Stein block thresholding for wavelet-based image deconvolution

Christophe Chesneau, Jalal M. Fadili, Jean-Luc Starck

▶ **To cite this version:**

Christophe Chesneau, Jalal M. Fadili, Jean-Luc Starck. Stein block thresholding for wavelet-based image deconvolution. Electronic Journal of Statistics , 2010, 4, pp.415-435. 10.1214/09-EJS550 . hal-00436661v2

HAL Id: hal-00436661
https://hal.science/hal-00436661v2

Submitted on 4 Apr 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.

Electronic Journal of Statistics
Vol. 4 (2010) 415–435
ISSN: 1935-7524
DOI: 10.1214/09-EJS550

Stein block thresholding for wavelet-based image deconvolution

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, 14050 Caen France
e-mail: Jalal.Fadili@greyc.ensicaen.fr

Jean-Luc Starck

DAPNIA/SEDI-SAP CEA-Saclay, 91191 Gif-sur-Yvette France
e-mail: jstarck@cea.fr

Abstract: In this paper, we propose a fast image deconvolution algorithm that combines adaptive block thresholding and Vaguelet-Wavelet Decomposition. The approach consists in first denoising the observed image using a wavelet-domain Stein block thresholding, and then inverting the convolution operator in the Fourier domain. 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. The resulting algorithm is simple to implement and fast. Its computational complexity is dominated by that of the FFT. We report a simulation study to support our theoretical findings. The practical performance of our block vaguelet-wavelet deconvolution compares very favorably to existing competitors on a large set of test images.

AMS 2000 subject classifications: Primary 62M10, 62F12; secondary 62F12.
Keywords and phrases: Image deconvolution, block thresholding, wavelets, minimax, LaTeX 2_ε.

Received November 2009.

Contents

1. Introduction

1.1. Problem statement and prior work

In this paper, we consider the two-dimensional convolution model with Gaussian white noise $\sim \mathcal{N}(0, \sigma^2)$. We observe the stochastic process $Y(.)$ where

$$dY(\mathbf{x}) = T(f)(\mathbf{x})d\mathbf{x} + \sigma dW(\mathbf{x}), \tag{1.1}$$

$\mathbf{x} \in [0,1]^2$, $W(.)$ is a (non-observed) white Gaussian noise, $T(f)(\mathbf{x}) = (f \star g)(\mathbf{x})$ is the two-dimensional convolution operator on $[0,1]^2$, g is a known kernel (called also point spread function PSF), both f and g are one-periodic functions belonging to $\mathbb{L}_2([0,1]^2)$. In the sequel, the Fourier transform of a function h will be denoted $\mathcal{F}(h)(\mathbf{l}) = \int_{[0,1]^2} h(\mathbf{x})e^{-2i\pi<\mathbf{l},\mathbf{x}>}d\mathbf{x}$. The observation model (1.1) illustrates the action of a linear time-invariant system on an input image f when the data are corrupted with additional noise. Deconvolution is to estimate f from Y which is a longstanding inverse problem in many areas of signal and image processing. Application fields cover biomedical imaging, astronomical imaging, remote-sensing, seismology, etc. This list is by no means exhaustive.

There is an extensive statistical literature on wavelet-based deconvolution problems. For obvious space limitations, we only focus on *some* of them. In 1D, Donoho in [10] gave the first discussion of wavelet thresholding in linear inverse problems and introduced the Wavelet-Vaguelet Decomposition (WVD). The WaveD algorithm as described in [13] is an adaptation of WVD to the one-dimensional deconvolution problem. Abramovich and Silverman in [1] proposed another procedure; the Vaguelet-Wavelet Decomposition (VWD). The VWD estimator in its original form relies on standard term-by-term thresholding rules. It has been improved by Cai in [3] using a Stein block thresholding rule. In the same vein as the block extension of VWD, the original term-by-term thresholding-based WaveD procedure has been recently enhanced by Chesneau [5] using again block thresholding.

In 2D, the WVD approach was refined in [14] where the authors proposed a mirror wavelet basis adapted to capture the singularity of the spectrum of

the inverse of h. The authors in [17] advocated a hybrid approach known as ForWarD. In [9], the authors described a two-dimensional adaptation of the WaveD algorithm [13], and they showed that it enjoys good numerical performance. Deconvolution methods based on variational or bayesian formulations with sparsity-promoting regularization over wavelet coefficients have been recently proposed; see e.g. [4, 7, 11, 12] and others. These algorithms are based on iterative thresholding. A few papers have also focused on deconvolution by using the SURE principle to minimize an unbiased estimate of the MSE of deconvolution estimators operating by thresholding in orthogonal or redundant wavelet bases [18, 20].

1.2. Contributions

However, so far, these two-dimensional wavelet deconvolution algorithms were based on term-by-term (or individual) thresholding which clearly under-performs for many images. The drawback of individual thresholding cannot be circumvented by fine-tuning the regularization/threshold parameter. These reasons motivated us to develop an adaptive estimator of f based on combining two-dimensional Stein block thresholding and VWD. The approach consists in first denoising the observed image using a wavelet-domain block thresholding, and then inverting the convolution operator in the Fourier domain. It can be viewed as a two-dimensional version of the procedure developed by [3]. In fact, although we focus on the two-dimensional case for clarity and accessibility to a larger audience, our procedure and results apply equally well to dimensions higher than two; see also Remark 4.1.

From a theoretical point of view, taking the minimax approach over the Besov balls $\mathbf{B}^s_{p,q}(M)$ (to be defined in Section 2) and under the $\mathbb{L}_2$ risk, we prove that our estimator achieves near optimal rates of convergence. This is featured in Theorem 4.1. These rates are for instance better than those attained by the two-dimensional WaveD of [9]. Using lower bound techniques, we also provide a lower bound on the risk over the Besov ball as stated in Theorem 4.2. From a practical point of view, our algorithm is very simple to implement and runs very fast. Its performances compare very favorably to alternative deconvolution algorithms in the literature such as [7, 9, 12, 17] over a large set of test images.

1.3. Paper organization

The paper is organized as follows. Section 2 briefly reviews wavelets and Besov balls. Section 3 describes the block thresholding-based deconvolution estimator. The minimax performances of this estimator and the lower bound of its risk are investigated in Section 4. Section 5 describes the experimental results, before drawing some conclusions in Section 6. The proofs are assembled in 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. The main features of such wavelets are:

1. they are band-limited, i.e. the Fourier transforms of ϕ and ψ have compact supports respectively included in $[-4\pi 3^{-1}, 4\pi 3^{-1}]$ and $[-8\pi 3^{-1}, -2\pi 3^{-1}] \cup [2\pi 3^{-1}, 8\pi 3^{-1}]$.
2. for any frequency in $[-2\pi, -\pi] \cup [\pi, 2\pi]$, there exists a constant $c > 0$ such that the magnitude of the Fourier transform of ψ is lower bounded by c.
3. the functions (ϕ, ψ) are $\mathcal{C}^\infty$ 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:

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

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

$$|\phi(t)| = O\left((1+|t|)^{-r-1}\right), \qquad |\psi(t)| = O\left((1+|t|)^{-r-1}\right),$$

meaning that ϕ and ψ are not very well localized in time. This is why a Meyer wavelet transform is generally implemented in the Fourier domain.

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

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

be the elements of the wavelet basis, and

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

their periodized versions. There exists an integer j_* such that the collection $\{\phi^{\mathrm{per}}_{j_*,k}, k = 0, \ldots, 2^{j_*} - 1;\ \psi^{\mathrm{per}}_{j,k}, \quad j = j_*, \ldots, \infty,\ k = 0, \ldots, 2^j - 1\}$ forms an orthonormal basis of $\mathbb{L}_2^{\mathrm{per}}([0,1])$. In what follows, the superscript "per" will be dropped to lighten the notation.

