Statistica Sinica **17**(2007), 317-340

ADAPTIVE BOXCAR DECONVOLUTION ON FULL LEBESGUE MEASURE SETS

Gérard Kerkyacharian, Dominique Picard and Marc Raimondo

Université de Paris X-Nanterre, Universités de Paris VII and University of Sydney

Abstract: We consider the non-parametric estimation of a function that is observed in white noise after convolution with a boxcar, the indicator of an interval $(-a, a)$. In a recent paper Johnstone, Kerkyacharian, Picard and Raimondo (2004) have developed a wavelet deconvolution method (called `WaveD`) that can be used for "certain" boxcar kernels. For example, `WaveD` can be tuned to achieve near optimal rates over Besov spaces when a is a Badly Approximable (BA) irrational number. While the set of all BA's contains quadratic irrationals, e.g., $a = \sqrt{5}$, it has Lebesgue measure zero. In this paper we derive two tuning scenarios of `WaveD` that are valid for "almost all" boxcar convolutions (i.e., when $a \in A$ where A is a full Lebesgue measure set). We propose (i) a tuning inspired from Minimax theory over Besov spaces; (ii) a tuning derived from Maxiset theory providing similar rates as for WaveD in the BA widths setting. Asymptotic theory finds that (i) in the worst case scenario, departures from the BA assumption affect `WaveD` convergence rates, at most, by log factors; (ii) the Maxiset tuning, which yields smaller thresholds, is superior to the Minimax tuning over a whole range of Besov sub-scales. Our asymptotic results are illustrated in an extensive simulation of boxcar convolution observed in white noise.

Key words and phrases: Adaptive estimation, boxcar, deconvolution, non-parametric regression, Meyer wavelet.

1. Introduction

Suppose we observe the stochastic process

$$Y_n(dt) = f \star b(t)dt + \sigma n^{-1/2} W(dt), \quad t \in T = [0, 1], \tag{1}$$

where $b(t) = (2a)^{-1}\mathbb{I}\{|t| \leq a\}$, σ is a positive constant, $W(\cdot)$ is a Gaussian white noise and

$$f \star b(t) = \frac{1}{2a} \int_{-a}^{a} f(t-u)du \,. \tag{2}$$

Our goal is to recover the function f from the noisy blurred observations (1). The boxcar half-width a in (2) is assumed to be known. Further, we assume that the function f is periodic on the unit interval T. This is the so-called boxcar deconvolution problem. It is an important model for the problem of recovery of noisy signals (or images) in linear motion blur, see Bertero and Boccacci (1998).

Over the last decade, many wavelet methods have been developed to recover f from indirect observations: see Donoho (1995), Abramovich and Silverman (1998), Pensky and Vidakovic (1999), Walter and Shen (1999), Johnstone (1999), Fan and Koo (2002) and Kalifa and Mallat (2003). However, the boxcar assumption (2) escapes most of the previously cited works. More recent papers which deal specifically with the boxcar problem includes Hall, Ruymgaart, van Gaans and van Rooij (2001), Neelamani, Choi and Baraniuk (2004), Johnstone and Raimondo (2004) and Johnstone, Kerkyacharian, Picard and Raimondo (2004) ([JKPR] in the sequel). In the periodic model the boxcar convolution has the special feature that if the boxcar half-width is rational then certain frequencies are lost, Johnstone and Raimondo (2004). If one changes the observation model and imposes a compact support assumption on the signal f, as in Hall, Ruymgaart, van Gaans and van Rooij (2001), the latter difficulty disappears. Another approach to deal with the boxcar scenario is to introduce a regularisation step in the wavelet reconstruction algorithm, as in Neelamani, Choi and Baraniuk (2004). In this paper we follow the approach of [JKPR] and consider boxcar convolutions (2) where the boxcar half-width a is an irrational number. If a is chosen among Badly Approximable (BA) irrational numbers (these contain quadratic irrational like $\sqrt{5}$) then the WaveD method is near optimal for a wide range of target functions and error losses. In the finite sample implementation of the model (1) on a regular grid $t_i = i/n$, $i = 1, \dots, n$, the WaveD estimator can recover the unknown function f with an accuracy of order

$$\left(\frac{\log n}{n}\right)^{\beta}, \quad \text{where} \tag{3}$$

$$\beta = \frac{sp}{4+2s}, \quad \text{if } s \geq \frac{2p}{\pi} - 2, \tag{4}$$

$$\beta = \frac{(s - \frac{1}{\pi} + \frac{1}{p})p}{4 + 2(s - \frac{1}{\pi})}, \quad \text{if } \frac{1}{\pi} \leq s < \frac{2p}{\pi} - 2, \tag{5}$$

performance being measured in an integrated L^p-metric, for any $p > 1$. Here n denotes the usual sample size and s plays the role of a smoothness index for our target function f.

In fact near optimal properties of WaveD hold for any kernel function $b_a(t)$ whose Fourier coefficients $(b_l(a))$ satisfy a decay condition when averaged over dyadic blocks. Let

$$\tau_j^2(a) = |C_j|^{-1} \sum_{l=2^j}^{2^j+r} |b_l(a)|^{-2} \,. \tag{6}$$

Then for any kernel function $b_a(t)$ such that $\tau_j^2(a) \le 2^{3j}$ (as in the BA case), the rate result (3) holds.

