Electronic Journal of Statistics
Vol. 7 (2013) 1686–1715
ISSN: 1935-7524
DOI: 10.1214/13-EJS820

Anisotropic de-noising in functional deconvolution model with dimension-free convergence rates

Rida Benhaddou, Marianna Pensky

Department of Mathematics, University of Central Florida
e-mail: ridab2009@knights.ucf.edu; marianna.pensky@ucf.edu

and

Dominique Picard

University Paris-Diderot, CNRS, LPMA
e-mail: picard@math.univ-paris-diderot.fr

Abstract: In the present paper we consider the problem of estimating a periodic $(r+1)$-dimensional function f based on observations from its noisy convolution. We construct a wavelet estimator of f, derive minimax lower bounds for the L^2-risk when f belongs to a Besov ball of mixed smoothness and demonstrate that the wavelet estimator is adaptive and asymptotically near-optimal within a logarithmic factor, in a wide range of Besov balls. We prove in particular that choosing this type of mixed smoothness leads to rates of convergence which are free of the "curse of dimensionality" and, hence, are higher than usual convergence rates when r is large.

The problem studied in the paper is motivated by seismic inversion which can be reduced to solution of noisy two-dimensional convolution equations that allow to draw inference on underground layer structures along the chosen profiles. The common practice in seismology is to recover layer structures separately for each profile and then to combine the derived estimates into a two-dimensional function. By studying the two-dimensional version of the model, we demonstrate that this strategy usually leads to estimators which are less accurate than the ones obtained as two-dimensional functional deconvolutions. Indeed, we show that unless the function f is very smooth in the direction of the profiles, very spatially inhomogeneous along the other direction and the number of profiles is very limited, the functional deconvolution solution has a much better precision compared to a combination of M solutions of separate convolution equations. A limited simulation study in the case of $r = 1$ confirms theoretical claims of the paper.

MSC 2010 subject classifications: Primary 62G05; secondary 62G08, 62P35.
Keywords and phrases: Functional deconvolution, minimax convergence rate, hyperbolic wavelets, seismic inversion.

Received January 2013.

*This is an original paper.

†Marianna Pensky was supported in part by National Science Foundation (NSF), grant DMS-1106564.

Contents

1. Introduction

Consider the problem of estimating a periodic $(r+1)$-dimensional function $f(\mathbf{u},x)$ with $\mathbf{u}=(u_1,\dots,u_r)\in[0,1]^r$ $x\in[0,1]$, based on observations from the following noisy convolution

$$y(\mathbf{u},t)=\int_0^1 g(\mathbf{u},t-x)f(\mathbf{u},x)dx+\varepsilon z(\mathbf{u},t),\quad \mathbf{u}\in[0,1]^r, t\in[0,1]. \tag{1.1}$$

Here, ε is a positive small parameter such that asymptotically $\varepsilon\to 0$, Function $g(.,.)$ in (1.1) is assumed to be known and $z(\mathbf{u},t)$ is an $r+1$-dimensional Gaussian white noise, i.e., a generalized $r+1$-dimensional Gaussian field with covariance function

$$\mathbf{E}[z(\mathbf{u}_1,t_1)z(\mathbf{u}_2,t_1)] \quad = \quad \delta(t_1-t_2)\prod_{l=1}^{r}\delta(u_{1l}-u_{2l}),$$

where $\delta(\cdot)$ denotes the Dirac δ-function and $\mathbf{u}_{il}=(u_{i1},\dots,u_{ir})\in[0,1]^r$, $i=1,2$.

Denote

$$h(\mathbf{u},t)=\int_0^1 g(\mathbf{u},t-x)f(\mathbf{u},x)dx.$$

Then, equation (1.1) can be rewritten as

$$y(\mathbf{u},t)=h(\mathbf{u},t)+\varepsilon z(\mathbf{u},t) \tag{1.2}$$

In order to simplify the narrative, we start with the two dimensional version of equation (1.1)

$$y(u,t) = \int_0^1 g(u,t-x)f(u,x)dx + \varepsilon z(u,t), \quad u,t \in [0,1]. \tag{1.3}$$

The sampling version of problem (1.3) appears as

$$y(u_l,t_i) = \int_0^1 g(u_l,t_i-x)f(u_l,x)dx + \sigma z_{li}, \quad l = 1,\dots,M,\ i = 1,\dots,N, \tag{1.4}$$

where σ is a positive constant independent of N and M, $u_l = l/M$, $t_i = i/N$ and z_{li} are i.i.d normal variables with $\mathbf{E}(z_{li}) = 0$, and $\mathbf{E}(z_{l_1 i_1} z_{l_2 i_2}) = \delta(l_1 - l_2)\delta(i_1 - i_2)$.

Equation (1.4) seems to be equivalent to M separate convolution equations

$$y_l(t_i) = \int_0^1 f_l(x)g_l(t_i-x)dx + \sigma z_{li}, \quad l = 1,\dots,M,\ i = 1,\dots,N, \tag{1.5}$$

with $y_l(t_i) = y(u_l,t_i)$, $f_l(x) = f(u_l,x)$ and $g_l(t_i - x) = g(u_l,t_i - x)$. This is, however, not true since the solution of equation (1.4) is a **two-dimensional function** while solutions of equations (1.5) are M unrelated functions $f_i(t)$. In this sense, problem (1.3) and its sampling equivalent (1.4) are functional deconvolution problems.

Functional deconvolution problems have been introduced in Pensky and Sapatinas (2009) [19] and further developed in Pensky and Sapatinas (2010, 2011) [20, 21]. However, [19, 20, 21] considered a different version of the problem where $f(u,t)$ was a function of one variable, i.e. $f(u,t) \equiv f(t)$. Their interpretation of functional deconvolution problem was motivated by solution of inverse problems in mathematical physics and multichannel deconvolution in engineering practices. Functional deconvolution problem of types (1.3) and (1.4) are motivated by experiments where one needs to recover a two-dimensional function using observations of its convolutions along profiles $u = u_i$. This situation occurs, for example, in geophysical explorations, in particular, the ones which rely on inversions of seismic signals (see, e.g., monographs of Robinson *et al.* (1986) [24] and Robinson (1999) [23] and, e.g., papers of Wason *et al.* (1984) [27], Berkhout (1986) [2] and Heimer and Cohen (2008) [8]).

In seismic exploration, a short duration seismic pulse is transmitted from the surface, reflected from boundaries between underground layers, and received by an array of sensors on the Earth surface. The signals are transmitted along straight lines called profiles. The received signals, called seismic traces, are analyzed to extract information about the underground structure of the layers along the profile. Subsequently, these traces can be modeled under simplifying assumptions as noisy outcomes of convolutions between reflectivity sequences which describe configuration of the layers and the short wave like function (called wavelet in geophysics) which corresponds to convolution kernel. The objective of seismic deconvolution is to estimate the reflectivity sequences from the measured traces. In the simple case of one layer and a single profile, the boundary

will be described by an univariate function which is the solution of the convolution equation. The next step is usually to combine the recovered functions which are defined on the set of parallel planes passing through the profiles into a multivariate function which provides the exhaustive picture of the structure of the underground layers. This is usually accomplished by interpolation techniques. However, since the layers are intrinsically anisotropic (may have different structures in various directions) and spatially inhomogeneous (may experience, for example, sharp breaks), the former approach ignores the anisotropic and spatially inhomogeneous nature of the two-dimensional function describing the layer and loses precision by analyzing each profile separately.

The paper carries out the following program:

i) Construction of a feasible procedure $\widehat{f}(\mathbf{u},t)$ for estimating the $(r+1)$-dimensional function $f(\mathbf{u},t)$ which achieves optimal rates of convergence (up to inessential logarithmic terms). We require $\widehat{f}(\mathbf{u},t)$ to be adaptive with respect to smoothness constraints on f. In this sense, the paper is related to a multitude of papers which offered wavelet solutions to deconvolution problems (see, e.g., Donoho (1995) [5], Abramovich and Silverman (1998) [1], Pensky and Vidakovic (1999) [22], Walter and Shen (1999) [26], Fan and Koo (2002) [7], Kalifa and Mallat (2003) [13], Johnstone, Kerkyacharian, Picard and Raimondo (2004) [11], Donoho and Raimondo (2004) [6], Johnstone and Raimondo (2004) [12], Neelamani, Choi and Baraniuk (2004)[17] and Kerkyacharian, Picard and Raimondo (2007) [16]).

ii) Identification of the best achievable accuracy under smoothness constraints on f. We focus here on obtaining fast rates of convergence. In this context, we prove that considering multivariate functions with 'mixed' smoothness and hyperbolic wavelet bases allows to obtain rates which are free of dimension and, as a consequence, faster than the usual ones. In particular, the present paper is related to anisotropic de-noising explored by, e.g., Kerkyacharian, Lepski and Picard (2001, 2008) [14, 15]. We compare our functional classes as well as our rates with the results obtained there.

iii) Comparison of the two-dimensional version of the functional deconvolution procedure studied in the present paper to the separate solutions of convolution equations. We show especially that the former approach delivers estimators with higher precision. For this purpose, in Section 5, we consider a discrete version of functional deconvolution problem (1.4) (rather than the continuous equation (1.3)) and compare its solution with solutions of M separate convolution equations (1.5). We show that, unless the function f is very smooth in the direction of the profiles, very spatially inhomogeneous along the other direction and the number of profiles is very limited, functional deconvolution solution has a better precision than the combination of M solutions of separate convolution equations.

The rest of the paper is organized as follows. In order to make the paper more readable and due to the application to seismic inversion, we start, in Section 2, with the two-dimensional version of the functional deconvolution problem

(1.3), describe the construction of a two-dimensional wavelet estimator of $f(u,t)$ given by equation (1.3). In Section 3, we give a brief introduction on spaces of anisotropic smoothness. After that, we derive minimax lower bounds for the L^2-risk, based on observations from (1.3), under the condition that $f(u,t)$ belongs to a Besov ball of mixed regularity and $g(u,x)$ has certain smoothness properties. In Section 4, we prove that the hyperbolic wavelet estimator derived in Section 2 is adaptive and asymptotically near-optimal within a logarithmic factor (in the minimax sense) in a wide range of Besov balls. Section 5 is devoted to the discrete version of the problem (1.4) and comparison of functional deconvolution solution with the collection of individual deconvolution equations. Section 6 extends the results to the $(r+1)$-dimensional version of the problem (1.1). Section 7 contains a limited simulation study which supports theoretical claims of the paper. We conclude the paper by discussion of the results in Section 8. Finally, Section 9 contains the proofs of the theoretical results obtained in the earlier sections.

2. Estimation Algorithm

In what follows, $\langle\cdot,\cdot\rangle$ denotes the inner product in the Hilbert space $L^2([0,1])$ (the space of squared-integrable functions defined on the unit interval $[0,1]$), i.e., $\langle f,g\rangle = \int_0^1 f(t)\overline{g(t)}dt$ for $f,g \in L^2([0,1])$. We also denote the complex conjugate of a by $\bar{a}$. Let $e_m(t) = e^{i2\pi mt}$ be a Fourier basis on the interval $[0,1]$. Let $h_m(u) = \langle e_m, h(u,\cdot)\rangle$, $y_m(u) = \langle e_m, y(u,\cdot)\rangle$, $z_m(u) = \langle e_m, z(u,\cdot)\rangle$, $g_m(u) = \langle e_m, g(u,\cdot)\rangle$ and $f_m(u) = \langle e_m, f(u,\cdot)\rangle$ be functional Fourier coefficients of functions h, y, z, g and f respectively. Then, applying the Fourier transform to equation (1.2), one obtains for any $u \in [0,1]$

$$y_m(u) = g_m(u)f_m(u) + \varepsilon z_m(u)$$

and

$$h_m(u) = g_m(u)f_m(u). \tag{2.1}$$

Consider a bounded bandwidth periodized wavelet basis (e.g., Meyer-type) $\psi_{j,k}(t)$ and finitely supported periodized s_0-regular wavelet basis (e.g., Daubechies) $\eta_{j',k'}(u)$. The choice of the Meyer wavelet basis for t is motivated by the fact that it allows easy evaluation of the the wavelet coefficients in the Fourier domain while finitely supported wavelet basis gives more flexibility in recovering a function which is spatially inhomogeneous in u. Let m_0 and m_0' be the lowest resolution levels for the two bases and denote the scaling functions for the bounded bandwidth wavelet by $\psi_{m_0-1,k}(t)$ and the scaling functions for the finitely supported wavelet by $\eta_{m_0'-1,k'}(u)$. Then, $f(u,x)$ can be expanded into wavelet series as

$$f(u,x) = \sum_{j=m_0-1}^{\infty}\sum_{j'=m_0'-1}^{\infty}\sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^{j'}-1}\beta_{j,k,j',k'}\psi_{j,k}(x)\eta_{j',k'}(u). \tag{2.2}$$