In higher dimension, and 2D in particular as treated in this paper, we consider tensor product wavelet bases on $\mathbb{L}_2([0,1]^2)$. Let us briefly recall their construction (see, for instance, [15] for more details). Define the tensor product wavelets Φ, Ψ^1, Ψ^2 and Ψ^3 as respectively

$$\Phi(\mathbf{x}) = \phi(x)\phi(y), \quad \Psi^1(\mathbf{x}) = \psi(x)\phi(y), \quad \Psi^2(\mathbf{x}) = \phi(x)\psi(y), \quad \Psi^3(\mathbf{x}) = \psi(x)\psi(y),$$

$\forall \mathbf{x} = (x, y) \in [0,1]^2$. For any orientation $i \in \{1,2,3\}$, scale $j \geq 0$ and spatial location $\mathbf{k} = (k_1, k_2) \in D_j = \{0, \ldots, 2^j - 1\}^2$, we define the translated and scaled versions $\Phi_{j,\mathbf{k}}(\mathbf{x}) = 2^j \Phi(2^j x - k_1, 2^j y - k_2)$ and $\Psi_{j,i,\mathbf{k}}(\mathbf{x}) = 2^j \Psi^i(2^j x - k_1, 2^j y - k_2)$.
Any function $f \in \mathbb{L}_2([0,1]^2)$ can be expanded into a wavelet series

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in D_{j_0}} \alpha_{j_0,\mathbf{k}} \Phi_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{i=1}^{3} \sum_{j \geq j_0} \sum_{\mathbf{k} \in D_j} \beta_{j,i,\mathbf{k}} \Psi_{j,i,\mathbf{k}}(\mathbf{x}),$$

$\mathbf{x} \in [0,1]^2$, where $\alpha_{j,\mathbf{k}} = \int_{[0,1]^2} f(\mathbf{x}) \Phi_{j,\mathbf{k}}(\mathbf{x}) d\mathbf{x}$ and $\beta_{j,i,\mathbf{k}} = \int_{[0,1]^2} f(\mathbf{x}) \Psi_{j,i,\mathbf{k}}(\mathbf{x}) d\mathbf{x}$ are the wavelet coefficients of f. See [16, Vol. 1 Chapter III.11] for a detailed account on periodized orthonormal wavelet bases.

2.2. Besov balls

We say that a function f in $\mathbb{L}_2([0,1]^2)$ belongs to the bi-dimensional (isotropic) Besov ball $\mathbf{B}^s_{p,q}(M)$ if, and only if, $\int_{[0,1]^2} f^2(\mathbf{x}) d\mathbf{x} \leq M$ and there exists a constant M_*, depending on M, such that the wavelet coefficients of f satisfy

$$\left(\sum_{i=1}^{3} \sum_{j \geq 0} \left(2^{j(s+1-2/p)} \left(\sum_{\mathbf{k} \in D_j} |\beta_{j,i,\mathbf{k}}|^p \right)^{1/p} \right)^q \right)^{1/q} \leq M_*,$$

with a smoothness parameter $s > 0$, and the norm parameters: $0 < p \leq \infty$ and $0 < q \leq \infty$. Such Besov spaces contain both smooth images and those with sharp edges.

3. The deconvolution block estimator

3.1. Smoothness of the kernel g

For the theoretical study, the following smoothness assumption on g will play an essential role. We suppose that there exist four constants, $c > 0$, $C > 0$, $\delta_1 > 1/2$ and $\delta_2 > 1/2$, such that, for any $\mathbf{l} = (l_1, l_2) \in \mathbb{Z}^2$, the Fourier transform of g satisfies

$$c(1 + |l_1|^{\delta_1})^{-1}(1 + |l_2|^{\delta_2})^{-1} \leq |\mathcal{F}(g)(\mathbf{l})| \leq C(1 + |l_1|^{\delta_1})^{-1}(1 + |l_2|^{\delta_2})^{-1}. \quad (3.1)$$

In words, this means that the Fourier transform of the blurring PSF does not decay too fast, typically in a polynomial fashion within its bandwidth. This assumption controls the decay of the Fourier coefficients of g, or equivalently its smoothness. It is a standard hypothesis usually adopted in the field of nonparametric estimation for deconvolution problems, see e.g. [3, 9, 10, 17]. The parameters (δ_1, δ_2) quantify the spectral decay rate, hence the ill-conditioning,

of the convolution operator associated to g. For example, it is easy to check that the square integrable one-periodic function g defined by $g(x,y) = h(x)h(y)$ where $h(x) = \sum_{m\in\mathbb{Z}} e^{-|x+m|}$, $x \in [0,1]$, satisfies (3.1). Indeed, for any $l \in \mathbb{R}$, we have $\mathcal{F}(h)(l) = 2\left(1+4\pi^2 l^2\right)^{-1}$. Hence, for any $\mathbf{l} = (l_1, l_2) \in \mathbb{R}^2$, $\mathcal{F}(g)(\mathbf{l}) = \mathcal{F}(h)(l_1)\mathcal{F}(h)(l_2)$ satisfies (3.1) with $c = 4(1+4\pi^2)^{-2}$, $C = (2\pi^2)^{-2}$ and $\delta_1 = \delta_2 = 2$. This assumption goes by the name of *ordinary smooth case*.

3.2. Vaguelet-wavelet decomposition

Although the VWD is valid for more general operators T, we here restrict our description to the case of convolution where the VWD takes a simple form.

Proposition 3.1. *For any $i \in \{1,2,3\}$, $j \geq j_0$ and $\mathbf{k} \in D_j$, set*

$$w_{j,i,\mathbf{k}}(\mathbf{x}) = 2^{-j(\delta_1+\delta_2)} T^{-1}(\Psi_{j,i,\mathbf{k}})(\mathbf{x}), \qquad \mathbf{x} \in [0,1]^2,$$

where, for any $h \in \mathbb{L}_2([0,1]^2)$,

$$\begin{aligned} T^{-1}(h)(\mathbf{x}) &= \mathcal{F}^{-1}\left(\mathcal{F}(h)(.)/\mathcal{F}(g)(.)\right)(\mathbf{x}) \\ &= \int_{\mathbb{R}^2} \left(\mathcal{F}(h)(\mathbf{l})/\mathcal{F}(g)(\mathbf{l})\right) e^{2i\pi<\mathbf{l},\mathbf{x}>} d\mathbf{l}. \end{aligned} \tag{3.2}$$

Then, under assumption (3.1), $(\omega_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$ is a Riesz sequence, i.e. there exist two constants $c > 0$ and $C > 0$ such that

$$\begin{aligned} c\sum_{i=1}^{3}\sum_{j\geq j_0}\sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}}^2 &\leq \int_{[0,1]^2} \left(\sum_{i=1}^{3}\sum_{j\geq j_0}\sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}}\omega_{j,i,\mathbf{k}}(\mathbf{x})\right)^2 d\mathbf{x} \\ &\leq C\sum_{i=1}^{3}\sum_{j\geq j_0}\sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}}^2. \end{aligned}$$

for every square-summable sequence $(a_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$.

Thanks to Proposition 3.1, under assumption (3.1), any function $f \in \mathbb{L}_2([0,1]^2)$ can be expanded into a vaguelet-wavelet series

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in D_{j_0}} \vartheta_{j_0,\mathbf{k}}\omega_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{i=1}^{3}\sum_{j\geq j_0}\sum_{\mathbf{k}\in D_j} \theta_{j,i,\mathbf{k}} w_{j,i,\mathbf{k}}(\mathbf{x}), \quad \forall \mathbf{x} \in [0,1]^2\ , \tag{3.3}$$

where

$$\begin{aligned} \vartheta_{j_0,\mathbf{k}} &= 2^{j_0(\delta_1+\delta_2)} \int_{[0,1]^2} T(f)(\mathbf{x})\Phi_{j_0,\mathbf{k}}(\mathbf{x})d\mathbf{x}, \\ \theta_{j,i,\mathbf{k}} &= 2^{j(\delta_1+\delta_2)} \int_{[0,1]^2} T(f)(\mathbf{x})\Psi_{j,i,\mathbf{k}}(\mathbf{x})d\mathbf{x}, \\ \omega_{j_0,\mathbf{k}}(\mathbf{x}) = 2^{-j_0(\delta_1+\delta_2)} T^{-1}(\Phi_{j_0,\mathbf{k}})(\mathbf{x}), &\qquad w_{j,i,\mathbf{k}}(\mathbf{x}) = 2^{-j_0(\delta_1+\delta_2)} T^{-1}(\Psi_{j,i,\mathbf{k}})(\mathbf{x}). \end{aligned}$$

Further details on vaguelet-wavelet series can be found in [1].

3.3. Gaussian sequence model

The first step to estimate f consists in estimating the unknown wavelet coefficients of $T(f)$: $(\vartheta_{j_0,\mathbf{k}})_{\mathbf{k}}$ and $(\theta_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$ from the observation Y in (1.1). It follows from (1.1) that

$$y_{j,i,\mathbf{k}} = \theta_{j,i,\mathbf{k}} + \sigma z_{j,i,\mathbf{k}}, \tag{3.4}$$

where

$$y_{j,i,\mathbf{k}} = 2^{j(\delta_1+\delta_2)} \int_{[0,1]^2} \Psi_{j,i,\mathbf{k}}(\mathbf{x}) dY(\mathbf{x}), \quad z_{j,i,\mathbf{k}} = 2^{j(\delta_1+\delta_2)} \int_{[0,1]^2} \Psi_{j,i,\mathbf{k}}(\mathbf{x}) dW(\mathbf{x}).$$

Thanks to the orthonormality of the wavelet basis, the random variables $(z_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$ are Gaussian i.i.d. with mean 0 and variance $2^{2j(\delta_1+\delta_2)}$.