For statistical applications, an important issue is whether the WaveD estimator is robust against departures from the BA assumption. In this paper we combine the Maxiset Theorem of Kerkyacharian and Picard (2000) with the Equidistribution Lemma of Johnstone and Raimondo (2004) to extend the results of [JKPR] outside the BA assumption. We propose two tuning scenarios of WaveD that can be applied to "almost all" boxcar convolutions (i.e., $a \in A$ where A is a full Lebesgue measure set). In the first scenario we show that departures from the BA assumption affect the rate (3) by, at most, log factors when tuning the WaveD method over standard Besov spaces. In the second scenario we show that the WaveD estimator can achieve the rate (3) on certain Besov sub-scales without the extra log-penalty. A theoretical comparison of the two scenarios suggests that smaller thresholds (such as arise in the second scenario) will always give better results than larger thresholds (such as arise in the first scenario). This is confirmed by an extensive simulation study of boxcar convolutions observed in white noise. All figures and tables presented in this paper can be reproduced using the `WaveD1.3` software package available at http://www.maths.usyd.edu.au/u/marcr/.

An illustration of the model (1) is given in Figures 2 and 3 using the four test functions depicted in Figure 1. In Figure 2 we took $\sigma = 0$ to illustrate the effect of the boxcar convolution when the boxcar-width is chosen at random according to a uniform distribution. In Figure 3 we took $\sigma = 0.05$ to illustrate the model

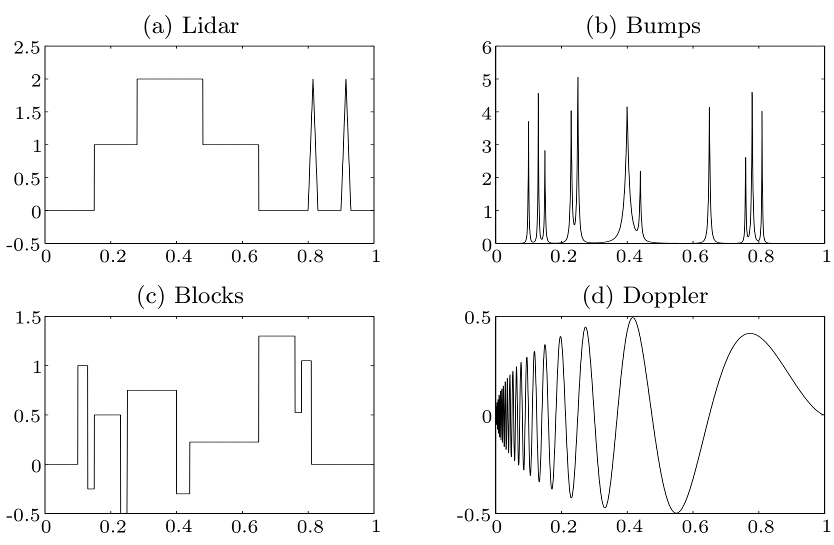

Figure 1. Four (inhomogeneous) signals, $t_i = i/n$, $n = 2,048$.

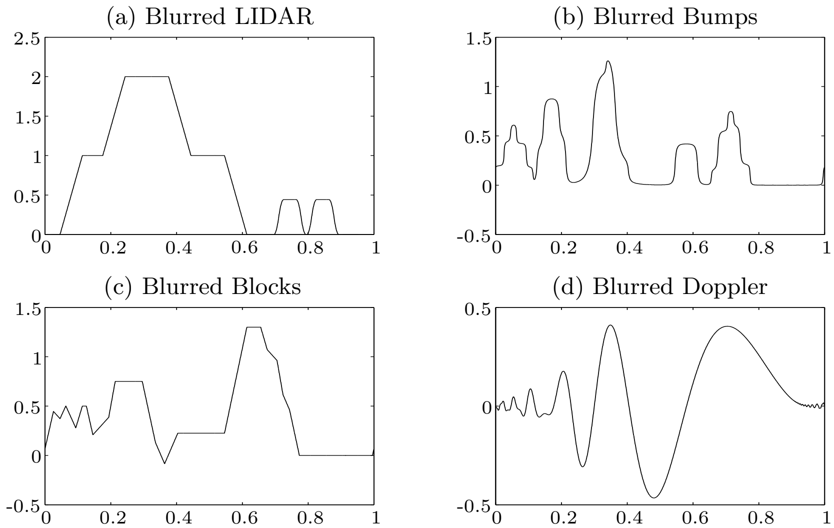

Figure 2. Signals of Figure 1 after boxcar blurring with a randomly chosen from a $\mathcal{U}(0.025, 075)$ distribution.

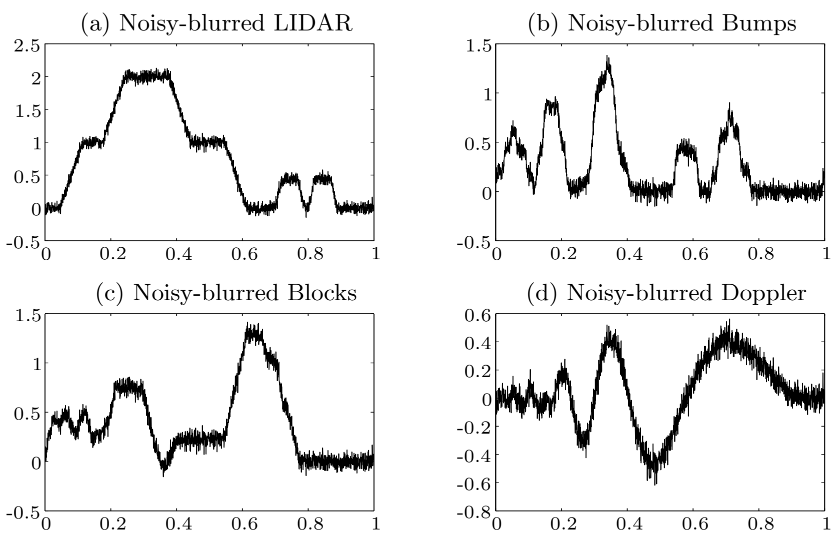

Figure 3. Illustration of the model (1): signals of Figure 2 in Gaussian noise with $\sigma = 0.05$.