Denote $\beta_{j,k}(u) = \langle f(u,\cdot), \psi_{j,k}(\cdot)\rangle$, then, $\beta_{j,k,j',k'} = \langle \beta_{j,k}(\cdot), \eta_{j',k'}(\cdot)\rangle$. If $\psi_{j,k,m} = \langle e_m, \psi_{j,k}\rangle$ are Fourier coefficients of $\psi_{j,k}$, then, by formula (2.1) and Plancherel's formula, one has

$$\beta_{j,k}(u) = \sum_{m\in W_j} f_m(u)\overline{\psi_{j,k,m}} = \sum_{m\in W_j} \frac{h_m(u)}{g_m(u)}\,\overline{\psi_{j,k,m}}, \tag{2.3}$$

where, for any $j \geq j_0$,

$$W_j = \{m : \psi_{jkm} \neq 0\} \subseteq 2\pi/3[-2^{j+2}, -2^j] \cup [2^j, 2^{j+2}], \tag{2.4}$$

due to the fact that Meyer wavelets are band-limited (see, e.g., [11], Section 3.1). Therefore, $\beta_{j,k,j',k'}$ are of the form

$$\beta_{j,k,j',k'} = \sum_{m\in W_j} \overline{\psi_{j,k,m}} \int \frac{h_m(u)}{g_m(u)}\,\eta_{j',k'}(u)du, \tag{2.5}$$

and allow the unbiased estimator

$$\widetilde{\beta}_{j,k,j',k'} = \sum_{m\in W_j} \overline{\psi_{j,k,m}} \int \frac{y_m(u)}{g_m(u)}\,\eta_{j',k'}(u)du. \tag{2.6}$$

We now construct a hard thresholding estimator of $f(u,t)$ as

$$\widehat{f}(u,t) = \sum_{j=m_0-1}^{J-1} \sum_{j'=m_0'-1}^{J'-1} \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \widehat{\beta}_{jk,j'k'}\psi_{jk}(t)\eta_{j'k'}(u) \tag{2.7}$$

where

$$\widehat{\beta}_{j,k,j',k'} = \widetilde{\beta}_{j,k,j',k'}\mathbf{1}\left(\left|\widetilde{\beta}_{j,k,j',k'}\right| > \lambda_{j\varepsilon}\right). \tag{2.8}$$

and the values of J, J' and $\lambda_{j\varepsilon}$ will be defined later.

In what follows, we use the symbol C for a generic positive constant, independent of ε, which may take different values at different places.

3. Smoothness classes and minimax lower bounds

3.1. Smoothness classes

It is natural to consider *anisotropic* multivariate functions, i.e., functions whose smoothness is different in different directions. It is, however, much more difficult to construct appropriate spaces of mixed regularity which are meaningful for applications. One of the objectives of the present paper is to prove that classes of mixed regularity allow to obtain rates of convergence which are free of dimension. This is specifically due to the application of hyperbolic wavelets, i.e., wavelets which allow different resolution levels for each direction (see, e.g., Heping (2004) [9]).

Although comprehensive study of functional classes of mixed regularity is not the purpose of this paper, below we provide a short introduction of functional classes that we are going to consider. Due to relation of this paper to anisotropic de-noising explored by [14, 15], we also compare classes of mixed regularity used therein to the Nikolski classes considered in the papers cited above.

First, let us recall definition of the Nikolski classes $\mathcal{N}^{(s_1,\ldots,s_d)}_{(p_1,\ldots,p_d),\infty}$ (see Nikolskii (1975) [18]). In this section we consider d dimensional multivariate functions. In what follows, we set $d = r + 1$ or $d = 2$.

Let f be a measurable function defined on $\mathbb{R}^d$. For any $x, y \in \mathbb{R}^d$, we define

$$\Delta_y f(x) = f(x + y) - f(x).$$

If $l \in \mathbb{N}$ then Δ^l_y is the $l-$iterated version of the operator Δ_y. (Of course $\Delta^0_y = I_d$ where I_d is the identity operator.) Then, Nikolski classes can be defined as follows:
(recall that $\|g\|_{\mathbb{L}^p(\mathbb{R}^d,dx)} = \|g\|_p$ denotes $\left[\int_{\mathbb{R}^d} |g(x_1,\ldots,x_d)|^p dx_1 \ldots dx_d\right]^{1/p}$ for $1 \le p < \infty$, with the usual modification for $p = \infty$.)

1. Let $e_1, \ldots e_d$ be the canonical basis of $\mathbb{R}^d$. For $0 < s_i < \infty; 1 \le p_i \le \infty$, we say that f belongs to $\mathcal{N}^{s_i}_{p_i,\infty}$ if and only if there exists $l \in \mathbb{N}$, $s_i < l$, and $C(s_i, l) < \infty$, such that for any $h \in \mathbb{R}$ one has
$$\|\Delta^l_{he_i} f\|_{\mathbb{L}^{p_i}(\mathbb{R}^d,dx)} \le C(s_i, l)|h|^{s_i}.$$
2. $\mathcal{N}^{(s_1,\ldots,s_d)}_{(p_1,\ldots,p_d),\infty} = \cap_{i=1}^d \mathcal{N}^{s_i}_{p_i,\infty}$

The Nikolski classes defined above were investigated by [14, 15], they are anisotropic but do not involve mixed smoothness. Quite differently, in the present paper we shall consider classes of mixed regularity defined as follows. Denote $h = (h_1, \ldots, h_d)$, $t = (t_1, \ldots, t_d)$, $s = (s_1, \ldots, s_d)$ and let $t_i > 0$, $s_i > 0$, $i = 1, \ldots, d$. For a subset $e \subset \{1, \ldots, d\}$, we set h^e to be the vector with coordinates h_i when i belongs to e, and 0 otherwise. For a fixed integer l and $1 \le p \le \infty$, we denote

$$\Delta^{l,e}_{h^e} f(x) := \left(\prod_{j\in e} \Delta^l_{h_j e_j}\right) f(x), \quad \Omega^{l,e}(f, t^e)_p := \sup_{|h_j|\le t_j} \|\Delta^{l,e}_{h^e} f\|_p.$$

Now, in order to construct Besov classes of mixed regularity, we choose $l \ge \max_j s_j$ and define

$$B^{s_1,\ldots,s_d}_{p,\infty} = \left\{ f \in \mathbb{L}_p, \sum_{e\subset\{1,\ldots,d\}} \sup_{t>0} \sup_{j\in e} t_j^{-s_j} \Omega^{l,e}(f, t^e)_p < \infty \right\}. \tag{3.1}$$

It is proved in, e.g. [9], that under appropriate (regularity) conditions which we are omitting here, classes (3.1) can be expressed in terms of hyperbolic-wavelet coefficients, thus, providing a convenient generalization of the one-dimensional

Besov $B^s_{p,\infty}$ spaces. Furthermore, [9] considers more general Besov classes of mixed regularity $B^{s_1,\dots,s_d}_{p,q}$ that correspond to $q<\infty$ rather than $q=\infty$. In this paper, we shall assume that the hyperbolic wavelet basis satisfies required regularity conditions and follow [9] for the definition of Besov spaces of mixed regularity

$$B^{s_1,\dots,s_d}_{p,q}=\left\{f\in L^2(U):\ \sum_{j_1,\dots,j_d}2^{\aleph}\left(\sum_{k_1,\dots,k_d}|\beta_{j_1,k_1\dots,j_dk_d}|^p\right)^{\frac{q}{p}}<\infty\right\},\quad(3.2)$$

where $\aleph=\sum_{i=1}^d j_i(s_i+1/2-1/p)q$. Besov classes (3.2) compare quite easily to the Nikolski classes: it is easy to prove that the former form a subset of the latter.

3.2. Lower bounds for the risk: Two-dimensional case

Denote $U=[0,1]\times[0,1]$ and

$$s_i^*=s_i+1/2-1/p,\quad s_i'=s_i+1/2-1/p',\quad i=1,2,\quad p'=\min\{p,2\}.\quad(3.3)$$

In what follows, we assume that the function $f(u,t)$ belongs to a two-dimensional Besov ball as described above ($d=2$), so that wavelet coefficients $\beta_{j,k,j'k'}$ satisfy the following condition

$$B^{s_1,s_2}_{p,q}(A)=\left\{f\in L^2(U):\ \left(\sum_{j,j'}2^{(js_1^*+j's_2^*)q}\left(\sum_{k,k'}|\beta_{j,k,j'k'}|^p\right)^{\frac{q}{p}}\right)^{1/q}\le A\right\}.\quad(3.4)$$

Below, we construct minimax lower bounds for the L^2-risk. For this purpose, we define the minimax L^2-risk over the set V as

$$R_\varepsilon(V)=\inf_{\tilde f}\sup_{f\in V}\mathbf{E}\|\tilde f-f\|^2,$$

where $\|g\|$ is the L^2-norm of a function $g(\cdot)$ and the infimum is taken over all possible estimators $\tilde f(\cdot)$ (measurable functions taking their values in a set containing V) of $f(\cdot)$.

Assume that functional Fourier coefficients $g_m(u)$ of function $g(u,t)$ are uniformly bounded from above and below, that is, there exist positive constants ν, and C_1 and C_2, independent of m and u such that

$$C_1\,|m|^{-2\nu}\le|g_m(u)|^2\le C_2\,|m|^{-2\nu}.\quad(3.5)$$

Then, the following theorem gives the minimax lower bounds for the L^2-risk of any estimator $\tilde f_n$ of f.

Theorem 1. *Let* $\min\{s_1, s_2\} \geq \max\{1/p, 1/2\}$ *with* $1 \leq p, q \leq \infty$, *let* $A > 0$ *and* s_i', $i = 1, 2$, *be defined in (3.3). Then, under assumption (3.5), as* $\varepsilon \to 0$

$$R_\varepsilon(B_{p,q}^{s_1,s_2}(A)) \geq CA^2 \left(\frac{\varepsilon^2}{A^2}\right)^d \tag{3.6}$$

where

$$d = \min\left(\frac{2s_2}{2s_2+1}, \frac{2s_1}{2s_1+2\nu+1}, \frac{2s_1'}{2s_1'+2\nu}\right). \tag{3.7}$$

Note that the value of d in (3.7) can be re-written as

$$d = \begin{cases} \frac{2s_2}{2s_2+1}, & \text{if } \ s_1 > s_2(2\nu+1), \\ \frac{2s_1}{2s_1+2\nu+1}, & \text{if } \ (\frac{1}{p} - \frac{1}{2})(2\nu+1) \leq s_1 \leq s_2(2\nu+1), \\ \frac{2s_1'}{2s_1'+2\nu}, & \text{if } \ s_1 < (\frac{1}{p} - \frac{1}{2})(2\nu+1). \end{cases} \tag{3.8}$$

Remark 1. Note that the rates obtained here are in fact the worst rate associated to the one dimensional problem in each direction, which is not surprising since a function of only one variable and constant in the other direction, e.g., $f(u_1, u_2) = h(u_1)$ belongs to $B_{p,q}^{s_1,s_2}(A)$ as soon as h belongs to a ball of the usual one-dimensional Besov space $B_{p,q}^{s_1}$, for any s_2.

Also it is worthwhile to observe that the third rate (involving s_1') corresponds in dimension one to a "sparse" rate. Hence we observe here the so-called "elbow phenomenon" occurring only along the direction 2, because we are considering an L^2-loss and the problem has a degree of ill-posedness ν precisely in this direction.

4. Minimax upper bounds

Before deriving expressions for the minimax upper bounds for the risk, we formulate several useful lemmas which give some insight into the choice of the thresholds $\lambda_{j\varepsilon}$ and upper limits J and J' in the sums in (2.7).