3.4. Two-dimensional block thresholding estimator

Let the observed image be defined on a $n \times n$ discrete grid of equally-spaced pixels $\{Y(i/n, j/n);\ (i,j) \in \{1,\dots,n\}^2\}$. Let $L = \lfloor (2\log(n))^{1/2} \rfloor$ be the block length, $j_0 = \lfloor \log_2 L \rfloor$ is the coarsest decomposition scale, and $J_* = \lfloor (1/(1+\delta_1+\delta_2)) \log_2(n) \rfloor$. Consider the sequence model (3.4) with $\sigma = 1/n$. For any $\mathbf{k} \in D_{j_0}$, the empirical approximation coefficients are

$$\widehat{\vartheta}_{j_0,\mathbf{k}} = 2^{j_0(\delta_1+\delta_2)} \int_{[0,1]^2} \Phi_{j_0,\mathbf{k}}(\mathbf{x}) dY(\mathbf{x}).$$

For any scale $j \in \{j_0,\dots,J_*\}$, let $\mathcal{A}_j = \{1,\dots,\lfloor 2^j L^{-1} \rfloor\}^2$ be the set indexing the blocks at scale j, and for each block index $\mathbf{K} = (K_1, K_2) \in \mathcal{A}_j$, $U_{j,\mathbf{K}} = \{\mathbf{k} \in D_j;\ (K_1-1)L \le k_1 \le K_1 L - 1,\ (K_2-1)L \le k_2 \le K_2 L - 1\}$ is the set indexing the positions of coefficients within the $\mathbf{K}$th block $U_{j,\mathbf{K}}$.

Now, for any position $\mathbf{k} \in U_{j,\mathbf{K}}$ within the $\mathbf{K}$th block and orientation $i \in \{1,2,3\}$, we estimate the wavelet coefficients $\theta_{j,i,\mathbf{k}}$ of $T(f)$ from $y_{j,i,\mathbf{k}}$ in (3.4) using block Stein thresholding as follows:

$$\widehat{\theta}_{j,i,\mathbf{k}} = \begin{cases} y_{j,i,\mathbf{k}} & \text{if } j \in \{0,\dots,j_0-1\} \\ y_{j,i,\mathbf{k}} \left(1 - \dfrac{\lambda_* \sigma^2 2^{2j(\delta_1+\delta_2)}}{\frac{1}{L^2}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} y_{j,i,\mathbf{k}}^2}\right)_+ & \text{if } j \in \{j_0,\dots,J_*\} \\ 0 & \text{if } j > J_*\ , \end{cases}$$

where $(a)_+ = \max(a,0)$, and λ_* is the root of $x - \log x = 3$, i.e. $\lambda_* = 4.50524\dots$. To estimate f, we reconstruct it from these block-thresholded coefficients as

$$\widehat{f}(\mathbf{x}) = \sum_{\mathbf{k}\in D_{j_0}} \widehat{\vartheta}_{j_0,\mathbf{k}} \omega_{j_0,\mathbf{k}}(\mathbf{x}) + \sum_{i=1}^{3} \sum_{j=j_0}^{J_*} \sum_{\mathbf{K}\in\mathcal{A}_j} \sum_{\mathbf{k}\in U_{j,\mathbf{K}}} \widehat{\theta}_{j,i,\mathbf{k}} w_{j,i,\mathbf{k}}(\mathbf{x}), \quad \mathbf{x} \in [0,1]^2\ . \tag{3.5}$$

4. Minimaxity results

Theorem 4.1 below investigates the minimax rates of convergence attained by $\widehat{f}$ over $\mathbf{B}^s_{p,q}(M)$ under the $\mathbb{L}_2$ risk.

Theorem 4.1 (Upper bound). *Consider the model (1.1). Let $\widehat{f}$ be the estimator defined by (3.5). Then there exists a constant $C > 0$ such that*

$$\sup_{f\in\mathbf{B}^s_{p,q}(M)} \mathbb{E}\left(\int_{[0,1]^2} \left(\widehat{f}(\mathbf{x}) - f(\mathbf{x})\right)^2 d\mathbf{x}\right) \leq C v_\sigma,$$

where

$$v_\sigma = \begin{cases} \sigma^{2s/(s+\delta_1+\delta_2+1)}, & \text{for } 2 \leq p, \\ (\sigma|\log(\sigma)|)^{2s/(s+\delta_1+\delta_2+1)}, & \text{for } p < 2,\ sp > c, \end{cases} \tag{4.1}$$

$c = 2 \vee (2-p)(\delta_1+\delta_2+1)$.

Remark 4.1. *Our work in this paper focuses on the two-dimensional case image for the sake of clarity, simplicity and for illustrative purposes in image processing. Yet the rates of Theorem 4.1 can be extended rather straightforwardly to the d-dimensional case. More precisely, in the d-dimensional case, the rates (4.1) read*

$$v_\sigma = \begin{cases} \sigma^{4s/(2s+2(\delta_1+\delta_2)+d)}, & \text{for } 2 \leq p, \\ (\sigma|\log(\sigma)|)^{4s/(2s+2(\delta_1+\delta_2)+d)}, & \text{for } p<2,\ sp > d \vee (1-p/2)(2(\delta_1+\delta_2)+d), \end{cases}$$

We now turn to the lower bound of the $\mathbb{L}_2$ risk. This is formalized in Theorem 4.2 which determines this lower bound over the ball $\mathbf{B}^s_{p,q}(M)$.

Theorem 4.2 (Lower bound). *Consider the model (1.1). Then there exists a constant $c > 0$ such that*

$$\inf_{\widetilde{f}} \sup_{f\in\mathbf{B}^s_{p,q}(M)} \mathbb{E}\left(\int_{[0,1]^2} \left(\widetilde{f}(\mathbf{x}) - f(\mathbf{x})\right)^2 d\mathbf{x}\right) \geq c\sigma^{2s/(s+\delta_1+\delta_2+1)},$$

where the infimum is taken over all the estimators $\widetilde{f}$ of f.

It can be concluded from Theorems 4.1 and 4.2 that the rate of convergence attained by $\widehat{f}$, i.e. v_σ, is optimal except in the cases $p < 2$ where there is an extra logarithmic term. Additionally, v_σ is better than the one achieved by conventional term-by-term thresholding estimators (e.g. WaveD [9] and others). The main difference is for the case $p \geq 2$ where there is no extra logarithmic term.

5. Experimental results

The proposed block VWD deconvolution method has been compared to three deconvolution methods from the literature: ForWarD [17], wavelet-domain iterative soft-thresholding (IST) with 100 iterations [7, 12], and WaveD [9]. For fair comparison, the regularization parameter of the IST method was tweaked

FIG 1. *Test images.*

TABLE 1
Average execution times (seconds) over then replications for 512×512 *and* 256×256 *images. The algorithms were run under Matlab with an 2.53GHz Intel Core Duo CPU, 4Gb RAM*

Algorithm	Ours	ForWarD [17]	IST [7, 12]	WaveD [9]
512×512	4.08	5.2	119	4.08
256×256	1.2	1.05	29	1.27

manually to reach its best performance. For reliable comparison, we applied the deconvolution algorithms to six standard grayscale images of size 512×512 (Barbara, Lena, Boat) and 256×256 (Cameraman, House, Peppers); see Fig. 1. The blurred images were corrupted by a zero-mean white Gaussian noise such that the blurred signal-to-noise ratio (BSNR $= 10 \log_{10}(\|f \star g\|_\infty / \sigma^2)$) ranged from 10 to 40 dB. At each combination of test image and noise level, ten noisy versions were generated and each deconvolution algorithm was applied to each noisy realization. The output SNR improvement (ISNR) was averaged over the ten replications. The results in the case where the PSF is $g(i,j) = e^{-|i/n|^{0.5}+|j/n|^{0.5}}$ are summarized in Fig. 2. Each plot corresponds to the ISNR as a function of BSNR for each image. These results clearly show that our approach compares very favorably to the best competitors which are ForWarD and iterative soft thresholding. It is even able to outperform them particularly at low BSNR or textured images, while maintaining a low computational cost as reported in Table 1. The computational complexity of block thresholding as well as the WaveD is dominated by the FFT involved in the Meyer wavelet transform that