(1) in medium noise level. In Figure 4, we depicted the corresponding WaveD estimates based on the Maxiset tuning of Section 3.2.

We begin in Section 2 with preliminaries on `WaveD` estimation, Besov spaces and the Maxiset Theorem. Asymptotic results are summarised in Section 3 and numerical performances are studied in Section 4. Proofs are given in Section 5.

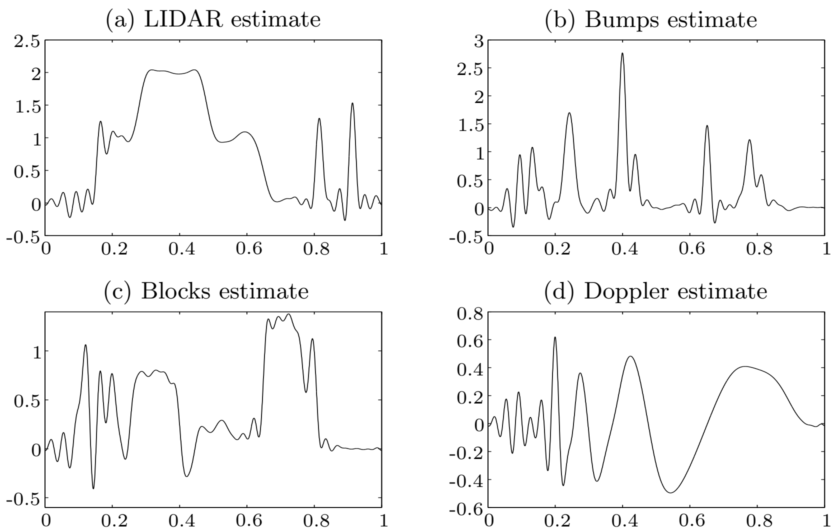

Figure 4. WaveD estimates based on the noisy data of Figure 3.

2. Preliminaries

2.1. Wavelet Deconvolution in a periodic setting

The wavelet deconvolution (WaveD) method proposed by [JKPR] combines both Fourier and Wavelet analysis. Let Φ, Ψ denote the (periodised) Meyer scaling and wavelet function, see e.g., Meyer (1990), Mallat (1998). Let $e_l(t) = e^{2\pi i l t}$, $l \in \mathbb{Z}$, and write $f_l = \langle f, e_l \rangle, b_l = \langle b, e_l \rangle$ for the Fourier coefficients of f, b, respectively, where $\langle f, g \rangle = \int_T f \bar{g}$. The WaveD estimator is based on hard thresholding of a wavelet expansion as follows (here and in the sequel κ will denote the multiple index (j, k) and $\Psi_{-1} = \Phi$):

$$\hat{f}_n = \sum_{\kappa \in \Lambda_n} \hat{\beta}_\kappa \, \Psi_\kappa \, \mathbb{I}\{|\hat{\beta}_\kappa| \geq \hat{\sigma} \eta \, \sigma_j \, c_n\} \,. \tag{7}$$

The wavelet coefficients are computed in the Fourier domain as

$$\hat{\beta}_\kappa = \sum_{l \in C_j} \left(\frac{y_l}{b_l} \right) \bar{\Psi}_l^\kappa \,, \tag{8}$$

using eigen values of the boxcar function,

$$b_l = \frac{\sin \pi a l}{\pi a l}, l \in \mathbf{Z}. \tag{9}$$

Throughout this paper we assume that a is an irrational number so that there are no zeros in (9). We use the fast algorithm of Donoho and Raimondo (2004) to compute the wavelet transform (8) and its inverse (7). This algorithm takes full advantage of the compact support of the Meyer wavelet in the Fourier domain: $C_j = \{l : \Psi_l^\kappa \neq 0\} \subset (2\pi/3)\cdot[-2^{j+2}, -2^j] \bigcup [2^j, 2^{j+2}]$.

The tuning parameters of WaveD are as follows.

- The range of resolution levels (frequencies) where the approximation (7) is used, $\Lambda_n = \{(j,k),\ -1 \le j \le j_1, 0 \le k \le 2^j\}$. Here j_1 determines the highest resolution level of WaveD. Theoretical properties of j_1 are given in Section 3. In software, the default value of the finest scale j_1 is determined from the data: j_1 is set to be the level preceding $j(100\%)$ where $j(100\%)$ is the smallest level where 100% of thresholding occurs. An alternative data-driven method for chosing j_1 is described in Cavalier and Raimondo (2007).
- The threshold value has four input parameters $\hat{\sigma}\,, \eta\,, \sigma_j$ and c_n, prescribed below.
  - $\hat{\sigma}$: an estimate of the noise standard deviation, σ. If $y_{J,k} = \langle Y_n, \Psi_{J,k}\rangle$, denote the finest scale wavelet coefficients of the observed data, then $\hat{\sigma} = m.a.d.\{y_{J,k}\}/0.6745$, where $m.a.d.$ is median absolute deviation.
  - η: a constant which depends on the tail of the noise distribution. For Gaussian noise, the range $\sqrt{2} \le \eta \le \sqrt{6}$ gives good results in practice. In software, the default value is $\sqrt{6}$. Theoretical properties of η are given in Section 3.
  - σ_j: a level-dependent scaling factor based on the convolution kernel. Theoretical properties of σ_j are given in Sections 2 and 3. In software the (standard) default value is

$$\sigma_j^2 := \tau_j^2(a) = |C_j|^{-1} \sum_{l\in C_j} |b_l(a)|^{-2}.$$

  - c_n: a sample size-dependent scaling factor. Theoretical properties of c_n are given in Section 3. In software the default value is $c_n = (\log n/n)^{1/2}$.