Lemma 1. *Let* $\widetilde{\beta}_{j,k,j',k'}$ *be defined in (2.6). Then, under assumption (3.5), one has*

$$Var\left(\widetilde{\beta}_{j,k,j',k'}\right) \asymp \varepsilon^2 2^{2j\nu}. \tag{4.1}$$

Lemma 1 suggests that thresholds $\lambda_{j\varepsilon}$ should be chosen as

$$\lambda_{j\varepsilon} = C_\beta \sqrt{\ln(1/\varepsilon)}\, 2^{j\nu}\, \varepsilon \tag{4.2}$$

where C_β is some positive constant independent of ε. We choose J and J' as

$$2^J = (\varepsilon^2)^{-\frac{1}{2\nu+1}}, \quad 2^{J'} = (\varepsilon^2)^{-1}. \tag{4.3}$$

Note that the choices of J, J' and $\lambda_{j\varepsilon}$ are independent of the parameters, s_1, s_2, p, q and A of the Besov ball $B_{p,q}^{s_1 s_2}(A)$, and therefore our estimator (2.7) is adaptive with respect to those parameters.

The next two lemmas provide upper bounds for the wavelet coefficients and the large deviation inequalities for their estimators.

Lemma 2. *Under assumption (3.4), one has*

$$\sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^{j'}-1}|\beta_{j,k,j',k'}|^2 \le A^2 2^{-2(js_1'+j's_2')}$$

for any $j, j' \ge 0$.

Lemma 3. *Let* $\widetilde{\beta}_{j,k,j',k'}$ *and* $\lambda_{j\varepsilon}$ *be defined by formulae (2.6) and (4.2), respectively. For some positive constant* α*, define the set*

$$\Theta_{j,k,j'k',\alpha} = \{\Theta : \left|\widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'}\right| > \alpha\lambda_{j\varepsilon}\}. \tag{4.4}$$

Then, under assumption (3.5), as $\varepsilon \to 0$*, one has*

$$\Pr\left(\Theta_{j,k,j'k',\alpha}\right) = O\left(\varepsilon^{\frac{\alpha^2 C_\beta^2}{2\sigma_0^2}}\left[\ln(1/\varepsilon)\right]^{-\frac{1}{2}}\right) \tag{4.5}$$

where $\sigma_0^2 = \left(\frac{8\pi}{3}\right)^{2\nu}\frac{1}{C_1}$ *and* C_1 *is defined in (3.5).*

Using the statements above, we can derive upper bounds for the minimax risk of the estimator (2.7).

Theorem 2. *Let* $\widehat{f}(.,.)$ *be the wavelet estimator defined in (2.7), with* J *and* J' *given by (4.3). Let condition (3.5) hold and* $\min\{s_1, s_2\} \ge \max\{1/p, 1/2\}$*, with* $1 \le p, q \le \infty$*. If* C_β *in (4.2) is such that*

$$C_\beta^2 \ge 80(C_1)^{-1}(2\pi/3)^{2\nu} \tag{4.6}$$

where C_1 *is defined in (3.5), then, as* $\varepsilon \to 0$*,*

$$\sup_{f\in B_{p,q}^{s1,s2}(A))}\mathbf{E}\|\widehat{f}-f\|^2 \le CA^2\left(\frac{\varepsilon^2\ln(1/\varepsilon)}{A^2}\right)^d \ln\left(\frac{1}{\varepsilon}\right)^{d_1} \tag{4.7}$$

where d *is defined in (3.7) and*

$$d_1 = \mathbf{1}(s_1 = s_2(2\nu+1)) + \mathbf{1}(s_1 = (2\nu+1)(1/p-1/2)). \tag{4.8}$$

Remark 2. Looking at the previous results, we conclude that the rates obtained by the wavelet estimator defined in (2.7) are optimal, in the minimax sense, up to logarithmic factors. These factors are standard and coming from the thresholding procedure.

5. Sampling version of the equation and comparison with separate deconvolution recoveries

Consider now the sampling version (1.4) of the problem (1.3). In this case, the estimators of wavelet coefficients $\beta_{j,k,j',k'}$ can be constructed as

$$\widetilde{\beta}_{j,k,j',k'} = \frac{1}{M} \sum_{m \in W_j} \overline{\psi_{j,k,m}} \sum_{l=1}^{M} \frac{y_m(u_l)}{g_m(u_l)} \eta_{j',k'}(u_l). \tag{5.1}$$

In practice, $\widetilde{\beta}_{j,k,j',k'}$ are obtained simply by applying discrete wavelet transform to vectors $y_m(\cdot)/g_m(\cdot)$.

For any two sequences a_n and b_n, one says that $a_n \asymp b_n$ as $n \to \infty$ if $0 < C_1 < a_n/b_n < C_2 < \infty$ for some constants C_1 and C_2 independent of n. Recall that the continuous versions (2.6) of estimators (5.1) have $\mathrm{Var}(\widetilde{\beta}_{j,k,j',k'}) \asymp \varepsilon^2 2^{2j\nu}$ (see formula (4.1)). In order to show that equation (1.4) is the sampling version of (1.3) with $\varepsilon^2 = \sigma^2/(MN)$, one needs to show that, in the discrete case, $\mathrm{Var}(\widetilde{\beta}_{j,k,j',k'}) \asymp \sigma^2 (MN)^{-1} 2^{2j\nu}$. This indeed is accomplished by the following Lemma.

Lemma 4. *Let $\widetilde{\beta}_{j,k,j',k'}$ be defined in (5.1). Then, under assumption (3.5), as $MN \to \infty$, one has*

$$Var\left(\widetilde{\beta}_{j,k,j',k'}\right) \asymp \sigma^2 (MN)^{-1} 2^{2j\nu}. \tag{5.2}$$

Using tools developed in [19] and Lemma 4, it is easy to formulate the lower and the upper bounds for convergence rates of the estimator (2.7) with $\widehat{\beta}_{j,k,j'k'}$ given by (2.8) and the values of $\lambda_{j\varepsilon}$ and J, J' defined in (4.2) and (4.3), respectively. In particular, we obtain the following statement.

Theorem 3. *Let $\min\{s_1, s_2\} \geq \max\{1/p, 1/2\}$ with $1 \leq p, q \leq \infty$, let $A > 0$ and s_i^* be defined in (3.3). Then, under assumption (3.5), as $MN \to \infty$, for some absolute constant $C > 0$ one has*

$$R_{(MN)}(B_{p,q}^{s_1,s_2}(A)) \geq C(\sigma^2 (MN)^{-1})^d. \tag{5.3}$$

Moreover, if $\widehat{f}(.,.)$ is the wavelet estimator defined in (2.7), $\min\{s_1, s_2\} \geq \max\{1/p, 1/2\}$, and J and J' given by (4.3), then, under assumption (3.5), as $MN \to \infty$,

$$\sup_{f \in B_{p,q}^{s_1,s_2}(A))} \mathbf{E}\|\widehat{f} - f\|^2 \leq C(\sigma^2 (MN)^{-1} \ln(MN))^d \, (\ln(MN))^{d_1}. \tag{5.4}$$

where d and d_1 are defined in (3.7) and (4.8), respectively.

Now, let us compare the rates in Theorem 3 with the rates obtained by recovering each deconvolution $f_l(t) = f(u_l, t)$, $u_l = l/M$, $l = 1, \ldots, M$, separately, using equations (1.5). In order to do this, we need to determine in which space functions $f_l(x)$ are contained. The following lemma provides the necessary conclusion.

Lemma 5. *Let $f \in B^{s_1,s_2}_{p,q}(A)$ with $s_1 \geq \max\{1/p, 1/2\}$, $s_2 > \max\{1/p, 1/2\}$ and $1 \leq p, q \leq \infty$. Then, for any $l = 1, \ldots, M$, we have*

$$f_l(t) = f(u_l, t) \in B^{s_1}_{p,q}(\tilde{A}).$$

Using Lemma 5 and standard arguments (see, e.g., [11]), we obtain for each f_l

$$\sup_{f_l \in B^{s_1}_{p,q}(\tilde{A})} \mathbf{E}\|\tilde{f}_l - f_l\|^2 \asymp \begin{cases} CN^{-\frac{2s_1}{2s_1+2\nu+1}}, & \text{if } \ s_1 \geq (\frac{1}{p} - \frac{1}{2})(2\nu+1), \\ CN^{-\frac{2s_1'}{2s_1'+2\nu}}, & \text{if } \ s_1 < (\frac{1}{p} - \frac{1}{2})(2\nu+1). \end{cases}$$

Now, consider estimator $\tilde{f}$ of f with $\tilde{f}(u_l, t_i) = f_l(t_i)$. If $f_u = \partial f/\partial u$ and $f_{uu} = \partial^2 f/\partial u^2$ exist and uniformly bounded for $u \in [0,1]$, then rectangle method for numerical integration yields

$$\mathbf{E}\|\tilde{f} - f\|^2 = M^{-1} \sum_{l=1}^{M} \mathbf{E}\|\tilde{f}_l - f_l\|^2 + R_M,$$

where

$$R_M \leq (12M^2)^{-1} \left[\mathbf{E}\|\tilde{f}_u - f_u\|^2 + \sqrt{\mathbf{E}\|\tilde{f} - f\|^2 \ \mathbf{E}\|\tilde{f}_{uu} - f_{uu}\|^2}\right].$$

If M is large enough, then $R_M = o\left(\mathbf{E}\|\tilde{f} - f\|^2\right)$ as $M \to \infty$ and we derive

$$\mathbf{E}\|\tilde{f} - f\|^2 \asymp \begin{cases} CN^{-\frac{2s_1}{2s_1+2\nu+1}}, & \text{if } \ s_1 \geq (\frac{1}{p} - \frac{1}{2})(2\nu+1), \\ CN^{-\frac{2s_1'}{2s_1'+2\nu}}, & \text{if } \ s_1 < (\frac{1}{p} - \frac{1}{2})(2\nu+1). \end{cases} \tag{5.5}$$

By straightforward calculations, one can check that the only case when convergence rates of separate deconvolution recoveries can possibly be better than that of the simultaneous estimator is when $s_1 > s_2(2\nu+1)$. In this case, $s_1 > (\frac{1}{p} - \frac{1}{2})(2\nu+1)$, so that comparing the rates, by straightforward calculations we derive that simultaneous recovery delivers better precision than separate ones unless

$$\lim_{\substack{M\to\infty \\ N\to\infty}} MN^{-\frac{s_1 - s_2(2\nu+1)}{s_2(2s_1+2\nu+1)}} < 1, \quad s_1 > s_2(2\nu+1). \tag{5.6}$$

It is easy to see that relation (5.6) holds only if s_1 is large, s_2 is small and M is relatively small in comparison with N.

6. Extension to the (r + 1)-dimensional case

In this section, we extend the results obtained above to the $(r+1)$-dimensional version of the model (1.1). In this case, expanding both sides of equation (1.1) over Fourier basis, as before, we obtain for any $\mathbf{u} \in [0,1]^r$

$$y_m(\mathbf{u}) = g_m(\mathbf{u}) f_m(\mathbf{u}) + \varepsilon z_m(\mathbf{u}).$$

Construction of the estimator follows the path of the two-dimensional case. With $\psi_{j,k}(t)$ and $\eta_{j',k'}(u)$ defined earlier, we consider vectors $\mathbf{j}' = (j'_1, \ldots, j'_r)$, $\mathbf{k}' = (k'_1, \ldots, k'_r)$, $\mathbf{m}' = (m'_1, \ldots, m'_r)$ and $\mathbf{J}' = (J'_1, \ldots, J'_r)$, and subsets $\Upsilon(\mathbf{m}', \mathbf{J}')$ and $\mathcal{K}(\mathbf{j}')$ of the set of r-dimensional vectors with nonnegative integer components:

$$\begin{aligned} \Upsilon(\mathbf{m}', \mathbf{J}') &= \{\mathbf{j}' : m'_l \le j'_l \le J'_l,\ l = 1, \ldots, r\}, \\ \mathcal{K}(\mathbf{j}') &= \{\mathbf{k}' : 0 \le k'_l \le j'_l - 1,\ l = 1, \ldots, r\}. \end{aligned}$$