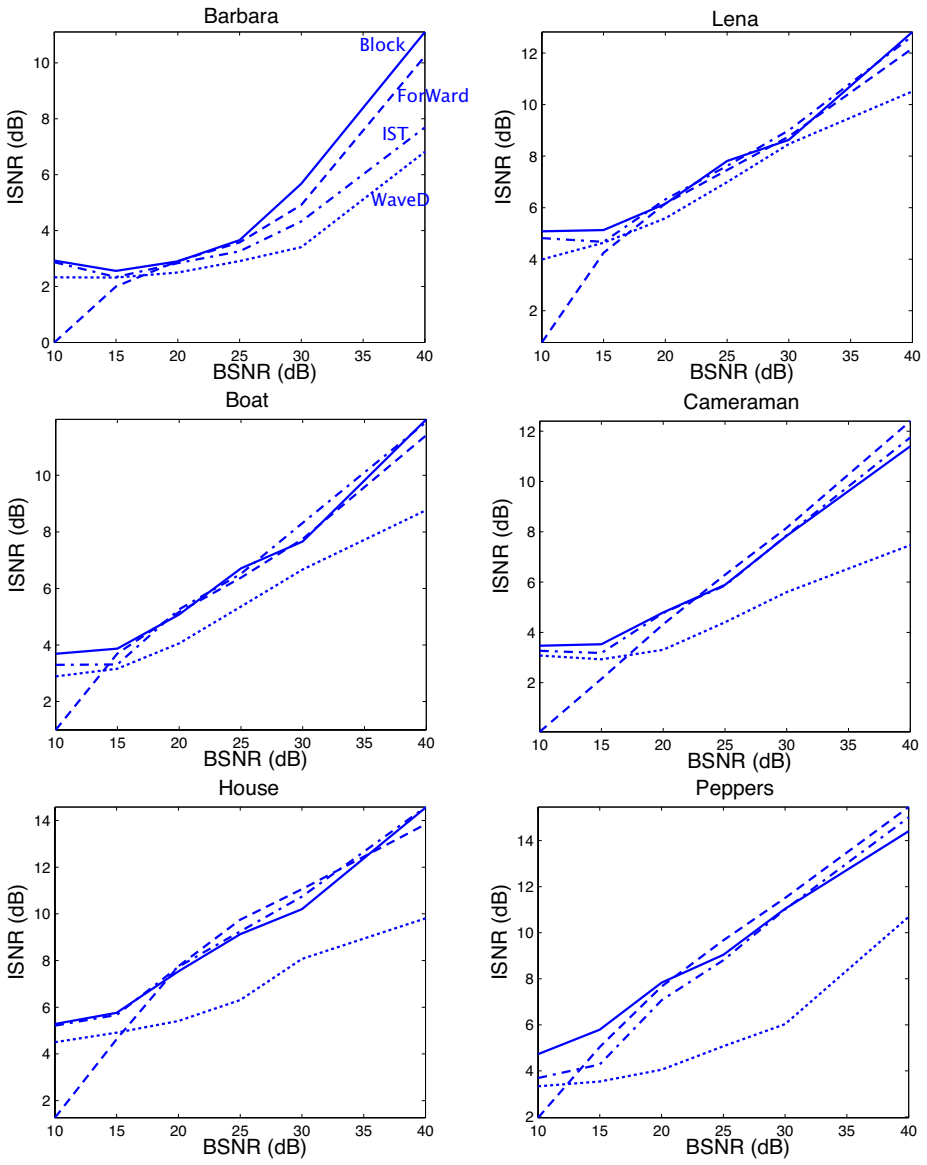

FIG 2. *Comparison of average ISNR in dB over ten realizations for the six test images.*

operates in the Fourier domain. The quantitative results of Fig. 2 are confirmed by visual inspection of Fig. 3 and Fig. 4 which display the results on Barbara and Boat for BSNR=30dB. Again, owing to block thresholding, our VWD deconvolution is able to recover many details (e.g. textured areas on Barbara trousers) better that the other methods.

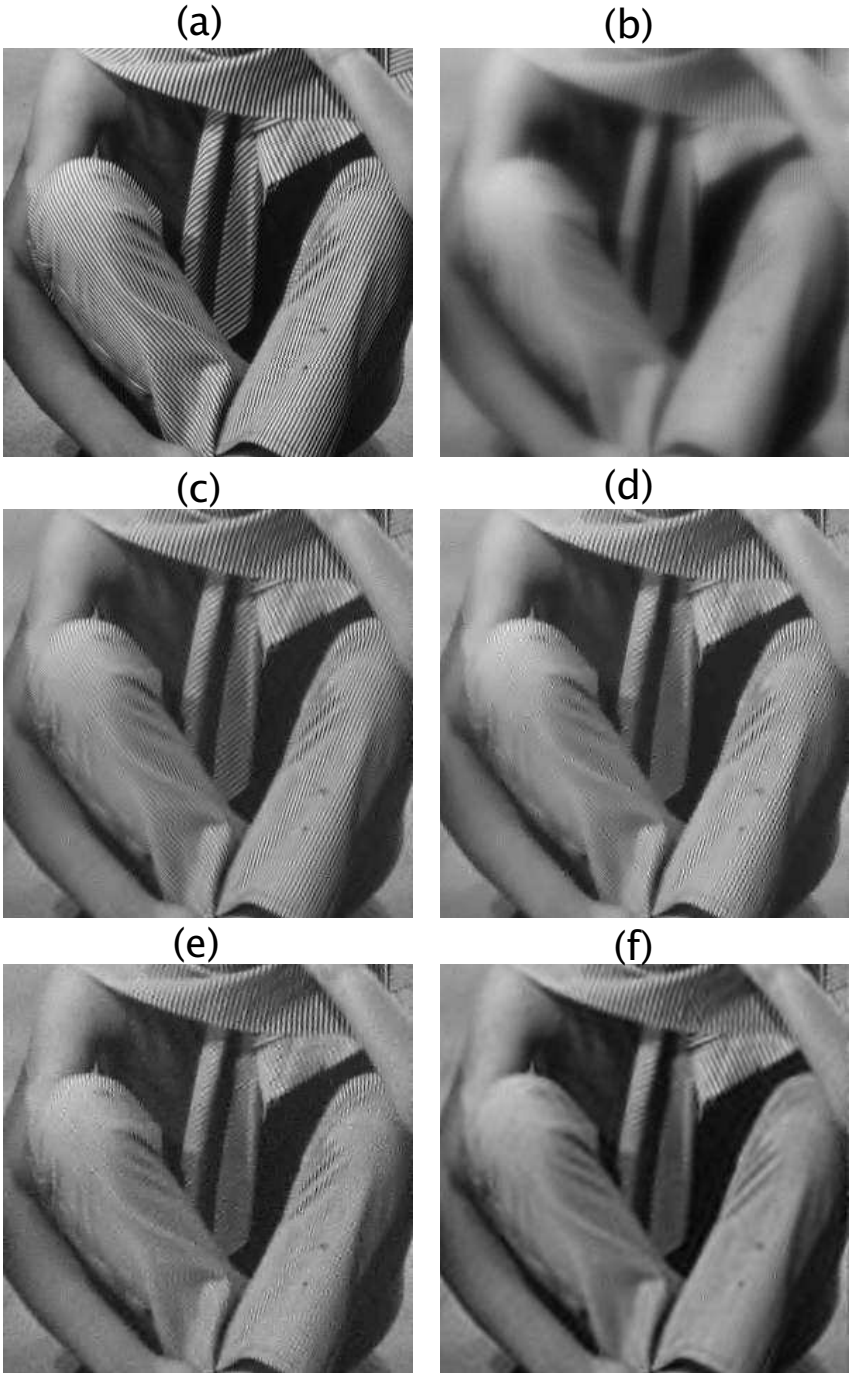

FIG 3. *Deconvolution of Barbara* 512×512. *(a) original, (b) blurred and noisy BSNR=30dB, (c) our method ISNR=5.66dB, (d) ForWarD [17] ISNR=4.9dB, (e) iterative thresholding [7, 12] ISNR=4.33dB, (f) WaveD [9] ISNR=3.4dB.*

