d}}{|\mathcal{F}(g_v)(\ell)|^2} \left| \sum_{k \in B_{j,\kappa}} \bar{a}_k \mathcal{F}(\psi_{j,k})(\ell) \right|^2. \quad (17)
\end{aligned}$$

For any $a \in \mathbb{B}(1)$, the Parseval identity and the fact that $\{\psi_{j,k}\}_{k=0,\dots,2^j-1}$ are orthonormal yields

$$\begin{aligned}
\sum_{\ell \in \mathcal{C}_j} \left| \sum_{k \in B_{j,\kappa}} \bar{a}_k \mathcal{F}(\psi_{j,k})(\ell) \right|^2 &= \sum_{\ell \in \mathcal{C}_j} \left| \mathcal{F} \left(\sum_{k \in B_{j,\kappa}} \bar{a}_k \psi_{j,k} \right) (\ell) \right|^2 \\
&= \int_0^1 \left| \sum_{k \in B_{j,\kappa}} \bar{a}_k \psi_{j,k}(t) \right|^2 dt \\
&= \sum_{k \in B_{j,\kappa}} |a_k|^2 \leq 1. \quad (18)
\end{aligned}$$

Piecing (15), (17) and (18) together, we get

$$\begin{aligned}
\sup_{a \in \mathbb{B}(1)} \mathbb{V}(Z(a)) &\leq C \epsilon^2 \rho_n^{-1} 2^{2(\delta+d)j} \sup_{a \in \mathbb{B}(1)} \sum_{\ell \in \mathcal{C}_j} \left| \sum_{k \in B_{j,\kappa}} \bar{a}_k \mathcal{F}(\psi_{j,k})(\ell) \right|^2 \\
&\leq C \epsilon^2 \rho_n^{-1} 2^{2(\delta+d)j}.
\end{aligned}$$

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

Taking λ large enough and $x = 2^{-1} \lambda \epsilon 2^{(\delta+d)j} (\log \rho_n / \rho_n)^{1/2}$, the Tsirelson

inequality (see Lemma 7.1) yields

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

Proposition 7.3 is proved. \square

7.2. Proof of Theorem 4.1

Proof. Plugging Propositions 7.2 and 7.3 into [12, Theorem 3.1], we end the proof of Theorem 4.1. \square

7.3. Proof of Theorem 4.2

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

Lemma 7.2. *Let $m \in \mathbb{N}^*$ and A be a σ -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 $i \neq j$,*

$$A_i \cap A_j = \emptyset.$$

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

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

where

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

A^c denotes the complement of A in Ω and K is the Kullbak-Leibler divergence defined by

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

The proof of Lemma 7.2 can be found in [15, Lemma 3.3]. For further details and applications of the Fano lemma, see [27].

In what follows, we distinguish ψ and ψ^{per} (see Section 2.1). For any integrable function h on \mathbb{R} , we set

$$\mathcal{F}_*(h)(\ell) = \int_{-\infty}^{\infty} h(t)e^{-2i\pi\ell t} dt, \quad \ell \in \mathbb{Z}.$$

Consider the Besov balls $\mathbf{B}_{p,r}^s(M)$ (see (7)). 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

$$h_\varepsilon(t) = M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \varepsilon_k \frac{1}{(d-1)!} \sum_{l \in \mathbb{Z}} \int_{-\infty}^{t+l} (t+l-y)^{d-1} \psi_{j_0,k}(y) dy, \quad t \in [0, 1],$$

and, if $d = 0$, set $h_\varepsilon(t) = M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \varepsilon_k \psi_{j_0,k}^{\text{per}}(t)$, $t \in [0, 1]$. Notice that, owing to (5), h_ε exists. Moreover, it is 1-periodic. Using the Cauchy formula for repeated integration, we have

$$h_\varepsilon^{(d)}(t) = M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \varepsilon_k \psi_{j_0,k}^{\text{per}}(t), \quad t \in [0, 1].$$

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

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

Therefore $h_\varepsilon^{(d)} \in \mathbf{B}_{p,r}^s(M)$. The Varshamov-Gilbert theorem (see [27, Lemma 2.7]) asserts that there exist a set $E_{j_0} = \{\varepsilon^{(0)}, \dots, \varepsilon^{(T_{j_0})}\}$ and two constants, $c \in]0, 1[$ and $\alpha \in]0, 1[$, such that, for any $u \in \{0, \dots, T_{j_0}\}$, $\varepsilon^{(u)} = (\varepsilon_k^{(u)})_{k \in \{0, \dots, 2^{j_0} - 1\}} \in \{0, 1\}^{2^{j_0}}$ and any $(u, v) \in \{0, \dots, T_{j_0}\}^2$ with $u < v$, the following inequalities hold:

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

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

by orthonormality of the collection $\{\psi_{j_0,k}\}_{k=0,\dots,2^{j_0}-1}$

$$\begin{aligned} & \left(\int_0^1 \left(h_{\varepsilon^{(u)}}^{(d)}(t) - h_{\varepsilon^{(v)}}^{(d)}(t) \right)^2 dt \right)^{1/2} \\ &= M_* c 2^{-j_0(s+1/2)} \left(\sum_{k=0}^{2^{j_0}-1} \left| \varepsilon_k^{(u)} - \varepsilon_k^{(v)} \right|^2 \right)^{1/2} \\ &= M_* c 2^{-j_0(s+1/2)} \left(\sum_{k=0}^{2^{j_0}-1} \left| \varepsilon_k^{(u)} - \varepsilon_k^{(v)} \right| \right)^{1/2} \\ &\geq 2\delta_{j_0}, \end{aligned}$$

where

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

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

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

where

$$A_u = \left\{ \left(\int_0^1 \left(\widetilde{f}^{(d)}(t) - h_{\varepsilon^{(u)}}^{(d)}(t) \right)^2 dt \right)^{1/2} < \delta_{j_0} \right\}$$

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

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

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

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

The Parseval identity yields

$$\begin{aligned} K(\mathbb{P}_{f_1}, \mathbb{P}_{f_2}) &= \frac{1}{2\epsilon^2} \sum_{v=1}^n \sum_{\ell \in \mathbb{Z}} |\mathcal{F}((f_1 - f_2) \star g_v)(\ell)|^2 \\ &= \frac{1}{2\epsilon^2} \sum_{v=1}^n \sum_{\ell \in \mathbb{Z}} |\mathcal{F}(f_1 - f_2)(\ell)|^2 |\mathcal{F}(g_v)(\ell)|^2. \end{aligned}$$

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

$$K\left(\mathbb{P}_{h_{\epsilon^{(u)}}}, \mathbb{P}_{h_{\epsilon^{(v)}}}\right) = \frac{1}{2\epsilon^2} \sum_{v=1}^n \sum_{\ell \in \mathbb{Z}} |\mathcal{F}(h_{\epsilon^{(u)}} - h_{\epsilon^{(v)}})(\ell)|^2 |\mathcal{F}(g_v)(\ell)|^2. \quad (20)$$

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

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

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

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

Then, for any $u \in \{0, \dots, d\}$, thanks to (4), $\lim_{|t| \rightarrow \infty} \theta_k^{(u)}(t) = 0$. Consequently, for any $\ell \in \mathbb{Z}$,

$$(2\pi i \ell)^d \mathcal{F}_*(\theta_k)(\ell) = \mathcal{F}_* \left(\theta_k^{(d)} \right) (\ell) = \mathcal{F}(\psi_{j_0, k}^{\text{per}})(\ell).$$

So, for any $\ell \in \mathcal{C}_{j_0}$ (excluding 0), (21) implies that

$$\begin{aligned} &\mathcal{F}(h_{\epsilon^{(u)}} - h_{\epsilon^{(v)}})(\ell) \\ &= M_* 2^{-j_0(s+1/2)} \sum_{k=0}^{2^{j_0}-1} \left(\varepsilon_k^{(u)} - \varepsilon_k^{(v)} \right) \frac{1}{(2\pi i \ell)^d} \mathcal{F} \left(\psi_{j_0, k}^{\text{per}} \right) (\ell), \end{aligned} \quad (22)$$

which entails in particular that $\text{supp}(\mathcal{F}(h_{\varepsilon^{(u)}} - h_{\varepsilon^{(v)}})) = \mathcal{C}_{j_0}$. This in conjunction with equalities (20) and (22) imply that

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

By assumptions (8) and (11), for any $v \in \{1, \dots, n\}$,

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

Moreover, the Parseval identity and orthonormality of the family $\left\{ \psi_{j_0,k}^{\text{per}} \right\}_{k=0, \dots, 2^{j_0}-1}$ imply that

$$\begin{aligned} & \sum_{\ell \in \mathcal{C}_{j_0}} \left| \sum_{k=0}^{2^{j_0}-1} (\varepsilon_k^{(u)} - \varepsilon_k^{(v)}) \mathcal{F}(\psi_{j_0,k}^{\text{per}})(\ell) \right|^2 \\ &= \int_0^1 \left| \sum_{k=0}^{2^{j_0}-1} (\varepsilon_k^{(u)} - \varepsilon_k^{(v)}) \psi_{j_0,k}^{\text{per}}(t) \right|^2 dt = \sum_{k=0}^{2^{j_0}-1} (\varepsilon_k^{(u)} - \varepsilon_k^{(v)})^2 \leq 2^{j_0}. \end{aligned} \quad (25)$$