If $\boldsymbol{\infty}$ is the r-dimensional vector with all components being ∞, one can expand $f(\mathbf{u}, t)$ into wavelet series as

$$f(\mathbf{u}, t) = \sum_{j=m_0-1}^{\infty} \sum_{k=0}^{2^j-1} \sum_{\mathbf{j}' \in \Upsilon(\mathbf{m}', \boldsymbol{\infty})} \sum_{\mathbf{k}' \in \mathcal{K}(\mathbf{j}')} \beta_{j,k,\mathbf{j}',\mathbf{k}'} \psi_{jk}(t) \prod_{l=1}^{r} \eta_{j'_l,k'_l}(u_l), \quad (6.1)$$

where coefficients $\beta_{j,k,\mathbf{j}',\mathbf{k}'}$ are of the form

$$\beta_{j,k,\mathbf{j}',\mathbf{k}'} = \sum_{m \in W_j} \overline{\psi_{j,k,m}} \int_{[0,1]^d} \frac{h_m(\mathbf{u})}{g_m(\mathbf{u})} \prod_{l=1}^{r} [\eta_{j'_l,k'_l}(u_l)]\, d\mathbf{u}, \quad (6.2)$$

the set W_j is defined by formula (2.4) and $h_m(\mathbf{u}) = \langle (f * g)(\cdot, \mathbf{u}), e_m(\cdot) \rangle$. Similarly to the two-dimensional case, we estimate $f(\mathbf{u}, t)$ by

$$\widehat{f}(\mathbf{u}, t) = \sum_{j=m_0-1}^{J-1} \sum_{k=0}^{2^j-1} \sum_{\mathbf{j}' \in \Upsilon(\mathbf{m}', \mathbf{J}')} \sum_{\mathbf{k}' \in \mathcal{K}(\mathbf{j}')} \widehat{\beta}_{j,k,\mathbf{j}',\mathbf{k}'}\, \psi_{jk}(t) \prod_{l=1}^{r} \eta_{j'_l,k'_l}(u_l) \quad (6.3)$$

with

$$\widehat{\beta}_{j,k,\mathbf{j}',\mathbf{k}'} = \widetilde{\beta}_{j,k,\mathbf{j}',\mathbf{k}'} \mathbf{1}\left(\left| \widetilde{\beta}_{j,k,\mathbf{j}',\mathbf{k}'} \right| > \lambda_{j,\varepsilon} \right). \quad (6.4)$$

Here

$$\widetilde{\beta}_{j,k,\mathbf{j}',\mathbf{k}'} = \sum_{m \in W_j} \overline{\psi_{j,k,m}} \int \frac{y_m(\mathbf{u})}{g_m(\mathbf{u})} \prod_{l=1}^{r} [\eta_{j'_l,k'_l}(u_l)] d\mathbf{u} \quad (6.5)$$

are the unbiased estimators of $\beta_{j,k,\mathbf{j}',\mathbf{k}'}$, J is defined in (4.3), J'_l are such that $2^{J'_l} = \varepsilon^{-2}$, $l = 1, \ldots, r$, and $\lambda_{j,\varepsilon}$ is given by formula (4.2).

Assume, as before, that functional Fourier coefficients $g_m(\mathbf{u})$ of function $g(\mathbf{u}, t)$ are uniformly bounded from above and below

$$C_1 |m|^{-2\nu} \le |g_m(\mathbf{u})|^2 \le C_2 |m|^{-2\nu} \quad (6.6)$$

and that function $f(\mathbf{u}, t)$ belongs to an $(r+1)$-dimensional Besov ball. As described in section 3.1 to define these Besov balls, we introduce the vector $\mathbf{s_2} = (s_{21}, \ldots, s_{2r})$ and denote by $\mathbf{s'_2}$ and $\mathbf{s^*_2}$ vectors with components $s'_{2l} = s_{2l} + 1/2 - 1/p'$ and $s^*_{2l} = s_{2l} + 1/2 - 1/p$, $l = 1, \ldots, r$, respectively, where $p' = \min\{p, 2\}$. If $s_0 \ge \max_l s_{2l}$, then the $(r+1)$-dimensional Besov ball

of radius A is characterized by its wavelet coefficients $\beta_{j,k,\mathbf{j}',\mathbf{k}'}$ as follows (see, e.g. [9])

$$B_{p,q}^{s_1,\mathbf{s_2}}(A)=\left\{f\in L^2([0,1]^{r+1}):\left(\sum_{j,\mathbf{j}'}2^{[js_1^*+\mathbf{j}'^T\mathbf{s}_\mathbf{2}^*]q}\left(\sum_{k,\mathbf{k}'}|\beta_{j,k,\mathbf{j}',\mathbf{k}'}|^p\right)^{\frac{q}{p}}\right)^{1/q}\leq A\right\}. \tag{6.7}$$

It is easy to show that, with the above assumptions, similarly to the two-dimensional case, as $\varepsilon\to 0$, one has

$$\mathrm{Var}\left(\widetilde{\beta}_{j,k,\mathbf{j}',\mathbf{k}'}\right)\asymp\varepsilon^2 2^{2j\nu},\quad \sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^{j'}-1}|\beta_{j,k,\mathbf{j}',\mathbf{k}'}|^2\leq A^2 2^{-2(js_1'+\mathbf{j}'^T\mathbf{s}_\mathbf{2}^*)}, \tag{6.8}$$

$$\Pr\left(\left|\widetilde{\beta}_{j,k,\mathbf{j}',\mathbf{k}'}-\beta_{j,k,\mathbf{j}',\mathbf{k}'}\right|>\alpha\lambda_{j\varepsilon}\right)=O\left(\varepsilon^{\frac{\alpha^2C_\beta^2}{2\sigma_0^2}}\left[\ln(1/\varepsilon)\right]^{-\frac{1}{2}}\right). \tag{6.9}$$

The upper and the lower bounds for the risk are expressed via

$$s_{2,0}=\min_{l=1,\dots,r}s_{2,l}=s_{2,l_0}, \tag{6.10}$$

where $l_0=\arg\min s_{2,l}$. In particular, the following statements hold.

Theorem 4. *Let* $\min\{s_1,s_{2,l_0}\}\geq\max\{1/p,1/2\}$ *with* $1\leq p,q\leq\infty$. *Then, under assumption (6.6), as* $\varepsilon\to 0$,

$$R_\varepsilon(B_{p,q}^{s_1,\mathbf{s_2}}(A))\geq CA^2\left(\frac{\varepsilon^2}{A^2}\right)^D \tag{6.11}$$

where

$$D=\min\left(\frac{2s_{2,0}}{2s_{2,0}+1},\frac{2s_1}{2s_1+2\nu+1},\frac{2s_1'}{2s_1'+2\nu}\right). \tag{6.12}$$

or,

$$D=\begin{cases}\frac{2s_{2,0}}{2s_{2,0}+1}, & \text{if } s_1>s_{2,0}(2\nu+1),\\ \frac{2s_1}{2s_1+2\nu+1}, & \text{if } (\frac{1}{p}-\frac{1}{2})(2\nu+1)\leq s_1\leq s_{2,0}(2\nu+1),\\ \frac{2s_1'}{2s_1'+2\nu}, & \text{if } s_1<(\frac{1}{p}-\frac{1}{2})(2\nu+1).\end{cases} \tag{6.13}$$

Theorem 5. *Let* $\widehat{f}(.,.)$ *be the wavelet estimator defined in (6.3), with* J *defined in (4.3),* J_l' *such that* $2^{J_l'}=(\varepsilon^2)^{-1}$, $l=1,\dots,r$, *and* $\lambda_{j,\varepsilon}$ *given by formula (4.2). Let condition (3.5) hold and* $\min\{s_1,s_{2,0}\}\geq\max\{1/p,1/2\}$, *with* $1\leq p,q\leq\infty$. *If* C_β *in (4.2) satisfies condition (4.6), then, as* $\varepsilon\to 0$,

$$\sup_{f\in B_{p,q}^{s_1,\mathbf{s_2}}(A))}\mathbf{E}\|\widehat{f}-f\|^2\leq CA^2\ \left(A^{-2}\,\varepsilon^2\ln(1/\varepsilon)\right)^D\ \ln\left(1/\varepsilon\right)^{D_1} \tag{6.14}$$

where D is defined in (6.12) and

$$D_1 = \mathbf{1}(s_1 = s_{2,0}(2\nu+1)) + \mathbf{1}(s_1 = (2\nu+1)(1/p - 1/2)) + \sum_{l \neq l_0} \mathbf{1}(s_{2,l} = s_{2,0}). \tag{6.15}$$

Remark 3. Observe that convergence rates in Theorems 4 and 5 depend on s_1, p, ν and $\min_l s_{2l}$ but not on the dimension r.

It could be also natural to ask what would the corresponding results be if s_1 itself was multidimensional, that is, if one considers the case of convolution in more than one direction where

$$h(\mathbf{u}, \mathbf{t}) = \int_{[0;1]^d} g(\mathbf{u}, \mathbf{t} - \mathbf{x}) f(\mathbf{u}, x) d\mathbf{x}, \quad \mathbf{t} \in [0;1]^d; \ \mathbf{u} \in [0;1]^r.$$

Although this is beyond the scope of this paper, let us just mention that, as soon as one establishes upper bounds for the variances of the wavelet coefficients like (6.8) as well as concentration inequalities for the wavelet coefficients estimators like in (6.9), one expects to obtain convergence rates similar to Theorems 4 and 5 with s_1 replaced with $\min_k s_{1k}$.

7. Simulations

In order to investigate finite-sample performance of our estimator, we carried out a limited simulation study. We used WaveLab package for Matlab and carried out simulations using degree 3 Meyer wavelet and degree 6 Daubechies wavelets. We generated data using equation (1.4) with kernel $q(u,t) = 0.5 \exp(-|t|\,(1 + (u - 0.5)^2))$, various functions $f(u,t)$ and various values of M, N and σ. In particular, we used $N = 512$, $M = 128$ or $M = 256$, $\sigma = 0.5$ or $\sigma = 1.0$ and $f(u,t) = f_1(u) f_2(t)$ where $f_1(u)$ and $f_2(t)$ are standard test functions routinely used in testing signal processing techniques (see, e.g., introduced by Donoho & Johnstone (1994) [4]). In particularly, we utilize functions `blip`, `bumps`, and `quadratic` with `quadratic` just being a quadratic function $(y - 0.5)^2$ scaled to have a unit norm. Note that, though $f(u,t)$ is a product of two dimensional functions, the method does not "know" this and, therefore, cannot take advantage of this information.

Graphs of all test functions are presented in Figure 1.

Table 1 contains simulations results. We generated data and constructed functional deconvolution estimator (2.7) and also M Fourier-wavelet deconvolution estimators of [11]). We evaluated mean integrated square error (MISE) $\mathbf{E}\|\hat{f} - f\|^2$ of the functional deconvolution estimator and the average MISE of M Fourier-wavelet deconvolution estimators. Table 1 reports the averages of those errors over 100 simulation runs together with their standard deviations (in the parentheses).

Simulation results confirm that, as M grows, functional deconvolution becomes more advantageous than M separate deconvolutions. Indeed, while the error of a functional deconvolution estimator declines as M grows, the average error of M deconvolution estimators remains the same.