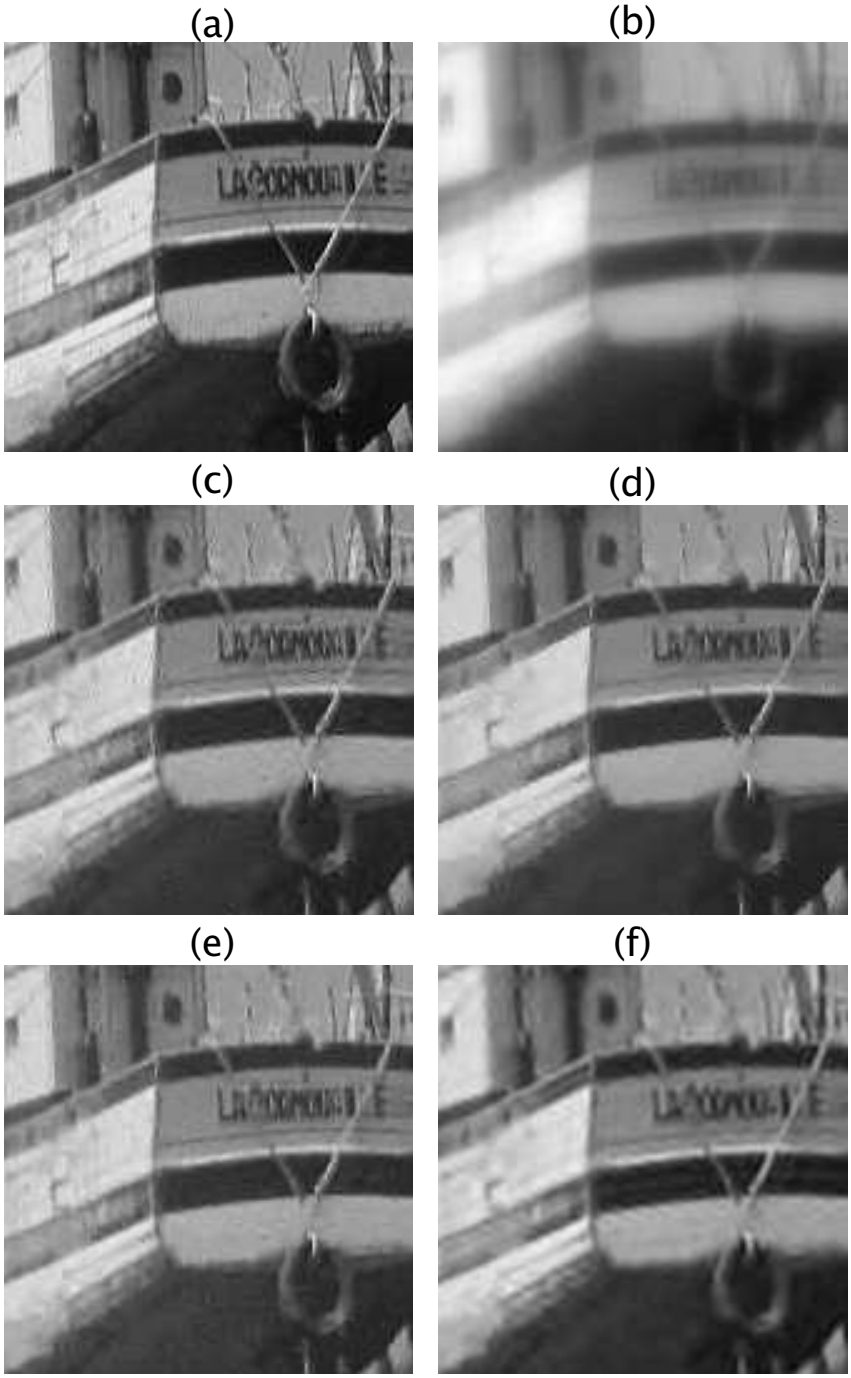

FIG 4. *Deconvolution of Boat* 512×512. *(a) original, (b) blurred and noisy BSNR=30dB, (c) our method ISNR=7.66dB, (d) ForWarD [17] ISNR=7.7dB, (e) iterative thresholding [7, 12] ISNR=8.3dB, (f) WaveD [9] ISNR=6.6dB.*

Following the philosophy of reproducible research, a toolbox is made available freely for download at the address

http://www.greyc.ensicaen.fr/~jfadili/software.html

This toolbox is a collection of Matlab functions, scripts and datasets for image block denoising. It requires at least WaveLab 8.02 [21] to run properly. The toolbox implements the proposed block denoising procedure with several transforms and contains all scripts to reproduce the figures and tables reported in this paper.

6. Conclusion

In this paper, a wavelet-based deconvolution algorithm was presented. It combines the benefits of block-thresholding with vaguelet-wavelet decomposition. Its theoretical and practical performances were established. Although we focused on convolution, the approach can handle other operators $T(f)$. A possible perspective of the present work that we are currently investigating is the theoretical properties of the procedure when other transforms than orthonormal wavelets (the curvelet frame for instance) are used.

7. Proofs

Proof of Proposition 3.1. For every sequence $(a_{j,i,\mathbf{k}})$, the Plancherel formula implies that

$$\begin{aligned}
&\int_{[0,1]^2} \left(\sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}} \omega_{j,i,\mathbf{k}}(\mathbf{x}) \right)^2 d\mathbf{x} \\
&= 2^{-2j(\delta_1+\delta_2)} \int_{[0,1]^2} \left(\mathcal{F}^{-1} \left(\sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}} \mathcal{F}\left(\Psi_{j,i,\mathbf{k}}\right)(.)/\mathcal{F}(g)(.) \right)(\mathbf{x}) \right)^2 d\mathbf{x} \\
&= 2^{-2j(\delta_1+\delta_2)} \int_{\mathbb{R}^2} \left| \sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}} \mathcal{F}\left(\Psi_{j,i,\mathbf{k}}\right)(\mathbf{l})/\mathcal{F}(g)(\mathbf{l}) \right|^2 d\mathbf{l} \\
&= 2^{-2j(\delta_1+\delta_2)} \int_{\mathbb{R}^2} (1/|\mathcal{F}(g)(\mathbf{l})|^2) \left| \mathcal{F}\left(\sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}} \Psi_{j,i,\mathbf{k}} \right)(\mathbf{l}) \right|^2 d\mathbf{l}.
\end{aligned} \tag{7.1}$$

Let $\mathcal{C}_{j,i} = \operatorname{supp}(\Psi_{j,i,\mathbf{k}})$. Then by definition of the wavelet basis and under assumption (3.1), there exist two constants $c > 0$ and $C > 0$ such that

$$c2^{2j(\delta_1+\delta_2)} \leq \inf_{\mathbf{x}\in\mathcal{C}_{j,i}} (1/|\mathcal{F}(g)(\mathbf{x})|^2) \leq \sup_{\mathbf{x}\in\mathcal{C}_{j,i}} (1/|\mathcal{F}(g)(\mathbf{x})|^2) \leq C2^{2j(\delta_1+\delta_2)}. \tag{7.2}$$

Using again the Plancherel formula, we obtain

$$\int_{\mathbb{R}^2} \left| \mathcal{F}\left(\sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}} \Psi_{j,i,\mathbf{k}} \right)(\mathbf{l}) \right|^2 d\mathbf{l}$$
$$= \int_{[0,1]^2} \left(\sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}} \Psi_{j,i,\mathbf{k}}(\mathbf{x}) \right)^2 d\mathbf{x} = \sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}}^2. \quad (7.3)$$

Putting (7.1), (7.2) and (7.3) together, it follows immediately that there are indeed two constants $c > 0$ and $C > 0$ such that

$$c \sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}}^2 \leq \int_{[0,1]^2} \left(\sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}} \omega_{j,i,\mathbf{k}}(\mathbf{x}) \right)^2 d\mathbf{x}$$
$$\leq C \sum_{i=1}^{3} \sum_{j\geq j_0} \sum_{\mathbf{k}\in D_j} a_{j,i,\mathbf{k}}^2.$$

□

Proof of Theorem 4.1. From the Gaussian sequence (3.4), Theorem 4.1 can be proved by applying [6, Theorem 3.1]: it is enough to prove the existence of two constants $Q_3 > 0$ and $Q_4 > 0$ (independent of n) such that the conditions (i) and (ii) below are satisfied:

(i)
$$\sup_{j\in\{0,\ldots,J_*\}} \sup_{i\in\{1,2,3\}} \sup_{\mathbf{k}\in D_j} 2^{-4(\delta_1+\delta_2)j} \mathbb{E}\left(z_{j,i,\mathbf{k}}^4\right) \leq Q_3.$$

(ii) For any $a = (a_{\mathbf{k}})_{\mathbf{k}\in D_j}$ such that $\sup_{j\in\{0,\ldots,J_*\}} \sup_{\mathbf{K}\in\mathcal{A}_j} \sum_{\mathbf{k}\in U_{j,\mathbf{K}}} a_{\mathbf{k}}^2 \leq 1$, we have

$$\sup_{j\in\{0,\ldots,J_*\}} \sup_{i\in\{1,2,3\}} \sup_{\mathbf{K}\in\mathcal{A}_j} 2^{-2(\delta_1+\delta_2)j} \mathbb{E}\left(\left(\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} a_{\mathbf{k}} z_{j,i,\mathbf{k}} \right)^2 \right) \leq Q_4.$$

Since the random variables $(z_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$ are Gaussian i.i.d. with mean 0 and variance $2^{2j(\delta_1+\delta_2)}$, these bounds are obvious.