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

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

Hence

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

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 $\widetilde{f^{(d)}}$ of $f^{(d)}$, we have

$$\begin{aligned} \delta_{j_0}^{-2} \sup_{f^{(d)} \in \mathbf{B}_{p,r}^s(M)} \mathbb{E} \left(\int_0^1 \left(\widetilde{f^{(d)}}(t) - f^{(d)}(t) \right)^2 dt \right) &\geq c \exp \left((\alpha/2) 2^{2j_0} - C c_0^2 2^{2j_0} \right) \\ &\geq c, \end{aligned}$$

where

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

This completes the proof of Theorem 4.2. \square

Acknowledgements The authors are very grateful to the Associate Editor and the anonymous referees for their careful reading and valuable comments that have led to an improvement of the presentation. This work is supported by ANR grant NatImages, ANR-08-EMER-009.

References

- [1] A. Al-Mbaideen and M. Benaissa. Frequency self deconvolution in the quantitative analysis of near infrared spectra. *Analytica Chimica Acta*, 705(1-2):135–148, 2011.
- [2] T. Cai. Adaptive wavelet estimation: A block thresholding and oracle inequality approach. *The Annals of Statistics*, 27:898–924, 1999.
- [3] 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.
- [4] L. Cavalier. Penalized blockwise stein’s method, monotone oracles and sharp adaptive estimation. *Math. Methods of Stat.*, 10:247–282, 2001.
- [5] L. Cavalier. Nonparametric statistical inverse problems. *Inverse Problems*, 24, 2008.
- [6] 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.
- [7] L. Cavalier and A. B. Tsybakov. Sharp adaptation for inverse problems with random noise. *Probability Theory and Related Fields*, 123:323–354, 2002.
- [8] 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.
- [9] 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.
- [10] C. Chesneau. Wavelet estimation via block thresholding: A minimax study under the L^p risk. *Statistica Sinica*, 18(3):1007–1024, 2008.
- [11] 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.
- [12] C. Chesneau, M. J. Fadili, and J.-L. Starck. Stein block thresholding for image denoising. *Applied and Computational Harmonic Analysis*, 28(1):67–88, 2010.
- [13] C. Cirelson, I. Ibragimov, and V. Sudakov. *Norm of Gaussian sample functions*. Springer Verlag, 1976.

- [14] D. De Canditiis and M. Pensky. 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 S. Smale 6(1):3–58, 2006.
- [16] D.L. Donoho. Nonlinear solution of inverse problems by wavelet-vaguelette decomposition. *Applied and Computational Harmonic Analysis*, 2:101–126, 1995.
- [17] J. Fan and J.Y. Koo. Wavelet deconvolution. *IEEE transactions on information theory*, 48:734–747, 2002.
- [18] H.Y. Gao. Wavelet shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7:469–488, 1998.
- [19] I. Johnstone, G. Kerkyacharian, D. Picard, and M. Raimondo. Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society, Serie B*, 66(3):547–573, 2004.
- [20] J.S. Marron, S. Adak, I.M. Johnstone, M.H. Neumann, and P. Patil. Exact risk analysis of wavelet regression. *J. Comput. Graph. Statist.*, 7:278–309, 1998.
- [21] Y. Meyer. *Wavelets and Operators*. Cambridge University Press, Cambridge, 1992.
- [22] M. Pensky and T. Sapatinas. Functional deconvolution in a periodic setting: Uniform case. *The Annals of Statistics*, 37(1):73–104, 2009.
- [23] 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.
- [24] M. Pensky and T. Sapatinas. Multichannel boxcar deconvolution with growing number of channels. *Electronic Journal of Statistics*, 5:53–82, 2011.
- [25] M. Pensky and B. Vidakovic. Adaptive wavelet estimator for nonparametric density deconvolution. *The Annals of Statistics*, 27:2033–2053, 1999.
- [26] 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.
- [27] A. B. Tsybakov. *Introduction à l'estimation non paramétrique*. Springer Verlag, Berlin, 2004.