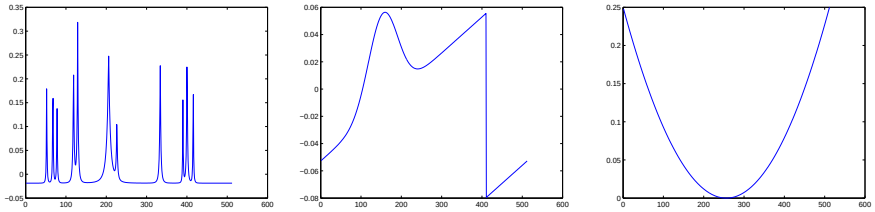

FIG 1. *Test functions:* `bumps` *(left),* `blip` *(middle),* `quadratic` *(right)*

TABLE 1

MISE averaged over 100 *runs. Third and fifth columns: average MISE of the functional deconvolution estimator. Fourth and sixth columns: average MISE of separate estimators for every* u*. Standard deviations of the errors are listed in the parentheses*

FUNCTIONAL DECONVOLUTION AND M SEPARATE DECONVOLUTIONS

M	σ	MISE (functional)	MISE (separate)	MISE (functional)	MISE (separate)
$N = 512$		$f_1 =$ `Quadratic`, $f_2 =$ `Blip`		$f_1 =$ `Quadratic`, $f_2 =$ `Bumps`	
128	0.5	0.0535 (0.00148)	0.0450 (0.00197)	0.0534 (0.00123)	0.0455 (0.00175)
128	1.0	0.213 (0.00614)	0.181 (0.00816)	0.212 (0.00589)	0.179 (0.00757)
256	0.5	0.0363 (0.00105)	0.0452 (0.00148)	0.0363 (0.000801)	0.0451 (0.00133)
256	1.0	0.145 (0.00331)	0.181 (0.00454)	0.145 (0.00343)	0.180 (0.00458)
$N = 512$		$f_1 =$ `Blip`, $f_2 =$ `Blip`		$f_1 =$ `Blip`, $f_2 =$ `Bumps`	
128	0.5	0.0539 (0.00160)	0.0453 (0.00190)	0.0531 (0.00149)	0.0447 (0.00208)
128	1.0	0.214 (0.00695)	0.180 (0.00756)	0.214 (0.00661)	0.180 (0.00836)
256	0.5	0.0364 (0.000887)	0.0452 (0.00120)	0.0364 (0.00107)	0.0452 (0.00149)
256	1.0	0.145 (0.00381)	0.180 (0.00572)	0.145 (0.00420)	0.180 (0.00591)
$N = 512$		$f_1 =$ `Bumps`, $f_2 =$ `Blip`		$f_1 =$ `Bumps`, $f_2 =$ `Bumps`	
128	0.5	0.0535 (0.00145)	0.0452 (0.00144)	0.0537 (0.00145)	0.0454 (0.00197)
128	1.0	0.213 (0.00551)	0.179 (0.00727)	0.214 (0.00683)	0.181 (0.00751)
256	0.5	0.0363 (0.000925)	0.0452 (0.00135)	0.0364 (0.00101)	0.0451 (0.00144)
256	1.0	0.144 (0.00366)	0.180 (0.00467)	0.146 (0.00355)	0.181 (0.00479)

8. Discussion

i) In the present paper, we constructed functional deconvolution estimators based on the hyperbolic wavelet thresholding procedure. We derived the lower and the upper bounds for the minimax convergence rates which confirm that estimators derived in the paper are adaptive and asymptotically near-optimal, within a logarithmic factor, in a wide range of Besov balls of mixed regularity.

ii) Although results of [14, 15] have been obtained in a slightly different framework (no convolution), they can nevertheless be compared with the results presented above. Set $\nu = 0$ to account for the absence of convolution, $p_i = p$ and $d = r + 1$. Then, convergence rates in the latter can be identified as rates of a one-dimensional setting with a regularity parameter which is equal to the harmonic mean

$$\bar{s} = \left(\frac{1}{s_1} + \cdots + \frac{1}{s_d} \right)^{-1} < \min_{i=1,\ldots,d} s_i.$$

In our case, the rates can also be identified as the rates in the one-dimensional setting with a regularity parameter $\min_i s_i$ which is always larger than $\bar{s}$. Moreover, if $s_i = s$, one obtains $\bar{s} = sd > s = \min s_i$, showing that estimators of [14, 15] in the Nikolski spaces are affected by "the curse of dimensionality" while the estimators in the anisotropic Besov spaces of mixed regularity considered in this paper are free of dimension and, therefore, have higher convergence rates.

iii) The problem studied in the paper is related to seismic inversion which can be reduced to solution of noisy convolution equations which deliver underground layer structures along the chosen profiles. The common practice in seismology, however, is to recover layer structures separately for each profile and then to combine them together. Usually, it is, however, not the best strategy and leads to estimators which are inferior to the ones obtained as two-dimensional functional deconvolutions. Indeed, as it is shown above, unless function f is very smooth in the direction of the profiles, very spatially inhomogeneous along another dimension and the number of profiles is very limited, functional deconvolution solution has precision superior to combination of M solutions of separate convolution equations. The precise condition when separate recoveries are preferable to the two-dimensional one is given by formula (5.6) which, essentially, is very reasonable. Really, if the number M of profiles is small, there is no reason to treat f as a two-dimensional function. Small value of s_2 indicates that f is very spatially inhomogeneous and, therefore, the links between its values on different profiles are very weak. Finally, if s_1 is large, deconvolutions are quite precise, so that combination of various profiles cannot improve the precision.

9. Proofs

9.1. Proof of the lower bounds for the risk

In order to prove Theorem 1, we consider two cases, the case when $f(u,t)$ is dense in both variables (the dense-dense case) and the case when $f(u,t)$ is dense in u and sparse in t. The proof is based on Lemma A.1 of Bunea, Tsybakov and Wegkamp (2007) [3] which we reformulate here for the case of squared risk.

Lemma 6 ([3], Lemma A.1)**.** *Let Ω be a set of functions of cardinality card$(\Omega) \geq 2$ such that*
(i) $\|f-g\|^2 \geq 4\delta^2, \quad$ for $f,g \in \Omega, \ \ f \neq g$,
(ii) the Kullback divergences $K(P_f, P_g)$ between the measures P_f and P_g satisfy the inequality $K(P_f,P_g) \leq \log(card(\Omega))/16, \quad$ for $\quad f,g \in \Omega$.
Then, for some absolute positive constant C, one has

$$\inf_{T_n} \sup_{f \in \ \Omega} \mathbf{E}_f \|T_n - f\|^2 \geq C\delta^2.$$

The dense-dense case. Let ω be the matrix with components $\omega_{k,k'} = \{0,1\}$, $k = 0,\dots,2^j-1$, $k' = 0,\dots,2^{j'}-1$. Denote the set of all possible values ω by Ω and let the functions $f_{j,j'}$ be of the form

$$f_{jj'}(t,u) = \gamma_{jj'} \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \omega_{k,k'} \psi_{jk}(t) \eta_{j'k'}(u). \tag{9.1}$$

Note that matrix ω has $N = 2^{j+j'}$ components, and, hence, cardinality of the set of such matrices is $\mathrm{card}(\Omega) = 2^N$. Since $f_{jj'} \in B^{s_1 s_2}_{p,q}(A)$, direct calculations show that $\gamma_{jj'} \le A 2^{-j(s_1+1/2)-j'(s_2+1/2)}$, so that we choose $\gamma_{jj'} = A 2^{-j(s_1+1/2)-j'(s_2+1/2)}$. If $\tilde f_{jj'}$ is of the form (9.1) with $\tilde\omega_{k,k'} \in \Omega$ instead of $\omega_{k,k'}$, then, the L^2-norm of the difference is of the form

$$\|\tilde f_{jj'} - f_{jj'}\|^2 \quad = \quad \gamma^2_{jj'} \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \mathbf{1}\left(\tilde\omega_{k,k'} \ne \omega_{k,k'}\right) = \gamma^2_{jj'} \rho(\tilde\omega, \omega)$$

where $\rho(\tilde\omega,\omega) = \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \mathbf{1}\left(\tilde\omega_{k,k'} \ne \omega_{k,k'}\right)$ is the Hamming distance between the binary sequences ω and $\tilde\omega$. In order to find a lower bound for the last expression, we apply the Varshamov-Gilbert lower bound (see Tsybakov (2008), page 104) which states that one can choose a subset Ω_1 of Ω, of cardinality at least $2^{N/8}$ such that $\rho(\tilde\omega,\omega) \ge N/8$ for any $\omega, \tilde\omega \in \Omega_1$. Hence, for any $\omega, \tilde\omega \in \Omega_1$ one has $\|\tilde f_{jj'} - f_{jj'}\|^2 \ge \gamma^2_{jj'} 2^{j+j'}/8$. Note that Kullback divergence can be written as

$$K(f,\tilde f) \quad = \quad (2\varepsilon^2)^{-1} \|(\tilde f - f) * g\|^2. \tag{9.2}$$

Since $|\omega_{jj'} - \tilde\omega_{jj'}| \le 1$, plugging f and $\tilde f$ into (9.2), using Plancherel's formula and recalling that $|\psi_{j,k,m}| \le 2^{-j/2}$, we derive

$$K(f,\tilde f) \quad \le \quad (2\varepsilon^2)^{-1} 2^{-j} \gamma^2_{jj'} \sum_{k=0}^{2^j-1} \sum_{k'=0}^{2^{j'}-1} \sum_{m \in W_j} \int_0^1 \eta^2_{j'k'}(u)\, g^2_m(u)\, du.$$

Using (3.5), we obtain

$$2^{-j} \sum_{m\in W_j} \int_0^1 \eta^2_{j'k'}(u) g^2_m(u) du \le C_2 2^{-j} \sum_{m \in W_j} |m|^{-2\nu} \int_0^1 \eta^2_{j'k'}(u) du \le C_3 2^{-2\nu j},$$

so that

$$K(f,\tilde f) \le C \varepsilon^{-2} \gamma^2_{jj'} 2^{j+j'} 2^{-2\nu j}. \tag{9.3}$$

Now, applying Lemma 6 with

$$\delta^2 = \gamma^2_{jj'} 2^{j+j'}/32 = A^2 2^{-2s_1 j - 2s_2 j'}/32 \tag{9.4}$$

one obtains constraint $2^{-j(2s_1+2\nu+1)-j'(2s_2+1)} \le C\varepsilon^2/A^2$ on j, j' and ε where C is an absolute constant. Denote

$$\tau_\varepsilon = \log_2(CA^2\varepsilon^{-2}). \tag{9.5}$$

Thus, we need to choose combination of j and j' which solves the following optimization problem

$$2js_1 + 2j's_2 \Rightarrow \min\ j(2s_1+2\nu+1) + j'(2s_2+1) \ge \tau_\varepsilon, \quad j, j' \ge 0. \tag{9.6}$$

It is easy to check that solution of this linear constraint optimization problem is of the form $\{j, j'\} = \left\{(2s_1+2\nu+1)^{-1}\tau_\varepsilon, 0\right\}$ if $s_2(2\nu+1) > s_1$, and $\{j, j'\} = \left\{0, (2s_2+1)^{-1}\tau_\varepsilon\right\}$ if $s_2(2\nu+1) \le s_1$. Plugging those values into (9.4), obtain

$$\delta^2 = \begin{cases} CA^2\,(\varepsilon^2/A^2)^{\frac{2s_2}{2s_2+1}}, & \text{if } \ s_1 > s_2(2\nu+1), \\ CA^2\,(\varepsilon^2/A^2)^{\frac{2s_1}{2s_1+2\nu+1}}, & \text{if } \ s_1 \le s_2(2\nu+1). \end{cases} \tag{9.7}$$

The sparse-dense case. Let ω be the vector with components $\omega_{k'} = \{0, 1\}$. Denote Ω the set of all possible ω and let the functions $f_{j,j'}$ be of the form

$$f_{jj'}(t,u) = \gamma_{jj'} \sum_{k'=0}^{2^{j'}-1} \omega_{k'}\psi_{jk}(t)\eta_{j'k'}(u) \tag{9.8}$$

Note that vector ω has $N = 2^{j'}$ components, and, hence, its cardinality is $\mathrm{card}(\Omega) = 2^N$. Since $f_{jj'} \in B^{s_1 s_2}_{p,q}(A)$, direct calculations show that $\gamma_{jj'} \le A2^{-js_1^* - j'(s_2+1/2)}$, so we choose $\gamma_{jj'} = A2^{-js_1^* - j'(s_2+1/2)}$. If $\tilde{f}_{jj'}$ is of the form (9.8) with $\tilde{\omega}_{k,k'} \in \Omega$ instead of $\omega_{k,k'}$, then, calculating the L^2 norm of the difference similarly to dense-dense case, obtain

$$\|\tilde{f}_{jj'} - f_{jj'}\|^2 \quad = \quad \gamma^2_{jj'} \sum_{k'=0}^{2^{j'}-1} \mathbf{1}\left(\tilde{\omega}_{k'} \ne \omega_{k'}\right) \ge \gamma^2_{jj'} 2^{j'}/8.$$

Similarly to dense-dense case, using formulae (3.5) and (9.2), Plancherel's formula and $|\psi_{j,k,m}| \le 2^{-j/2}$, derive

$$K(f, \tilde{f}) \le (2\varepsilon^2)^{-1}\gamma^2_{jj'} \sum_{k'=0}^{2^{j'}-1} 2^{-j} \sum_{m\in W_j} \int_0^1 \eta^2_{j'k'}(u) g^2_m(u)du \le C(2\varepsilon^2)^{-1}\gamma^2_{jj'} 2^{j'} 2^{-2\nu j}.$$

Now, applying Lemma 6 with

$$\delta^2 = \gamma^2_{jj'} 2^{j'}/32 = A^2 2^{-2s_1' j - 2s_2 j'}/32 \tag{9.9}$$

one obtains constraint $2^{-j(2s_1'+2\nu)-j'(2s_2+1)} \le C\varepsilon^2/A^2$ on j, j' and ε where C is an absolute constant. Thus, we need to choose combination of j and j' which

delivers solution to the following linear optimization problem $\min\{2js_1+2j's_2\}$ subject to constraint

$$2js_1+2j's_2 \Rightarrow \min \quad s.t. \quad j(2s_1'+2\nu)+j'(2s_2+1)\geq \tau_\varepsilon, \quad j,j'\geq 0. \tag{9.10}$$

It is easy to check that solution of this linear constraint optimization problem is of the form $\{j,j'\}=\left\{(2s_1'+2\nu)^{-1}\tau_\varepsilon,0\right\}$ if $2\nu s_2 > s_1'$, and $\{j,j'\}=\left\{0,(2s_2+1)^{-1}\tau_\varepsilon\right\}$ if $2\nu s_2 \leq s_1'$. Plugging those values into (9.9), obtain

$$\delta^2=\begin{cases} CA^2\left(\varepsilon^2/A^2\right)^{\frac{2s_2}{2s_2+1}}, & \text{if } \ 2\nu s_2\leq s_1', \\ CA^2\left(\varepsilon^2/A^2\right)^{\frac{2s_1'}{2s_1'+2\nu}}, & \text{if } \ 2\nu s_2> s_1'. \end{cases} \tag{9.11}$$

In order to complete the proof, recall expressions (3.7) and (3.8) for d.