2.2. Besov spaces of periodic functions

Let us first introduce the standard Besov spaces of periodic functions $B^s_{\pi,r}(T)$, $s > 0$, $\pi \ge 1$ and $r \ge 1$. Define for every measurable function f, $\Delta_\varepsilon f(x) = f(x+\varepsilon) - f(x)$. Then, recursively, take $\Delta_\varepsilon^2 f(x) = \Delta_\varepsilon(\Delta_\varepsilon f)(x)$, and similarly $\Delta_\varepsilon^N f(x)$ for positive integer N. Let

$$\rho^N(t,f,\pi) = \sup_{|\varepsilon|\le t}\Big(\int_0^1 |\Delta_\varepsilon^N f(u)|^\pi du\Big)^{\frac{1}{\pi}}.$$

Then for $N > s$, we define

$$B^s_{\pi,r}(T) = \{f \text{ periodic: } \Big(\int_0^1 \Big(\frac{\rho^N(t,f,\pi)}{t^s}\Big)^r \frac{dt}{t}\Big)^{\frac{1}{r}} < \infty\}.$$

(with the usual modifications for r or $\pi = \infty$.)

In this setting, we note that the Besov spaces are characterised by the behaviour of the wavelet coefficients (as soon as the wavelet is periodic and has enough smoothness and vanishing moments).

Definition 1. For $f \in L^\pi(T)$,

$$f = \sum_{j,k} \beta_{j,k} \Psi_{j,k} \in B^s_{\pi,r}(T) \iff \sum_{j \geq 0} 2^{j(s+1/2-1/\pi)r} \Big[\sum_{0 \leq k \leq 2^j} |\beta_{j,k}|^\pi \Big]^{r/\pi} < \infty. \tag{10}$$

The Besov spaces have proved to be an interesting scale for studying the properties of statistical procedures. The index s indicates the degree of smoothness of the function. Due to the differential averaging effects of the integration parameters π and r, the Besov spaces capture a variety of smoothness features in a function including spatially inhomogeneous behaviour, see Donoho et al. (1995).

In order to fully describe asymptotic properties of WaveD for boxcar deconvolution it is useful to introduce the following Besov sub-scales.

Definition 2. For $s > 0,\ 1 \leq \pi \leq \infty,\ \tau \in \mathbb{R}$, define

$$\tilde{B}^{s,\tau}_{\pi,\infty}(T) = \Big\{ f = \sum_{j \geq -1, k} \beta_{jk} \Psi_{j,k},\ \sup |j|^\tau 2^{j(s+\frac{1}{2}-\frac{1}{\pi})} \|(\beta_{jk})_k\|_{l_\pi} < \infty \Big\}. \tag{11}$$

The latter range of Besov scales are embedded into standard Besov scales:

$$\tilde{B}^{s,\tau}_{\pi,\infty}(T) \subset B^s_{\pi,\infty}(T), \quad \text{for all } \tau \geq 0.$$

2.3. The Maxiset-approach

The following theorem is borrowed from Kerkyacharian and Picard (2000). We refer to the appendix for condition (51) (known as the Temlyakov property). First, we introduce some notation: μ will denote the measure such that for $j \in \mathbb{N},\ k \in \mathbb{N}$,

$$\mu\{(j,k)\} = \|\sigma_j \Psi_{j,k}\|_p^p = \sigma_j^p 2^{j(\frac{p}{2}-1)} \|\Psi\|_p^p\,, \tag{12}$$

$$l_{q,\infty}(\mu) = \Big\{ f,\ \sup_{\lambda > 0} \lambda^q \mu\{(j,k):\ |\beta_{j,k}| > \sigma_j \lambda\} < \infty \Big\}. \tag{13}$$

Theorem 1. *Let $p > 1,\ 0 < q < p,\ \{\ \psi_{j,k}, j \geq -1, k = 0, 1, \ldots, 2^j\}$ be a periodised wavelet basis of $L^2(T)$ and σ_j be a positive sequence such that the* *heteroscedastic basis $\sigma_j\psi_{j,k}$ satisfies* (51). *Suppose that Λ_n is a set of pairs (j,k) and c_n is a deterministic sequence tending to zero with*

$$\sup_n \mu\{\Lambda_n\}\, c_n^p < \infty. \tag{14}$$

Suppose for any n and any pair $\kappa = (j,k) \in \Lambda_n$, we have

$$\mathbb{E}|\hat{\beta}_\kappa - \beta_\kappa|^{2p} \le C\,(\sigma_j\, c_n)^{2p}\,, \tag{15}$$

$$P\Big(|\hat{\beta}_\kappa - \beta_\kappa| \ge \eta\,\sigma_j\, c_n/2\Big) \le C\,(c_n^{2p} \wedge c_n^4)\,, \tag{16}$$

for some positive constants η and C. Then the wavelet based estimator

$$\hat{f}_n = \sum_{\kappa\in\Lambda_n} \hat{\beta}_\kappa\,\psi_\kappa\,\mathbb{I}\{|\hat{\beta}_\kappa| \ge\, \eta\,\sigma_j\, c_n\} \tag{17}$$

is such that, for all positive integers n, $\mathbb{E}\|\hat{f}_n - f\|_p^p \le C\, c_n^{p-q}$, if and only if

$$f \in l_{q,\infty}(\mu), \quad \textit{and} \tag{18}$$

$$\sup_n c_n^{q-p}\|f - \sum_{\kappa\in\Lambda_n} \beta_\kappa\psi_\kappa\|_p^p < \infty. \tag{19}$$