The steps of the proof are as follows.

We expand f into a vaguelet-wavelet series as in (3.3) with $j_0 = \lfloor \log_2 L \rfloor$. Using Proposition 3.1, we bound the $\mathbb{L}_2$ risk of $\widehat{f}$ in the following way

$$\mathbb{E}\left(\int_{[0,1]^2} \left(\widehat{f}(\mathbf{x}) - f(\mathbf{x}) \right)^2 d\mathbf{x} \right) \leq C(R + S + T), \quad (7.4)$$

where

$$R = \sum_{\mathbf{k}\in D_{j_0}} \mathbb{E}\left(\left(\widehat{\vartheta}_{j_0,k} - \vartheta_{j_0,k}\right)^2\right),$$

$$S = \sum_{i=1}^{3}\sum_{j=j_0}^{J_*}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} \mathbb{E}\left(\left(\widehat{\theta}_{j,i,\mathbf{k}} - \theta_{j,i,\mathbf{k}}\right)^2\right)$$

and

$$T = \sum_{i=1}^{3}\sum_{j=J_*+1}^{\infty}\sum_{\mathbf{k}\in D_j} \theta_{j,i,\mathbf{k}}^2.$$

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

Since $\widehat{\vartheta}_{j_0,k} - \vartheta_{j_0,k} = \sigma 2^{j_0(\delta_1+\delta_2)} \int_{[0,1]^2} \Phi_{j_0,\mathbf{k}}(\mathbf{x}) dW(\mathbf{x}) \sim \mathcal{N}(0, \sigma^2 2^{2j_0(\delta_1+\delta_2)})$, we have

$$R \le C\sigma^2 2^{2j_0(\delta_1+\delta_2+1)} \le \sigma^{2s/(s+\delta_1+\delta_2+1)} \le v_\sigma. \tag{7.5}$$

For any $s > 0$, any $p \ge 1$ and any $r \ge 1$, we have the embedding $\mathbf{B}^s_{p,r}(M) \subseteq \widetilde{\mathbf{B}}^s_{p,r}(M^*)$, where the Besov ball $\widetilde{\mathbf{B}}^s_{p,r}(M^*)$ is defined as $\mathbf{B}^s_{p,r}(M)$ but with the coefficients $(\vartheta_{j,\mathbf{k}})_{j,\mathbf{k}}$, $(\theta_{j,i,\mathbf{k}})_{j,i,\mathbf{k}}$ and a modified radius M^* (see [1] for more details). This embedding together with the inclusions $\widetilde{\mathbf{B}}^s_{p,r}(M) \subseteq \widetilde{\mathbf{B}}^s_{2,\infty}(M)$ if $p \ge 2$, and $\widetilde{\mathbf{B}}^s_{p,r}(M) \subseteq \widetilde{\mathbf{B}}^{s+1-2/p}_{2,\infty}(M)$ if $p \in [1,2)$, imply that

$$T \le C\max(2^{-2J_*s}, 2^{-2J_*(s+1-2/p)}) \le Cv_\sigma. \tag{7.6}$$

Let us set, for any $j \in \{j_0, \ldots, J_*\}$, $\lambda_j = \lambda_* \sigma^2 L^2 2^{2j(\delta_1+\delta_2)}$. The term S can be decomposed as

$$S = u_1 + u_2 + u_3 + u_4, \tag{7.7}$$

where

$$u_1 = \sum_{i=1}^{3}\sum_{j=j_0}^{J_*}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} \mathbb{E}\Bigg(\left(\widehat{\theta}_{j,i,\mathbf{k}} - \theta_{j,i,\mathbf{k}}\right)^2 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} |y_{j,i,\mathbf{k}}|^2 \ge \kappa\lambda_j\right\}} \times 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} |\theta_{j,i,\mathbf{k}}|^2 < \kappa\lambda_j/2\right\}}\Bigg),$$

$$u_2 = \sum_{i=1}^{3}\sum_{j=j_0}^{J_*}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} \mathbb{E}\Bigg(\left(\widehat{\theta}_{j,i,\mathbf{k}} - \theta_{j,i,\mathbf{k}}\right)^2 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} |y_{j,i,\mathbf{k}}|^2 \ge \kappa\lambda_j\right\}} \times 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} |\theta_{j,i,\mathbf{k}}|^2 \ge \kappa\lambda_j/2\right\}}\Bigg),$$

$$u_3 = \sum_{i=1}^{3}\sum_{j=j_0}^{J_*}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} \mathbb{E}\Bigg(\theta_{j,i,\mathbf{k}}^2 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} |y_{j,i,\mathbf{k}}|^2 < \kappa\lambda_j\right\}} 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}} |\theta_{j,i,\mathbf{k}}|^2 \ge 2\kappa\lambda_j\right\}}\Bigg)$$

and

$$u_4 = \sum_{i=1}^{3}\sum_{j=j_0}^{J_*}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}\mathbb{E}\left(\theta_{j,i,\mathbf{k}}^2 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}|y_{j,i,\mathbf{k}}|^2<\kappa\lambda_j\right\}} 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}|\theta_{j,i,\mathbf{k}}|^2<2\kappa\lambda_j\right\}}\right).$$

Upper bounds for u_1 and u_3. We have

$$\max(u_1,u_3) \le C\sum_{i=1}^{3}\sum_{j=j_0}^{J_*}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}\mathbb{E}\left(z_{j,i,\mathbf{k}}^2 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}|z_{j,i,\mathbf{k}}|^2>\kappa\lambda_j/2\right\}}\right).$$

It follows from the Jensen inequality, conditions (i)-(ii) and the Cirelson inequality (see [2]) that

$$\mathbb{E}\left(z_{j,i,\mathbf{k}}^2 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}|z_{j,i,\mathbf{k}}|^2>\kappa\lambda_j/2\right\}}\right) \le C2^{2(\delta_1+\delta_2)j}\sigma^4.$$

Hence

$$\max(u_1,u_3) \le C2^{2(\delta_1+\delta_2+1)J_*}\sigma^4 \le \sigma^2 \le v_\sigma. \tag{7.8}$$

Upper bound for u_2. By the Jensen inequality and condition (i), we derive the bound

$$\mathbb{E}\left(z_{j,i,\mathbf{k}}^2\right) \le \left(\mathbb{E}\left(z_{j,i,\mathbf{k}}^4\right)\right)^{1/2} \le C2^{2(\delta_1+\delta_2)j}\sigma^2.$$

Hence

$$u_2 \le C\sigma^2\sum_{i=1}^{3}\sum_{j=j_0}^{J_*}2^{2(\delta_1+\delta_2)j}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}|\theta_{j,k}|^2>\kappa\lambda_j/2\right\}}.$$

Splitting the sum by considering the integer j_2 defined by $2^{-1}\left(n/\ln n\right)^{1/(s+\delta_1+\delta_2+1)} < 2^{j_2} \le \left(n/\ln n\right)^{1/(s+\delta_1+\delta_2+1)}$, using the block structure and inclusions of Besov balls $\widetilde{\mathbf{B}}_{p,r}^{s}(M^*)$ similar to those used for the bound of u_1, we obtain

$$u_2 \le Cv_\sigma. \tag{7.9}$$

Upper bound for u_4. We have

$$u_4 \le \sum_{i=1}^{3}\sum_{j=j_0}^{J_*}2^{2(\delta_1+\delta_2)j}\sum_{\mathbf{K}\in\mathcal{A}_j}\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}\theta_{j,k}^2 1_{\left\{\sum_{\mathbf{k}\in U_{j,\mathbf{K}}}|\theta_{j,k}|^2<2\kappa\lambda_j\right\}}.$$

We then use the same argument as for u_3 by the splitting the sum exactly in the same way to arrive at the bound

$$u_4 \le Cv_\sigma. \tag{7.10}$$

Combining (7.4) - (7.10), we have

$$\mathbb{E}\left(\int_{[0,1]^2}\left(\widehat{f}(\mathbf{x})-f(\mathbf{x})\right)^2 d\mathbf{x}\right)\leq Cv_\sigma.$$

□

Proof of Theorem 4.2. We need the following result, consequence of the Fano lemma.