9.2. Proofs of supplementary lemmas

Proof of Lemma 1. Let us derive an expression for the upper bound of the variance of (2.6). Subtracting (2.5) from (2.6) we obtain

$$\widetilde{\beta}_{j,k,j',k'}-\beta_{j,k,j',k'}=\varepsilon\sum_{m\in W_j}\overline{\psi_{j,k,m}}\int_0^1\frac{z_m(u)}{g_m(u)}\eta_{j',k'}(u)du. \tag{9.12}$$

Now, before we proceed to the derivation of the upper bound of the variance, let us first state a result that will be used in our calculation. Recall from stochastic calculus that for any function $F(t,u)\in L^2([0,1]\times[0,1])$, one has

$$\mathbf{E}\left[\int_0^1\int_0^1 F(t,u)dz(t,u)du\right]^2 \;=\; \int_0^1\int_0^1 F^2(t,u)dtdu. \tag{9.13}$$

Hence, recalling that $z_m(u)=\int z(u,t)e_m(t)dt$, choosing

$$F(t,u)=\sum_{m\in W_j}\overline{\psi_{j,k,m}}\,\frac{e_m(t)}{g_m(u)}\eta_{j',k'}(u),$$

squaring both sides of (9.12), taking expectation and using the relation (9.13), we obtain

$$\begin{aligned}\operatorname{Var}\left(\widetilde{\beta}_{j,k,j',k'}\right) &= \varepsilon^2\,\mathbf{E}\left|\sum_{m\in W_j}\overline{\psi_{j,k,m}}\int_0^1\int_0^1\frac{\eta_{j',k'}(u)}{g_m(u)}e_m(t)dz(u,t)du\right|^2 \\ &= \varepsilon^2\int_0^1\int_0^1\sum_m\sum_{m'}\frac{\overline{\psi_{j,k,m}}\psi_{j,k,m'}}{g_m(u)\overline{g_{m'}}(u)}\overline{e_m}(t)e_{m'}(t)|\eta_{j',k'}(u)|^2dtdu \\ &= \varepsilon^2\sum_{m\in W_j}|\psi_{j,k,m}|^2\int_0^1\frac{|\eta_{j',k'}(u)|^2}{|g_m(u)|^2}du,\end{aligned}$$

since in the double summation above, all terms involving $m \neq m'$ vanish due to $\int_0^1 e_m(t)e_{m'}(t)dt = 0$. Consequently, Taking into account (2.4), (3.5) and the fact that $|\psi_{j,k,m}| \leq 2^{-j/2}$, obtain

$$\mathrm{Var}\left(\widetilde{\beta}_{j,k,j',k'}\right) \asymp \varepsilon^2 \sum_{m\in W_j} |\psi_{j,k,m}|^2 |m|^{2\nu} \int_0^1 \left|\eta^2_{j',k'}(u)\right| du \asymp \varepsilon^2 2^{2j\nu} \tag{9.14}$$

so that (4.1) holds. □

Proof of Lemma 2. First note that, under assumption (3.4), one has

$$\sum_{k,k'} |\beta_{j,k,j',k'}|^p \leq A^p 2^{-p\left[(js_1+j's_2)+(\frac{1}{2}-\frac{1}{p})(j+j')\right]}$$

If $p \leq 2$, one has $p' = p$, $s'_i = s_i + 1/2 - 1/p$, $i = 1, 2$, and

$$\sum_{k,k'} |\beta_{j,k,j',k'}|^2 \leq \sum_{k,k'} |\beta_{j,k,j',k'}|^p \left\{\max_{k,k'} |\beta_{j,k,j',k'}|^p\right\}^{(2-p)/p} \leq A^2 2^{-2(js'_1+j's'_2)}.$$

If $p \geq 2$, then $p' = 2$, $s'_i = s_i$, $i = 1, 2$, and, applying the Cauchy-Schwarz inequality, one obtain

$$\sum_{k,k'} |\beta_{j,k,j',k'}|^2 \leq \left(\sum_{k,k'} |\beta_{j,k,j',k'}|^p\right)^{2/p} \left(\sum_{k,k'} 1\right)^{(1-2/p)} \leq A^2 2^{-2\left[(js_1+j's_2)\right]},$$

which completes the proof. □

Proof of Lemma 3. Observe that $\widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'}$ is a zero-mean Gaussian random variable with variance given by (9.14), so that

$$\mathrm{Var}\left(\widetilde{\beta}_{j,k,j',k'}\right) \leq \varepsilon^2 \left(\frac{8\pi}{3}\right)^{2\nu} \frac{2^{2\nu j}}{C_1} = \sigma_0^2 \varepsilon^2 2^{2\nu j} \tag{9.15}$$

Denoting by $\bar{\Phi}(x) = 1 - \Phi(x)$ where $\Phi(x)$ is the standard normal c.d.f. and recalling that $\bar{\Phi}(x) \leq (x\sqrt{2\pi})^{-1} \exp(-x^2/2)$ if $x > 0$, we derive

$$\begin{aligned} \Pr\left(\Omega_{jk,j'k',\alpha}\right) &= \Pr\left(|\xi_{j,k,j',k'}| > \alpha\lambda_{j\varepsilon}\right) = 2\bar{\Phi}\left(\alpha\lambda_{j\varepsilon}(\sigma_0\varepsilon 2^{\nu j})^{-1}\right) \\ &\leq 2\bar{\Phi}\left(\alpha C_\beta(\sigma_0)^{-1}\sqrt{\ln(1/\varepsilon)}\right) \leq \frac{2\sigma_0}{\alpha C_\beta\sqrt{2\pi\ln(1/\varepsilon)}}\, \varepsilon^{\frac{\alpha^2 C_\beta^2}{2\sigma_0^2}} \end{aligned}$$

which completes the proof. □

9.3. Proof of upper bounds for the risk

Proof of Theorem 2. Denote

$$\chi_{\varepsilon,A} = A^{-2}\varepsilon^2 \ln(1/\varepsilon), \tag{9.16}$$

$$2^{j_0} = (\chi_{\varepsilon,A})^{-\frac{d}{2s_1'}}, \quad 2^{j_0'} = (\chi_{\varepsilon,A})^{-\frac{d}{2s_2'}} \tag{9.17}$$

and observe that with J and J' given by (4.3), the estimation error can be decomposed into the sum of four components as follows

$$\mathbf{E}\|\widehat{f}_n - f\|^2 \leq \sum_{j,k,j',k'} \mathbf{E}\|\widehat{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'}\|^2 \leq R_1 + R_2 + R_3 + R_4 \tag{9.18}$$

where

$$\begin{aligned}
R_1 &= \sum_{k=0}^{2^{m_0}-1} \sum_{k'=0}^{2^{m_0'}-1} \mathrm{Var}(\widetilde{\beta}_{m_0,k,m_0',k'}),\\
R_2 &= \sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} \sum_{k,k'} \mathbf{E}\left[\left|\widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'}\right|^2 \mathbf{1}\left(\left|\widetilde{\beta}_{j,k,j',k'}\right| > \lambda_{j\varepsilon}\right)\right],\\
R_3 &= \sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} \sum_{k,k'} |\beta_{j,k,j',k'}|^2 \Pr\left(\left|\widetilde{\beta}_{j,k,j',k'}\right| < \lambda_{j\varepsilon}\right),\\
R_4 &= \left(\sum_{j=J}^{\infty} \sum_{j'=m_0'}^{J'-1} + \sum_{j=m_0}^{J-1} \sum_{j'=J'}^{\infty} + \sum_{j=J}^{\infty} \sum_{j'=J'}^{\infty}\right) \sum_{k,k'} |\beta_{j,k,j',k'}|^2.
\end{aligned}$$

For R_1, using (4.1), derive, as $\varepsilon \to 0$,

$$R_1 \leq C\varepsilon^2 = O\left(A^2 \chi_{\varepsilon,A}^d\right). \tag{9.19}$$

To calculate R_4, we apply Lemma 2 and use (4.3), obtaining, as $\varepsilon \to 0$,

$$\begin{aligned}
R_4 &= O\left(\left(\sum_{j\geq J} \sum_{j'\geq m_0'} + \sum_{j\geq m_0} \sum_{j'\geq J'}\right) A^2 2^{-2js_1' - 2j's_2'}\right)\\
&= O\left(A^2 2^{-2Js_1} + A^2 2^{-2J's_2}\right)\\
&= O\left(A^2(\varepsilon^2)^{\frac{2s_1'}{2\nu+1}} + A^2(\varepsilon^2)^{2s_2'}\right) = O\left(A^2\chi_{\varepsilon,A}^d\right).
\end{aligned} \tag{9.20}$$

Then, our objective is to prove that, as $\varepsilon \to 0$, one has $R_i = O\left(A^2\chi_{\varepsilon,A}^d[\ln(1/\varepsilon)]^{d_1}\right)$.

Note that each R_2 and R_3 can be partitioned into the sum of two errors as follows

$$R_2 \leq R_{21} + R_{22}, \quad R_3 \leq R_{31} + R_{32}, \tag{9.21}$$

where

$$R_{21} = \sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} \sum_{k,k'} \mathbf{E}\left[\left| \widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'} \right|^2 \mathbf{1}\left(\left| \widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'} \right| > \frac{\lambda_{j\varepsilon}}{2} \right) \right],$$

$$R_{22} = \sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} \sum_{k,k'} \mathbf{E}\left[\left| \widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'} \right|^2 \mathbf{1}\left(|\beta_{j,k,j',k'}| > \frac{1}{2}\lambda_{j\varepsilon} \right) \right],$$

$$R_{31} = \sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} \sum_{k,k'} |\beta_{j,k,j',k'}|^2 \Pr\left(\left| \widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'} \right| > \frac{\lambda_{j\varepsilon}}{2} \right),$$

$$R_{32} = \sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} \sum_{k,k'} |\beta_{j,k,j',k'}|^2 \mathbf{1}\left(|\beta_{j,k,j',k'}| \leq \frac{3\lambda_{j\varepsilon}}{2} \right).$$

Combining expressions for R_{21} and R_{31}, and applying Cauchy-Schwarz inequality and Lemma 3 with $\alpha = 1/2$, one derives

$$\begin{aligned} R_{21} + R_{31} &= O\left(\sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} 2^{j+j'} \, \varepsilon^{\frac{C_\beta^2}{16\sigma_0^2}} \, [\ln(1/\varepsilon)]^{-\frac{1}{4}} \sqrt{\varepsilon^4 2^{4j\nu+j'}} \right) \\ &= O\left(2^{J(2\nu+1)} \, 2^{3J'/2} \, (\varepsilon)^{2+\frac{C_\beta^2}{16\sigma_0^2}} \right) = O\left((\varepsilon^2)^{\frac{C_\beta^2}{32\sigma_0^2} - \frac{3}{2}} \right). \end{aligned}$$

Hence, due to condition (4.6), one has, as $\varepsilon \to 0$,

$$R_{21} + R_{31} \leq C\varepsilon^2 = O\left(A^2 \chi_{\varepsilon,A}^d \right). \tag{9.22}$$

For the sum of R_{22} and R_{32}, using (4.1) and (4.2), we obtain

$$\Delta = R_{22} + R_{32} = O\left(\sum_{j=m_0}^{J-1} \sum_{j'=m_0'}^{J'-1} \sum_{k,k'} \min\left\{ \beta_{j,k,j',k'}^2, \varepsilon^2 \ln(1/\varepsilon) \, 2^{2j\nu} \right\} \right). \tag{9.23}$$