This Theorem identifies the 'Maxiset' of a general wavelet estimator of the form (17), by conditions (18) and (19). In [JKPR] p.565, we see that the estimated wavelet coefficients $(\hat{\beta}_\kappa)$ defined at (8) are unbiased, normally distributed with variances bounded by the quantity $\tau_j^2(a)$ defined at (6). Clearly (15) and (16) rely heavily on the precise evaluation of this quantity. In the next section we show that for *almost all a*, $\tau_j^2(a) = O(2^{3j} j^{11(1+\delta)})$. Hence, for such τ_j, the proof arguments of [JKPR] that hold for $\tau_j^2 \le 2^{3j}$ no longer apply. In fact, an improvement of Theorem 1 is required to derive the WaveD asymptotic theory for *almost all a*. The following corollary is an adaptation of the proof of Theorem 1, Kerkyacharian and Picard (2000).

Corollary 1. *Let $0 < q < \infty$, $\mathcal{H}237 - \infty < \alpha < \infty$. Let $\xi(t)$ be a continuous non-decreasing function such that $\xi(0) = 0$,*

$$\xi(t) = \xi_{(q,\alpha)}(t) := \begin{cases} t^q(\log(\frac{1}{t}))^\alpha, & t \in [0,\zeta] \\ (\log(\frac{1}{\zeta}))^\alpha t^q, & t > \zeta, \end{cases} \tag{20}$$

where $0 < \zeta \le \exp -\alpha/q$ if $\alpha \ge 0$, and $0 < \zeta < 1$ if $\alpha < 0$. Under the hypotheses of Theorem 1, the estimator $\hat{f}_n$ is such that, for all positive integers n, $\mathbb{E}\|\hat{f}_n - f\|_p^p \le C\,(c_n^p/\xi(c_n))$ if and only if

$$f \in l_{\xi,\infty}(\mu), \quad \textit{and} \tag{21}$$

$$\sup_n \frac{c_n^p}{\xi(c_n)}\Big\|f - \sum_{\kappa\in\Lambda_n} \beta_\kappa\psi_\kappa\Big\|_p^p < \infty\,, \tag{22}$$

where

$$l_{\xi,\infty}(\mu) = \left\{ f,\ \sup_{\lambda>0} \xi(\lambda)\mu\{(j,k) : |\beta_{j,k}| > \sigma_j \lambda\} < \infty \right\}. \tag{23}$$

Remark 0. Corollary 1 offers more flexibility than Theorem 1, in particular, it allows us to deal with scaling factor σ_j of the form $\sigma_j \asymp 2^{\nu j} j^z$. While this has direct applications in the problem at hand, we note that there are other interesting applications of Corollary 1 such as in multichannel deconvolution, Pensky and Zayed (2002). See also the discussion by De Canditiis and Pensky in [JKPR].

2.4. Diophantine approximations

To every real number there corresponds a unique sequence (a_k):

$$a = [a_0; a_1, a_2, \ldots] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}. \tag{24}$$

For rational numbers the sequence is finite, $a = [a_0; a_1, a_2, \ldots, a_k]$, whereas for irrational a the sequence of (a_k), $a_k > 0$, is infinite. The rational numbers defined by terminating the expansion (24) at stage k: $p_k(a)/q_k(a) = [a_0, a_1, \ldots, a_k]$ are called the *convergents* of a. For any irrational a, the convergents have the property of *best approximation*: for $n \geq 1$,

$$\inf_{1 \leq k \leq q_n} ||ka|| = |q_n a - p_n| = ||q_n a||, \tag{25}$$

where $\|x\|$ denotes the distance from $x \in \mathbb{R}$ to the nearest integer. The study of such Diophantine approximations plays a central role in our analysis of the boxcar blur, since from (9) it follows that

$$\frac{2}{\pi} \frac{\|la\|}{la} \leq b_l \leq \frac{\|la\|}{la}. \tag{26}$$

Hence, the properties of the WaveD estimator (7) depend on the nature of the irrational number a.

Definition 3. An irrational number a is called *Badly Approximable* (BA) if $\sup_n q_n/q_{n-1} < \infty$.

For boxcar with BA widths, [JKPR] have shown that the WaveD estimator achieves near optimal rates of convergence over Besov spaces. Another interesting class of irrational numbers can be derived from the "measure theory" of continued fractions Khintchine (1963, Chap.3.)

Definition 4. For each $\delta > 0$, there is a set A_δ of full Lebesgue measure such that, for all $a \in A_\delta$ and $n \geq n(a)$,

$$\frac{q_n}{q_{n-1}} < (\log q_{n-1})^{(1+\delta)}. \tag{27}$$

In the sequel "almost all a" means"for all a in A_δ".

One of the main difficulties of boxcar deconvolution is that the Fourier coefficients $b_l(a), l = 0, 1, 2, \ldots$ vary erratically according to the approximations $\|la\|/l$, $l = 1, 2, \ldots$, see Figure 5. This difficulty disappears when averaging over dyadic blocks. In particular, if

$$\tau_j^2(a) = |C_j|^{-1} \sum_{l \in C_j} |b_l(a)|^{-2}, \tag{28}$$

where $C_j = \{l : \Psi_l^\kappa \neq 0\} \subset (2\pi/3) \cdot [-2^{j+2}, -2^j] \bigcup [2^j, 2^{j+2}]$, then we have the following bounds.