Lemma 7.1. *Let $m\in\mathbb{N}^*$ and A be a sigma algebra on the space Ω. For any $i\in\{0,\ldots,m\}$, let $A_i\in A$ such that, for any $(i,j)\in\{0,\ldots,m\}^2$ with $i\neq j$,*

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

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

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

where

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

and K is the Kullback-Leibler divergence defined by

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

The proof of Lemma 7.1 can be found in [8, Lemma 3.3]. For further details and applications of the Fano lemma, see [19].

Consider the Besov balls $\mathbf{B}^s_{p,q}(M)$. Let j_1 be an integer which will be suitably chosen at the end of this proof. For any $\varepsilon=(\varepsilon_{i,\mathbf{k}})_{(i,\mathbf{k})\in\{1,2,3\}\times D_{j_1}}\in\{0,1\}^{3\times 2^{2j_1}}$, set

$$h_\varepsilon(\mathbf{x})=M_*2^{-j_1(s+1)}\sum_{i=1}^{3}\sum_{\mathbf{k}\in D_{j_1}}\varepsilon_{i,\mathbf{k}}\Psi_{j_1,i,\mathbf{k}}(\mathbf{x}),\qquad\mathbf{x}\in[0,1]^2.$$

Now, for any scale $j\geq\tau$, subband $i\in\{1,2,3\}$ and position $\mathbf{k}\in D_{j_1}$, the (mother) wavelet coefficient of h_ε is by orthonormality of the Meyer wavelet basis

$$\beta_{j,i,\mathbf{k}}=\int_{[0,1]^2}h_\varepsilon(\mathbf{x})\Psi_{j,i,\mathbf{k}}(\mathbf{x})d\mathbf{x}=\begin{cases}M_*\varepsilon_{i,\mathbf{k}}2^{-j_1(s+1)}, & \text{if } j=j_1,\\ 0, & \text{otherwise.}\end{cases}$$

Therefore $h_\varepsilon\in\mathbf{B}^s_{p,q}(M)$. The Varshamov-Gilbert theorem (see [19, Lemma 2.7]) asserts that there exist a set $E_{j_1}=\left\{\varepsilon^{(0)},\ldots,\varepsilon^{(T_{j_1})}\right\}$ and two constants

$c \in]0,1[$ and $\alpha \in]0,1[$, such that for any $u \in \{0,\ldots,T_{j_1}\}$, $\varepsilon^{(u)} = (\varepsilon^{(u)}_{i,\mathbf{k}})_{i,\mathbf{k}} \in \{0,1\}^{3\times 2^{2j_1}}$, and any $(u,v) \in \{0,\ldots,T_{j_1}\}2$ with $u < v$, the following hold:

$$\sum_{i=1}^{3} \sum_{\mathbf{k}\in D_{j_1}} |\varepsilon^{(u)}_{i,\mathbf{k}} - \varepsilon^{(v)}_{i,\mathbf{k}}| \geq c2^{2j_1}, \qquad T_{j_1} \geq e^{\alpha 2^{2j_1}}.$$

Considering such a set E_{j_1}, for any $(u,v) \in \{0,\ldots,T_{j_1}\}^2$ with $u \neq v$, we have

$$\begin{aligned}
&\left(\int_{[0,1]^2} \left(h_{\varepsilon^{(u)}}(\mathbf{x}) - h_{\varepsilon^{(v)}}(\mathbf{x})\right)^2 d\mathbf{x}\right)^{1/2} \\
&= M_* 2^{-j_1(s+1)} \left(\sum_{i=1}^{3} \sum_{\mathbf{k}\in D_{j_1}} \left(\varepsilon^{(u)}_{i,\mathbf{k}} - \varepsilon^{(v)}_{i,\mathbf{k}}\right)^2\right)^{1/2} \\
&= M_* 2^{-j_1(s+1)} \left(\sum_{i=1}^{3} \sum_{\mathbf{k}\in D_{j_1}} \left|\varepsilon^{(u)}_{i,\mathbf{k}} - \varepsilon^{(v)}_{i,\mathbf{k}}\right|\right)^{1/2} \geq 2\nu_{j_1},
\end{aligned}$$

where

$$\nu_{j_1} = M_* c^{1/2} 2^{j_1} 2^{-j_1(s+1)} = M_* c^{1/2} 2^{-j_1 s}.$$

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

$$\nu_{j_1}^{-2} \sup_{f\in \mathbf{B}^s_{p,q}(M)} \mathbb{E}\left(\int_{[0,1]^2} \left(\widetilde{f}(\mathbf{x}) - f(\mathbf{x})\right)^2 d\mathbf{x}\right) \geq \sup_{u\in\{0,\ldots,T_{j_1}\}} \mathbb{P}_{h_{\varepsilon^{(u)}}}(A_u^c) = p,$$

where

$$A_u = \left\{\left(\int_{[0,1]^2} \left(\widetilde{f}(\mathbf{x}) - h_{\varepsilon^{(u)}}(\mathbf{x})\right)^2 d\mathbf{x}\right)^{1/2} < \nu_{j_1}\right\}$$

and $\mathbb{P}_f$ is the distribution of the model. Notice that, for any $(u,v) \in \{0,\ldots,T_{j_1}\}^2$ with $u \neq v$, $A_u \cap A_v = \emptyset$. Lemma 7.1 applied to the probability measures $\left(\mathbb{P}_{h_{\varepsilon^{(u)}}}\right)_{u\in\{0,\ldots,T_{j_1}\}}$ gives

$$p \geq \min\left(2^{-1}, \exp(-3e^{-1})\sqrt{T_{j_1}}\exp(-\chi_{T_{j_1}})\right), \tag{7.11}$$

where

$$\chi_{T_{j_1}} = \inf_{v\in\{0,\ldots,T_{j_1}\}} \frac{1}{T_{j_1}} \sum_{\substack{u\in\{0,\ldots,T_{j_1}\}\\ u\neq v}} K\left(\mathbb{P}_{h_{\varepsilon^{(u)}}}, \mathbb{P}_{h_{\varepsilon^{(v)}}}\right).$$

Let us now bound $\chi_{T_{j_1}}$. For any functions f_1 and f_2 in $\mathbb{L}_2([0,1]^2)$, we have

$$\begin{aligned}
K\left(\mathbb{P}_{f_1}, \mathbb{P}_{f_2}\right) &= \frac{1}{2\sigma^2}\int_{[0,1]^2} \left(T(f_1)(\mathbf{x}) - T(f_2)(\mathbf{x})\right)^2 d\mathbf{x} \\
&= \frac{1}{2\sigma^2}\int_{[0,1]^2} \left(\left((f_1 - f_2) \star g\right)(\mathbf{x})\right)^2 d\mathbf{x}.
\end{aligned}$$

The Plancherel formula yields

$$\begin{aligned} K\left(\mathbb{P}_{f_1},\mathbb{P}_{f_2}\right) &= \frac{1}{2\sigma^2}\int_{\mathbb{R}^2}|\mathcal{F}((f_1-f_2)\star g)(\mathbf{l})|^2 d\mathbf{l} \\ &= \frac{1}{2\sigma^2}\int_{\mathbb{R}^2}|\mathcal{F}(f_1-f_2)(\mathbf{l})|^2\,|\mathcal{F}(g)(\mathbf{l})|^2 d\mathbf{l}. \end{aligned}$$

Consequently, for any $(u,v)\in\{0,\dots,T_{j_1}\}^2$ with $u\neq v$, we have

$$K\left(\mathbb{P}_{h_{\varepsilon^{(u)}}},\mathbb{P}_{h_{\varepsilon^{(v)}}}\right)=\frac{1}{2\sigma^2}\int_{\mathbb{R}^2}\left|\mathcal{F}\left(h_{\varepsilon^{(u)}}-h_{\varepsilon^{(v)}}\right)(\mathbf{l})\right|^2|\mathcal{F}(g)(\mathbf{l})|^2 d\mathbf{l}. \tag{7.12}$$

On the other hand, for any $\mathbf{l}\in\mathbb{R}^2$, by definition of h_ε, we have

$$\begin{aligned} &\mathcal{F}(h_{\varepsilon^{(u)}}-h_{\varepsilon^{(v)}})(\mathbf{l}) \\ &\quad= M_* 2^{-j_1(s+1)}\sum_{i=1}^{3}\sum_{\mathbf{k}\in D_{j_1}}\left(\varepsilon_{i,\mathbf{k}}^{(u)}-\varepsilon_{i,\mathbf{k}}^{(v)}\right)\mathcal{F}\left(\Psi_{j_1,i,\mathbf{k}}\right)(\mathbf{l}). \end{aligned} \tag{7.13}$$