Then, Δ can be partitioned into the sum of three components Δ_1, Δ_2 and Δ_3 according to three different sets of indices:

$$\Delta_1 = O\left(\left\{ \sum_{j=j_0+1}^{J-1} \sum_{j'=m_0'}^{J'-1} + \sum_{j=m_0}^{J-1} \sum_{j'=j_0'+1}^{J'-1} \right\} A^2 2^{-2js_1' - 2j's_2'} \right), \tag{9.24}$$

$$\Delta_2 = O\left(\sum_{j=m_0}^{j_0} \sum_{j'=m_0'}^{j_0'} \varepsilon^2 \ln(1/\varepsilon) \, 2^{j(2\nu+1)+j'} \, \mathbf{1}\left(2^{j(2\nu+1)+j'} \leq \chi_{\varepsilon,A}^{d-1} \right) \right), \tag{9.25}$$

$$\Delta_3 = O\left(\sum_{j=m_0}^{j_0} \sum_{j'=m_0'}^{j_0'} A^{p'} 2^{-p'js_1' - p'j's_2'} \left(\varepsilon^2 \ln(1/\varepsilon) 2^{2j\nu} \right)^{1-p'/2} \mathbf{1}(\Xi) \right). \tag{9.26}$$

where d is defined in (3.7) and $\Xi = \{\Xi : 2^{j(2\nu+1)+j'} > \chi_{\varepsilon,A}^{d-1}\}$. It is easy to see that for Δ_1 given in (9.24) and j_0 and j_0' given by (9.17), as $\varepsilon \to 0$, one has

$$\Delta_1 = O\left(A^2\, \chi_{\varepsilon,A}^{d}\right), \tag{9.27}$$

For Δ_2 defined in (9.25), obtain

$$\Delta_2 = O\left(\varepsilon^2 \ln(1/\varepsilon)\chi_{\varepsilon,A}^{d-1}\right) = O\left(A^2\, \chi_{\varepsilon,A}^{d}\right), \quad \varepsilon \to 0. \tag{9.28}$$

In order to construct upper bounds for Δ_3 in (9.26), we need to consider three different cases.

Case 1: $s_1 \geq s_2(2\nu+1)$. In this case, $d = 2s_2/(2s_2+1)$ and

$$\begin{aligned}
\Delta_3 &\leq CA^2(\chi_{\varepsilon,A})^{1-p'/2} \sum_{j=m_0}^{j_0} 2^{-j[p's_1' - 2\nu(1-p'/2)]} \sum_{j'=m_0'}^{j_0'} 2^{-p'j's_2'}\, \mathbf{1}\left(\Theta_1\right) \\
&\leq CA^2(\chi_{\varepsilon,A})^{(1-p'/2)+p's_2'(1-d)} \sum_{j=m_0}^{j_0} 2^{-j[p's_1' - 2\nu(1-p'/2) - p'(2\nu+1)s_2']} \\
&= CA^2(\chi_{\varepsilon,A})^{d} \sum_{j=m_0}^{j_0} 2^{-j[p's_1 - p's_2(2\nu+1)]},
\end{aligned}$$

where $\Theta_1 = \{\Theta : 2^{j'} > (\chi_{\varepsilon,A})^{d-1} 2^{-j(2\nu+1)}\}$, so that, as $\varepsilon \to 0$,

$$\Delta_3 = O\left(A^2\, \chi_{\varepsilon,A}^{d}\, [\ln(1/\varepsilon)]^{\mathbf{1}(s_1 = s_2(2\nu+1))}\right). \tag{9.29}$$

Case 2: $(\frac{1}{p} - \frac{1}{2})(2\nu+1) < s_1 < s_2(2\nu+1)$. In this case, $d = 2s_1/(2s_1+2\nu+1)$ and

$$\begin{aligned}
\Delta_3 &\leq CA^2(\chi_{\varepsilon,A})^{1-p'/2} \sum_{j=m_0}^{j_0} 2^{-j[p's_1' - 2\nu(1-p'/2)]} \sum_{j'=m_0'}^{j_0'} 2^{-p'j's_2'}\, \mathbf{1}\left(\Theta_2\right) \\
&\leq CA^2(\chi_{\varepsilon,A})^{(1-p'/2)+p'\frac{(1-d)}{1+2\nu}(s_1 - (2\nu+1)(1/p'-1/2)} \sum_{j'=m_0'}^{j_0'} 2^{-j'p'\Upsilon} \\
&\leq CA^2(\chi_{\varepsilon,A})^{d} \sum_{j'=m_0'}^{j_0'} 2^{-j'p'[s_2 - s_1/(2\nu+1)]},
\end{aligned}$$

where $\Theta_2 = \{\Theta : 2^{j} > (\chi_{\varepsilon,A})^{\frac{d-1}{2\nu+1}} 2^{-\frac{j'}{2\nu+1}}\}$ and $\Upsilon = [s_2 - s_1/(2\nu+1)]$, so that, as $\varepsilon \to 0$,

$$\Delta_3 = O\left(A^2\, \chi_{\varepsilon,A}^{d}\right). \tag{9.30}$$

Case 3: $s_1 \leq (\frac{1}{p} - \frac{1}{2})(2\nu+1)$. In this case, $d = 2s_1'/(2s_1'+2\nu)$ and $p \leq 2$. Then, since $ps_1' - 2\nu(1-p/2) = p[s_1 - (1/p - 1/2)(2\nu+1)] \leq 0$, one has

$$\begin{aligned}\Delta_3 &\leq CA^2(\chi_{\varepsilon,A})^{1-p'/2} \sum_{j=m_0}^{j_0} 2^{-j[ps_1' - 2\nu(1-p/2)]} \\ &\leq CA^2(\chi_{\varepsilon,A})^{1-p'/2}\, 2^{j_0 p[(1/p-1/2)(2\nu+1)-s_1]}\, [\ln(1/\varepsilon)]^{\mathbf{1}(s_1=(1/p-1/2)(2\nu+1))}.\end{aligned}$$

Plugging in j_0 of the form (9.17), obtain as $\varepsilon \to 0$

$$\Delta_3 = O\left(A^2\, \chi_{\varepsilon,A}^d\, [\ln(1/\varepsilon)]^{\mathbf{1}(s_1=(1/p-1/2)(2\nu+1))}\right). \tag{9.31}$$

Now, to complete the proof, combine formulae (9.18)–(9.31). □

9.4. Proofs of the statements in Section 5

Proof of Lemma 4. Subtracting $\beta_{j,k,j',k'}$ from (5.1), one obtains

$$\widetilde{\beta}_{j,k,j',k'} - \beta_{j,k,j',k'} = \frac{\sigma}{M} \sum_{m \in W_j} \overline{\psi_{j,k,m}} \sum_{l=1}^{M} \frac{z_m(u_l)}{g_m(u_l)}\, \eta_{j',k'}(u_l). \tag{9.32}$$

where $z_m(u_l) = y_m(u_l) - h_m(u_l)$. Since Fourier transform is an orthogonal transform, one has $\mathbf{E}[z_{m_1}(u_{l_1}) z_{m_2}(u_{l_2})] = 0$ if $l_1 \neq l_2$ and $\mathbf{E}[z_{m_1}(u_l) z_{m_2}(u_l)] = 0$, so that

$$\mathbf{E}[z_{m_1}(u_{l_1}) z_{m_2}(u_{l_2})] = \frac{\sigma^2}{N} \delta(m_1 - m_2)\delta(l_1 - l_2).$$

Therefore,

$$\begin{aligned}\mathrm{Var}(\widetilde{\beta}_{j,k,j',k'}) &= \frac{\sigma^2}{M^2 N} \sum_{m \in W_j} |\psi_{j,k,m}|^2 \sum_{l=1}^{M} \frac{1}{|g_m(u_l)|^2} |\eta_{j',k'}(u_l)|^2 \\ &\asymp \frac{\sigma^2 2^{2j\nu}}{MN} \sum_{m \in W_j} |\psi_{j,k,m}|^2\, \frac{1}{M} \sum_{l=1}^{M} |\eta_{j',k'}(u_l)|^2 \asymp \frac{\sigma^2 2^{2j\nu}}{MN},\end{aligned}$$

which completes the proof. □

Proof of Lemma 5. Recall that

$$f(u,t) = \sum_{j,k} \sum_{j',k'} \beta_{j,k,j',k'} \psi_{j,k}(t) \eta_{j',k'}(u) \quad \text{and} \quad f_l(t) = \sum_{j,k} b_{j,k}^{(l)} \psi_{j,k}(t) \eta_{j',k'}(u_l),$$

so that

$$b_{j,k}^{(l)} = \sum_{j'=0}^{\infty} \sum_{k' \in K_l} \beta_{j,k,j',k'} 2^{j'/2} \eta(2^{j'} u_l - k'),$$

where the set $K_l = \{k' : \ \eta(2^{j'}u_l - k') \neq 0\}$ is finite for any l due to finite support of η.

Thus, since $p \geq 1$, for any $\delta > 0$, one has

$$\sum_{k=0}^{2^j-1} |b_{j,k}^{(l)}|^p \leq C \sum_{k=0}^{2^j-1} \left[\sum_{j'=0}^{\infty} \sum_{k' \in K_l} |\beta_{j,k,j',k'}| \, 2^{j'(1+\delta)/2} \, 2^{-j'\delta/2} \right]^2$$
$$\leq C \sum_{k=0}^{2^j-1} \left[\sum_{j'=0}^{\infty} \sum_{k' \in K_l} |\beta_{j,k,j',k'}|^p \, 2^{j'(1+\delta)p/2} \right] \left[\sum_{j'=0}^{\infty} \sum_{k' \in K_l} \left(2^{-j'\delta/2} \right)^{\frac{p}{p-1}} \right]^{p-1}$$

Then, for any $q \geq 1$, one has

$$B_j = \left(\sum_{j'=0}^{\infty} 2^{j'(1+\delta)p/2} \sum_{k,k'} |\beta_{j,k,j',k'}|^p \right)^{q/p}.$$

If $q/p \geq 1$, then, using Cauchy-Schwarz inequality again, it is straightforward to verify that

$$B_j \leq \tilde{C}_\delta \sum_{j'=0}^{\infty} \left[\sum_{k,k'} |\beta_{j,k,j',k'}|^p \right]^{q/p} 2^{j'(1+2\delta)q/2}.$$

Hence,

$$\sum_{j'=0}^{\infty} 2^{js_1'q} \left(\sum_{k=0}^{2^j-1} |b_{j,k}^{(l)}|^p \right)^{q/p} \leq \tilde{C}_\delta 2^{js_1'q} \sum_{j'=0}^{\infty} 2^{j'(1+2\delta)q/2} \left[\sum_{k,k'} |\beta_{j,k,j',k'}|^p \right]^{q/p}$$
$$\leq \tilde{C}_\delta A^q = \tilde{A}^q$$

provided $s_2^* \geq (1+2\delta)/2$. Since $s_2 > \max\{1/2, 1/p\}$ implies $s_2 > 1/2$, choose $\delta = (s_2 - 1/2)/2$. If $q/p < 1$, then similar considerations yield

$$B_j \leq \tilde{C}_\delta \sum_{j'=0}^{\infty} \left[\sum_{k,k'} |\beta_{j,k,j',k'}|^p \right]^{q/p} 2^{j'(1+\delta)q/2},$$

so that the previous calculation holds with δ instead of 2δ, and the proof is complete. □

9.5. Proofs of the statements in Section 6

Proof of Theorem 4. Repeating the proof of Theorem 1 with j' and k' replaced by $\mathbf{j}'$ and $\mathbf{k}'$, respectively, and $s_2 j'$ replaced by $\mathbf{j}'^T \mathbf{s}_2'$, we again arrive at two cases. Denote the r-dimensional vector with all unit components by $\mathbf{e}$.