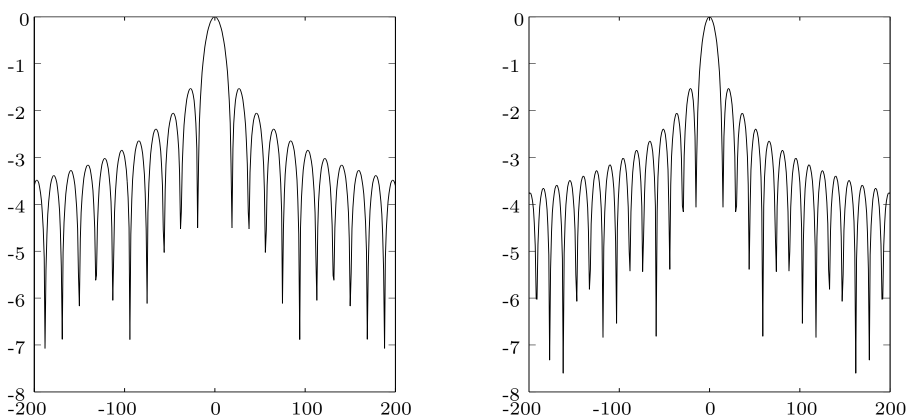

Figure 5. Eigen values of the boxcar function with: (left) BA scale $a = 1/\sqrt{353}$; (right) a randomly chosen from a $\mathcal{U}(0.025, 075)$ distribution.

Proposition 1. *Let (b_l) be the Fourier coefficients of the boxcar kernel $b(t) = (1/(2a))\mathbb{I}\{|t| \leq a\}$.*
1. *For BA boxcar scale a there is a constant $c > 0$ such that*

$$\tau_j^2(a) \leq c\, 2^{3j}. \tag{29}$$

2. *For each $\delta > 0$, there is a constant $c > 0$ such that for* almost all *a,*

$$\tau_j^2(a) \leq c\, 2^{3j}\, j^z,\ z = 11(1+\delta). \tag{30}$$

Remark 1. The bound (29) was derived in [JKPR] and is presented here for comparison purposes. The novelty appears in the bound (30) where we see that for almost all boxcar half-width a, the scaling factor $\tau_j(a)$ is slightly larger than in the BA case. In the next section we study the effect of these new scales on the asymptotic properties of WaveD.

3. Asymptotic Results

For boxcar convolutions observed in white noise we derive tuning parameters $\Lambda_n, \eta, \sigma_j$ and c_n which yield near optimal properties of WaveD for *almost all* a. Using the Maxiset approach we present two possible tuning scenarios of the waved estimator. Our first scenario (Section 3.1) is inspired from the Minimax theory: we fix a smoothness class (here $B^s_{\pi,r}(T)$) and find a set of tuning parameters which ensure near-optimal asymptotic properties of WaveD over the entire smoothness class. In our second scenario (Section 3.2) the primary focus is on the convergence rate rather than on the smoothness class: we propose a set of tuning parameters with which WaveD achieves the rate (3) known to be achievable for boxcar convolution with Badly Approximable width. Finally, we give a theoretical comparison of the two scenarios in Section 3.3.

3.1. Tuning inspired by Minimax theory

To study the properties of WaveD over standard Besov spaces, one must find conditions under which a particular Besov space $B^s_{\pi,r}(T)$ may be embedded in the space $l_{q,\infty}(\mu)$, as well as imply condition (19). We follow the steps detailed in Appendix B.1 of [JKPR], taking into account the Degree of Ill-Posedness (DIP) $\nu = 3/2$ of the Boxcar convolution, see Johnstone and Raimondo (2004). Recalling that the Maxisets $l_{q,\infty}(\mu)$ are defined in terms of the measure $\mu\{(j,k)\} = c\,\sigma_j^p 2^{j(p/2-1)}$, a key condition in recovering the standard Besov spaces is to let

$$\sigma_j^2 =: \sigma_j^2(1) = c_1\, 2^{3j}. \tag{31}$$

Letting as usual $\Lambda_n = \{(j,k),\ -1 \le j \le j_1, 0 \le k \le 2^j\}$ and $z = 11(1+\delta)$, we compensate for the extra log-term appearing in (30) by taking

$$c_n := c_n(1) = \left(\frac{\log n}{n}\right)^{\frac{1}{2}} (\log n)^{\frac{z}{2}} \text{ and } 2^{j_1} = n^{\frac{1}{4}} (\log n)^{-\frac{1+z}{4}}. \tag{32}$$

These choices ensure that (15) and (16) are satisfied.

Proposition 2. *Suppose we observe the random process* (1) *with* $\sigma = 1$. *Let* $p > 1$ *be an arbitrary number. If* f *belongs to* $B^s_{\pi,r}(T)$ *with* $\pi \ge 1$, $s \ge 1/\pi$ *and* $0 < r \le ((2p-1)/(2+s-1/\pi))$ *if* $s+2 = 2p/\pi$, $0 < r \le \infty$ *otherwise, then* *for* $\eta \geq 2\sqrt{8\pi(p \vee 2)}$ *and* $z = 11(1+\delta)$, *the WaveD estimator* (7) *with tuning parameters* (31), (32) *and* $\hat{\sigma} = 1$ *is such that, for almost all* a,

$$\mathbb{E}\|\hat{f}_1 - f\|_p^p \leq C[n^{-1}(\log(n))^{(1+z)}]^{\beta}, \text{ for all positive integers } n, \tag{33}$$

where β *depends on* s *as at* (4) *and* (5).