Equalities (7.12) and (7.13) then imply that

$$\begin{aligned} &K\left(\mathbb{P}_{h_{\varepsilon^{(u)}}},\mathbb{P}_{h_{\varepsilon^{(v)}}}\right) \\ &\quad= C\frac{1}{\sigma^2}2^{-2j_1(s+1)}\int_{\mathbb{R}^2}\left|\sum_{i=1}^{3}\sum_{\mathbf{k}\in D_{j_1}}\left(\varepsilon_{k}^{(u)}-\varepsilon_{k}^{(v)}\right)\mathcal{F}\left(\Psi_{j_1,i,\mathbf{k}}\right)(\mathbf{l})\right|^2|\mathcal{F}(g)(\mathbf{l})|^2 d\mathbf{l}. \end{aligned} \tag{7.14}$$

Let $\mathcal{C}_{j,i}=\operatorname{supp}(\Psi_{j,i,\mathbf{k}})$. Then by definition of the wavelet basis and under assumption (3.1), there exists a constant $C>0$ such that

$$\sup_{\mathbf{l}\in\mathcal{C}_{j_1,1}\cup\mathcal{C}_{j_1,2}\cup\mathcal{C}_{j_1,3}}|\mathcal{F}(g)(\mathbf{l})|^2\le C2^{-2j_1(\delta_1+\delta_2)}. \tag{7.15}$$

Moreover, the Plancherel formula implies that

$$\begin{aligned} &\int_{\mathbb{R}^2}\left|\sum_{i=1}^{3}\sum_{\mathbf{k}\in D_{j_1}}\left(\varepsilon_{i,\mathbf{k}}^{(u)}-\varepsilon_{i,\mathbf{k}}^{(v)}\right)\mathcal{F}\left(\Psi_{j_1,i,\mathbf{k}}\right)(\mathbf{l})\right|^2 d\mathbf{l} \\ &\quad= \int_{\mathbb{R}^2}\left|\mathcal{F}\left(\sum_{i=1}^{3}\sum_{\mathbf{k}\in D_{j_1}}\left(\varepsilon_{i,\mathbf{k}}^{(u)}-\varepsilon_{i,\mathbf{k}}^{(v)}\right)\Psi_{j_1,i,\mathbf{k}}\right)(\mathbf{l})\right|^2 d\mathbf{l} \\ &\quad= \int_{[0,1]^2}\left|\sum_{i=1}^{3}\sum_{\mathbf{k}\in D_{j_1}}\left(\varepsilon_{i,\mathbf{k}}^{(u)}-\varepsilon_{i,\mathbf{k}}^{(v)}\right)\Psi_{j_1,i,\mathbf{k}}(\mathbf{x})\right|^2 d\mathbf{x}=\sum_{i=1}^{3}\sum_{\mathbf{k}\in D_{j_1}}\left(\varepsilon_{i,\mathbf{k}}^{(u)}-\varepsilon_{i,\mathbf{k}}^{(v)}\right)^2 \\ &\quad\le C2^{2j_1}. \end{aligned} \tag{7.16}$$

The desired bound on $K\big(\mathbb{P}_{h_{\varepsilon(u)}},\mathbb{P}_{h_{\varepsilon(v)}}\big)$ then follows from (7.14), (7.15) and (7.16):

$$K\left(\mathbb{P}_{h_{\varepsilon(u)}},\mathbb{P}_{h_{\varepsilon(v)}}\right)\leq C\frac{1}{\sigma^2}2^{-2j_1(s+1)}2^{-2j_1(\delta_1+\delta_2)}2^{2j_1}=C\frac{1}{\sigma^2}2^{-2j_1(s+\delta_1+\delta_2+1)}2^{2j_1}.$$

Hence

$$\begin{aligned}\chi_{T_{j_1}} &= \inf_{v\in\{0,\ldots,T_{j_1}\}}\frac{1}{T_{j_1}}\sum_{\substack{u\in\{0,\ldots,T_{j_1}\}\\ u\neq v}}K\left(\mathbb{P}_{h_{\varepsilon(u)}},\mathbb{P}_{h_{\varepsilon(v)}}\right)\\ &\leq C\frac{1}{\sigma^2}2^{-2j_1(s+\delta_1+\delta_2+1)}2^{2j_1}. \qquad (7.17)\end{aligned}$$

Bringing (7.11) and (7.17) together and choosing j_1 such that

$$2^{-j_1(s+\delta_1+\delta_2+1)}=c_0\sigma,$$

where c_0 denotes a well chosen constant such that for any estimator $\widetilde{f}$ of f, we have

$$\begin{aligned}\nu_{j_1}^{-2}\sup_{f\in\mathbf{B}^s_{p,q}(M)}\mathbb{E}\left(\int_{[0,1]^2}\left(\widetilde{f}(\mathbf{x})-f(\mathbf{x})\right)^2d\mathbf{x}\right) &\geq c\exp\left((\alpha/2)2^{2j_1}-Cc_0^2 2^{2j_1}\right)\\ &\geq c,\end{aligned}$$

where

$$\nu_{j_1}=c2^{-j_1 s}=c\sigma^{s/(s+\delta_1+\delta_2+1)}.$$

This completes the proof.

□

Acknowledgment

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

References

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

[2] R. J. Adler. *An introduction to continuity, extrema, and related topics for general Gaussian processes.* Institute of Mathematical Statistics, Hayward, CA, 1990. MR1088478

[3] T. Cai. On adaptive wavelet estimation of a derivative and other related linear inverse problems. *J. Statistical Planning and Inference*, 108:329–349, 2002. MR1947406

[4] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs. A variational formulation for frame-based inverse problems. *Inv. Prob.*, 23:1495–1518, 2007. MR2348078

[5] C. Chesneau. Wavelet estimation via block thresholding: A minimax study under the L_p risk. *Statistica Sinica*, 18(3):1007–1024, 2008. MR2440401
[6] 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. MR2563260
[7] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57:1413–1541, 2004. MR2077704
[8] R. Devore, G. Kerkyacharian, D. Picard, and V. Temlyakov. On mathematical methods of learning. *Foundations of Computational Mathematics*, 1:3–58, 2006. MR2198214
[9] D. Donoho and M. Raimondo. A fast wavelet algorithm for image deblurring. *The Australian & New Zealand Industrial and Applied Mathematics Journal*, 46:C29–C46, 2005. MR2182159
[10] D.L. Donoho. Nonlinear solution of inverse problems by wavelet-vaguelette decomposition. *Applied and Computational Harmonic Analysis*, 2:101–126, 1995. MR1325535
[11] M. J. Fadili and J.-L. Starck. Sparse representation-based image deconvolution by iterative thresholding. In *ADA IV*, France, 2006. Elsevier.
[12] M. Figueiredo and R. Nowak. An EM algorithm for wavelet-based image restoration. *ITIP*, 12(8):906–916, 2003. MR2008658
[13] I.M. Johnstone, G. Kerkyacharian, D. Picard, and M. Raimondo. Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society. Series B. Methodological*, 66:547–573, 2004. MR2088290
[14] J. Kalifa, S. Mallat, and B. Rougé. Image deconvolution in mirror wavelet bases. In *IEEE ICIP*, volume 1, pages 565–569, 1998.
[15] S. Mallat. *A wavelet tour of signal processing*. Academic Press, 2nd edition, 1998. MR1614527
[16] Y. Meyer. *Wavelet and Operators*. Cambridge University Press, 1992. MR1228209
[17] R. Neelamani, H. Choi, and R. Baraniuk. Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Transactions on signal processing*, 52:418–433, 2004. MR2044455
[18] J.-C. Pesquet, A. Benazza-Benyahia, and C. Chaux. A sure approach for digital signal/image deconvolution problems. *IEEE Transactions on Image Processing*, 57(12):4616–4632, 2009.
[19] A.B. Tsybakov. *Introduction à l'estimation nonparametrique*. Springer, New York, 2004. MR2013911
[20] C. Vonesh, S. Ramani, and M. Unser. Recursive risk estimation for non-linear image deconvolution with a wavelet-domain sparsity constraint. In *IEEE International Conference on Image Processing, ICIP'08*, pages 665–668, San Diego, CA, October 2008.
[21] Wavelab 802. Wavelab toolbox. `http://www-stat.stanford.edu/~wavelab`, 2001.