In the dense-dense case, we use $(r+1)$-dimensional array w, so that $N = 2^{j+\mathbf{e}^T\mathbf{j}'}$. Choose $\gamma^2_{j,\mathbf{j}'} = A^2 2^{-j(2s_1+1)-\mathbf{j}'^T(2\mathbf{s_2}+\mathbf{e})}$ and observe that $K(f,\tilde{f}) \le C\varepsilon^{-2}\gamma^2_{jj'}2^{j+\mathbf{e}^T\mathbf{j}'}2^{-2\nu j}$. Now, applying Lemma 6 with

$$\delta^2 = \gamma^2_{jj'}2^{j+\mathbf{e}^T\mathbf{j}'}/32 = A^2 2^{-2s_1 j-2\mathbf{j}'^T\mathbf{s_2}}/32 \tag{9.33}$$

one arrives at the following optimization problem

$$2js_1 + 2j's_2 \Rightarrow \min\ j(2s_1+2\nu+1) + \sum_{l=1}^{r}(2s_{2,l}+1)j'_l \ge \tau_\varepsilon, \quad j, j'_l \ge 0, \tag{9.34}$$

where τ_ε is defined in formula (9.5). Setting $j = \tau_\varepsilon/(2s_1+2\nu+1) - \sum_{l=1}^{r}(2s_l+1)/(2s_1+2\nu+1)$, arrive at optimization problem

$$\frac{2s_1\tau_\varepsilon}{2s_1+2\nu+1} + \sum_{l=1}^{r}\frac{2j'_l[s_{2,l}(2\nu+1)-s_1]}{2s_1+2\nu+1} \Rightarrow \min, \quad j'_l \ge 0,\ l=1,\dots,r. \tag{9.35}$$

If $s_{2,l_0}(2\nu+1) \ge s_1$, then each j'_l is multiplied by a nonnegative number and minimum is attained when $j'_l = 0$, $l = 1,\dots,r$. Then, $j = \tau_\varepsilon/(2s_1+2\nu+1)$. On the other hand, if $s_{2,l_0}(2\nu+1) < s_1$, then j_{l_0} is multiplied by the smallest factor which is negative. Therefore, minimum in (9.35) is attained if $j = 0$, $j'_l = 0$, $l \ne l_0$ and $j_{l_0} = \tau_\varepsilon/(2s_{2,l_0}+1)$. Plugging those values into (9.33), obtain

$$\delta^2 = \begin{cases} CA^2\,(\varepsilon^2/A^2)^{\frac{2s_{2,0}}{2s_{2,0}+1}}, & \text{if } s_1 > s_{2,0}(2\nu+1), \\ CA^2\,(\varepsilon^2/A^2)^{\frac{2s_1}{2s_1+2\nu+1}}, & \text{if } s_1 \le s_{2,0}(2\nu+1). \end{cases} \tag{9.36}$$

In the sparse-dense case, we use r-dimensional array w, so that $N = 2^{\mathbf{e}^T\mathbf{j}'}$. Choose $\gamma^2_{j,\mathbf{j}'} = A^2 2^{-2js_1^*-\mathbf{j}'^T(2\mathbf{s_2}+\mathbf{e})}$ and observe that

$$K(f,\tilde{f}) \le C\varepsilon^{-2}\gamma^2_{jj'}2^{j+\mathbf{e}^T\mathbf{j}'}2^{-2\nu j}$$

Now, applying Lemma 6 with

$$\delta^2 = A^2 2^{-2s_1^* j-2\mathbf{j}'^T\mathbf{s_2}}/32 \tag{9.37}$$

one arrives at the following optimization problem

$$2js_1 + 2j's_2 \Rightarrow \min\ j(2s_1^*+2\nu+1) + \sum_{l=1}^{r}(2s_{2,l}+1)j'_l \ge \tau_\varepsilon, \quad j, j'_l \ge 0, \tag{9.38}$$

Again, setting $j = \tau_\varepsilon/(2s_1^*+2\nu) - \sum_{l=1}^{r}(2s_l+1)/(2s_1^*+2\nu)$, arrive at optimization problem

$$\frac{2s_1^*\tau_\varepsilon}{2s_1^*+2\nu} + \sum_{l=1}^{r}\frac{2j'_l[2s_{2,l}\nu - s_1^*]}{2s_1^*+2\nu} \Rightarrow \min, \quad j'_l \ge 0,\ l=1,\dots,r. \tag{9.39}$$

Repeating the reasoning applied in the dense-dense case, we obtain $j=0$, $j'_l=0$, $l\neq l_0$ and $j_{l_0}=\tau_\varepsilon/(2s_{2,l_0}+1)$ if $2s_{2,l_0}\nu<s_1^*$, and $j=\tau_\varepsilon/(2s_1+2\nu+1)$, $j'_l=0$, $l=1,\dots,r$, if $2s_{2,l_0}\nu>s_1^*$. Plugging those values into (9.37), obtain

$$\delta^2=\begin{cases} CA^2\left(\varepsilon^2/A^2\right)^{\frac{2s_{2,0}}{2s_{2,0}+1}}, & \text{if } \ 2\nu s_{2,0}\le s_1^*,\\ CA^2\left(\varepsilon^2/A^2\right)^{\frac{2s_1^*}{2s_1^*+2\nu}}, & \text{if } \ 2\nu s_{2,0}>s_1^*. \end{cases} \tag{9.40}$$

In order to complete the proof, combine (9.36) and (9.40) and note that $s_1^*=s'_1$ if $p\le 2$. □

Proof of Theorem 5. Repeat the proof of Theorem 2 with j' and k' replaced by $\mathbf{j}'$ and $\mathbf{k}'$, respectively, $s_2 j'$ replaced by ${\mathbf{j}'}^T\mathbf{s}'_{\mathbf{2}}$ and

$$2^{j_0}=\left(\chi_{\varepsilon,A}\right)^{-\frac{d}{2s'_1}},\quad 2^{j'_{0,l}}=\left(\chi_{\varepsilon,A}\right)^{-\frac{d}{2s'_{2,l}}},\quad l=1,\dots,r.$$

Then, formulae (9.18)–(9.22) are valid. One can also partition Δ in (9.23) into Δ_1, Δ_2 and Δ_3 given by expressions similar to (9.24), (9.25) and (9.26) with $r+1$ sums in (9.24) instead of two, $\sum_{j'=m'_0}^{j'_0}$ replaced by r respective sums and $\mathbf{1}\left(2^{j(2\nu+1)+j'}>\chi_{\varepsilon,A}^{d-1}\right)$ replaced by $\mathbf{1}\left(2^{j(2\nu+1)+\mathbf{e}^T\mathbf{j}'}>\chi_{\varepsilon,A}^{d-1}\right)$. Then, upper bounds (9.27) and (9.28) hold. In order to construct upper bounds for Δ_3, we again need to consider three different cases.

In Case 1, $s_1\ge s_{2,0}(2\nu+1)$, replace $\sum_{j'=m'_0}^{j'_0}$ by $\sum_{j'_{l_0}=m'_{l_0}}^{j'_{0,l_0}}$ and $\sum_{j=m_0}^{j_0}$ by the sum over j, $j'_1,\dots,j'_{l_0-1},j'_{l_0+1},\dots,j'_r$. Repeating calculations for this case, keeping in mind that $s'_{2,l}\ge s'_{2,0}$ for any l and noting that, whenever $s'_{2,l}=s'_{2,0}$, we gain an extra logarithmic factor, we arrive at

$$\Delta_3=O\left(A^2\,\chi_{\varepsilon,A}^d\,[\ln(1/\varepsilon)]^{\mathbf{1}(s_1=s_2(2\nu+1))+\sum_{l\neq l_0}\mathbf{1}(s_{2,l}=s_{2,0})}\right). \tag{9.41}$$

In Case 2, $(1/p-1/2)(2\nu+1)<s_1<s_{2,0}(2\nu+1)$, replace $\sum_{j'=m'_0}^{j'_0}$ by $\sum_{\mathbf{j}'\in\Upsilon(\mathbf{m}',\mathbf{j}'_\mathbf{0})}$ where $\mathbf{j}'_\mathbf{0}=(j'_{0,1},\dots,j'_{0,r})$ and arrive at (9.30). In Case 3, $s_1\le(\frac{1}{p}-\frac{1}{2})(2\nu+1)$, since the sum over $\mathbf{j}'$ is uniformly bounded, calculations for the two-dimensional case hold and (9.31) is valid. Combination of (9.41), (9.30) and (9.31) completes the proof. □

References

[1] Abramovich, F. and Silverman, B. W. (1998). Wavelet decomposition approaches to statistical inverse problems. *Biometrika.* **85** 115-129. MR1627226

[2] Berkhout, A. J. (1986). The seismic method in the search for oil and gas: Current techniques and future developments. *Proc. IEEE.* **74** 1133-1159.

[3] Bunea, F.,Tsybakov, A. and Wegkamp, M. H. (2007). Aggregation for Gaussian regression. *Ann. Statist.* **35** 1674-1697. MR2351101

[4] Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika.* **81** 425-456. MR1311089

[5] Donoho, D. L. (1995). Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Applied and Computational Harmonic Analysis.* **2** 101-126. MR1325535

[6] Donoho, D. L. and Raimondo, M. (2004). Translation invariant deconvolution in a periodic setting. *International Journal of Wavelets, Multiresolution and Information Processing.* **14** 415-432. MR2104873

[7] Fan, J. and Koo, J. (2002). Wavelet deconvolution. *IEEE Transactions on Information Theory.* **48** 734-747. MR1889978

[8] Heimer, A. and Cohen, I. (2008). Multichannel blind seismic deconvolution using dynamic programming. *Signal Processing.* **88** 1839-1851

[9] Heping, W. (2004). Representation and approximation of multivariate functions with mixed smoothness by hyperbolic wavelets *J. Math. Anal. Appl.* **291** 698-715. MR2039080

[10] Holdai, V. and Korostelev, A. (2008). Image Reconstruction in Multi-Channel Model under Gaussian Noise. *Math. Meth. Statist.* **17** 198-208. MR2448946

[11] Johnstone, I. M., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004). Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society.* Series B, **66** 547-573 (with discussion, 627-657). MR2088290

[12] Johnstone, I. M. and Raimondo, M. (2004). Periodic boxcar deconvolution and Diophantine approximation. *Annals of Statistics.* **32** 1781-1804. MR2102493

[13] Kalifa, J. and Mallat, S. (2003). Thresholding estimators for linear inverse problems and deconvolutions. *Annals of Statistics.* **31** 58-109. MR1962500

[14] Kerkyacharian, G., Lepski, O. and Picard, D. (2001). Nonlinear estimation in anisotropic multi-index denoising. *Probab. Theory Relat. Fields.* **121** 137-170. MR1863916

[15] Kerkyacharian, G.,Lepski, O. and Picard, D. (2008). Nonlinear Estimation in Anisotropic Multi-Index Denoising. Sparse Case. *Theory Probab. Appl.* **52** 58-77. MR2354574

[16] Kerkyacharian, G., Picard, D. and Raimondo, M. (2007). Adaptive boxcar deconvolution on full Lebesgue measure sets. *Statistica Sinica.* **7** 317-340. MR2352512

[17] Neelamani, R., Choi, H. and Baraniuk, R. (2004). Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Transactions on Signal Processing.* **52** 418-433. MR2044455

[18] Nikolskii S .M. (1975). *Approximation of functions of several variables and imbedding theorems* (Russian). Sec. ed., Moskva, Nauka 1977 English translation of the first ed., Berlin. MR0506247

[19] Pensky, M. and Sapatinas, T. (2009). Functional Deconvolution in a Periodic Setting: Uniform Case. *Annals of Statistics.* **37** 73-104. MR2488345

[20] Pensky, M. and Sapatinas, T. (2010). On Convergence Rates Equivalency and Sampling Strategies in a Functional Deconvolution Model. *Annals of Statistics.* **38** 1793-1844. MR2662360

[21] Pensky, M. and Sapatinas, T. (2011). Multichannel Boxcar Deconvolution with Growing Number of Channels. *Electronic Journal of Statistics.* **5** 53-82. MR2773608

[22] Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *Annals of Statistics.* **27** 2033-2053. MR1765627

[23] Robinson, E. A. (1999). *Seismic Inversion and Deconvolution: Part B: Dual-Sensor Technology.* Elsevier, Oxford.

[24] Robinson, E. A., Durrani, T. S. and Peardon, L. G. (1986). *Geophysical Signal Processing.* Prentice-Hall, London.

[25] Tsybakov, A. B. (2008). *Introduction to Nonparametric Estimation.* Springer, New York. MR2013911

[26] Walter, G. and Shen, X. (1999). Deconvolution using Meyer wavelets. *Journal of Integral Equations and Applications.* **11** 515-534. MR1738280

[27] Wason, C. B., Black, J. L. and King, G. A. (1984). Seismic modeling and inversion. *Proc. IEEE.* **72** 1385-1393.