Remark 2. A comparison of the rates (3) and (33) shows that, for almost all boxcar convolutions (in the sense of Definition 4), the rate properties of WaveD are, in the worst-case-scenario, only affected up to log factors. These results are consistent with those derived in Johnstone and Raimondo (2004) regarding the Degree of Ill-Posedness of the boxcar deconvolution problem in the periodic setting.

3.2. Tuning inspired by Maxiset theory

Corollary 1 allows a certain degree of freedom in the tuning of WaveD. More specifically, the moment condition (15) and the tail behavior (16) may be satisfied for a different choice of σ_j, c_n than proposed in Section 3.1. For example, it is possible to fit WaveD with a slightly smaller threshold than in Proposition 2 and yet derive a consistent estimator. While this second scenario may not ensure optimal properties over the scale standard Besov spaces, it is interesting to note that it yields a faster rate of convergence over Besov sub-scales $\tilde{B}^{s,\tau}_{\pi,r}$. In fact, it is possible to tune WaveD in such a way that it achieves the rate (3) as available for boxcar convolution with BA scales. Following Appendix A.1 of [JKPR], we start by letting

$$c_n =: c_n(2) = \left(\frac{\log n}{n}\right)^{\frac{1}{2}} \text{ and } 2^{j_1} = n^{\frac{1}{4}}(\log n)^{-\frac{1+z}{4}}. \tag{34}$$

Now let $z = 11(1+\delta)$, and compensate for the extra log terms in (30) by taking

$$\sigma_j^2 =: \sigma_j^2(2) = c_2\, 2^{3j} j^z. \tag{35}$$

Proposition 3. *Suppose we observe the random process* (1) *with* $\sigma = 1$. *Let* $p > 1$ *be an arbitrary number. Suppose* f *belongs to* $\tilde{B}^{s,\tau}_{\pi,\infty}(T)$ *with* $\pi \geq 1$, $s \geq 1/\pi$, *and with* $\tau \geq zs/4$ *if* $s+2 > 2p/\pi$, $\tau > 1 + zs/4$ *if* $s+2 = 2p/\pi$, $\tau \in \mathbb{R}$ *otherwise. Then for* $\eta \geq 2\sqrt{8\pi(p \vee 2)}$ *and* $z = 11(1+\delta)$, *the WaveD estimator* (7) *with tuning parameters* (35), (34) *and* $\hat{\sigma} = 1$ *is such that, for almost all* a,

$$\mathbb{E}\|\hat{f}_2 - f\|_p^p \leq C[n^{-1}\log(n)]^{\beta}, \text{ for all positive integers } n, \tag{36}$$

where β *is given at* (4), (5).

Remark 3. A quick look at Propositions 2 and 3 shows that they do not give a fair comparison between the two scenarios. The reason is that both the spaces and the convergence rates with which each tuning method is prescribed are different. Since Besov sub-scales are embedded in standard Besov spaces ($\tilde{B}^{s,\tau}_{\pi,\infty} \subset B^{s}_{\pi,\infty}$) it is not surprising to observe better rate performances in the second scenario. A further application of Corollary 1 allows a fair comparison of the two scenarios (next section).

3.3. Minimax versus Maxisets

The following proposition describes in detail the differences between the two tuning scenarios over Besov sub-scales in terms of their respective Maxisets.

Proposition 4. *Suppose we observe the random process* (1) *with* $\sigma = 1$. *Let* $p > 1$ *be an arbitrary number, and* q *be such that* $0 < q < p$. *Let* $v_n = (n^{-1} \log n)^{(p-q)/2}$, j_1 *be chosen as in* (32), *and define*

$$\mathcal{F} = \Big\{ f : \quad \sup_n v_n \| f - \sum_{\kappa, j \leq j_1} \beta_\kappa \psi_\kappa \|_p^p < \infty \Big\}. \tag{37}$$

Let $\hat{f}_1$ *be the WaveD estimator* (7) *with tuning parameters* (31), (32) *and* $\hat{f}_2$ *the WaveD estimator* (7) *with tuning parameters* (35), (34). *Let* $\hat{\sigma} = 1$. *Then for almost all* a,

(a) $\mathbb{E}\|\hat{f}_1 - f\|_p^p \leq C v_n$ for all positive integers n if and only if $f \in MAX(1) = \{ f \in l_{\xi_{(q, \frac{z(p-q)}{2})}, \infty}(\mu) \quad \cap \mathcal{F} \}$, where

$$l_{\xi_{(q, \frac{z(p-q)}{2})}, \infty}(\mu) = \left\{ f,\ \sup_{\lambda > 0}\ \xi_{(q, \frac{z(p-q)}{2})}(\lambda)\ \mu\{(j,k) :\ |\beta_{j,k}| > \sigma_j(1)\lambda\} < \infty \right\}, \tag{38}$$

with $\mu\{(j,k)\} = 2^{j(2p-1)}$ and $\xi(t) = \xi_{(q,\alpha)}(t)$ defined at (20);

(b) $\mathbb{E}\|\hat{f}_2 - f\|_p^p \leq C\, v_n$ for all positive integers n if and only if $f \in MAX(2) = \{ f \in l_{q,\infty}(\tilde{\mu}) \quad \cap \mathcal{F} \}$, where

$$f \in l_{q,\infty}(\tilde{\mu}) = \left\{ f,\ \sup_{\lambda > 0} \lambda^q \tilde{\mu}\{(j,k) :\ |\beta_{j,k}| > \sigma_j(2)\lambda\} < \infty \right\}, \tag{39}$$

and $\tilde{\mu}\{(j,k)\} = 2^{j(2p-1)} j^{zp/2}$.

Remark 4. Proposition 4 shows that over Besov sub-scales both tuning scenarios yield the rate of convergence (3). The difference between the two methods appears in their respective Maxisets described by (38) and (39). Hence the two maxisets to be compared are $MAX(1)$ for minimax-WaveD and $MAX(2)$ for maxiset-WaveD.

Proposition 5. *Under the assumptions Proposition* 4*, we have* $MAX(1) \subset MAX(2)$.

Remark 5. This result shows that if we compare the tuning scenarios using the maxiset point of view, the second tuning (Maxiset) is always better than the first tuning (Minimax) since it achieves a near optimal rate of convergence over a larger class of functions. This suggests that the smaller threshold setting (34) and (35) may give better results in practice than the larger threshold setting (31) and (32). This is confirmed by our simulation study, see Section 4 and Table 1.

Table 1. Monte-Carlo approximations to RMISE=$\sqrt{\mathbb{E}\|\hat{f} - f\|_2^2}$. The results are means of 1,000 independent simulations of the model (1) with $n = 2,048$ illustrated as Figure 3.

Tuning	Boxcar-scale	Signal	$\sigma_{low} = 0.005$	$\sigma_{med} = 0.05$	$\sigma_{high} = 0.1$
Standard	BA	Lidar	0.0990	0.2084	0.2744
Maxiset	AA	Lidar	0.1220	0.2671	0.3400
Minimax	AA	Lidar	0.1379	0.3274	0.3863
Standard	BA	Bumps	0.2042	0.4563	0.5233
Maxiset	AA	Bumps	0.2933	0.5103	0.5466
Minimax	AA	Bumps	0.3231	0.5332	0.5749
Standard	BA	Blocks	0.1207	0.2287	0.2676
Maxiset	AA	Blocks	0.1469	0.2643	0.3097
Minimax	AA	Blocks	0.1626	0.3044	0.3324
Standard	BA	Doppler	0.0601	0.1063	0.1372
Maxiset	AA	Doppler	0.0681	0.1346	0.1605
Minimax	AA	Doppler	0.0754	0.1572	0.1872

4. Numerical Performances

We study the finite sample properties of the WaveD algorithm (7) when applied to the noisy boxcar convolution (1). Figure 1 depicts four inhomogeneous signals borrowed from the statistical literature (Johnstone et al. (2004) and Donoho and Raimondo (2004)). Figure 2 depicts the signals of Figure 1 after blurring with a boxcar kernel. In Figure 3, we added Gaussian white noise to each signal of Figure 2 (with medium noise level). In our simulation study we used three noise levels: low, medium and high as seen Table 1. Our main results (summarised in Section 3) state that the WaveD estimator can be applied to noisy boxcar convolution where the boxcar scale a is chosen randomly with respect to a continuous distribution. An illustration of the WaveD estimator in such a setting is given in Figure 4. For illustration purposes, we used the Uniform $(0.025, 0.075)$ distribution to set the boxcar parameter in each simulation of (1), as illustrated in Figures 2 and 3. In Table 1 we give Monte Carlo approximation to the Root Mean Integrated Square Error (RMISE) of the WaveD estimator when fitted with the Minimax tuning (32) and (31) and when fitted with the Maxiset tuning (34) and (35). For comparison purposes we included the results of standard WaveD [JKRP] when applied to a noisy boxcar convolution when the boxcar scale is a BA number ($a = 1/\sqrt{353}$) which in Table 1 is indicated by boxcar-scale=BA. For the Minimax and Maxiset tunings we simulated noisy boxcar convolution with a randomly chosen from a $\mathcal{U}(0.025, 075)$ distribution which in Table 1 is indicated by boxcar-scale=AA ('Almost All').

Analysis of the results. Our numerical study confirms the theoretical results of Section 3. First, we see that the WaveD estimator can be applied successfully to *almost all* boxcar convolutions. An examination of Table 1 shows that the WaveD results obtained with smaller thresholds (Maxiset tuning) are slightly better than those obtained with larger thresholds (Minimax tuning). Finally, comparing the results obtained in the BA case with those of the AA setting we note slightly poorer performances in the AA case, as to be expected from the theory of Section 3.

Discussion. The WaveD estimator is based on Hard Thresholding and enjoys fast computation (Donoho and Raimondo (2004)). Alternative thresholding strategies such as block-wise thresholding (Pensky and Vidakovic (1999) and Cavalier and Tsybakov (2002)) and the multichannel approach (Pensky and Zayed (2002)) may also give good results. The properties of WaveD presented here can be extended to the 2-dimensional setting with applications to image deblurring, see Donoho and Raimondo (2005). A key feature of the WaveD method is to use band-limited wavelets [JKPR]. Current versions of the `WaveD` software use the Meyer wavelet but alternative approaches using other band-limited wavelet bases are possible. A general discussion on band-limited wavelet bases can be found in Hernàndez and Weiss (1996).

5. Proofs

Proof of Proposition 1. The bound (29) was derived in [JKPR]. Here we prove the bound (30). We start with a lemma deducible from Johnstone and Raimondo (2004).

Lemma 1. *Let p/q and p'/q' be successive principal convergents in the continued fraction expansion of a real number a, and let A_δ be a full Lebesgue measure set (as arises in the Definition 4). Let $q \geq 4$ and N be a non-negative integer with $N+q < q'$. Then there exists a constant $c_2 > 0$ such that, for all numbers $a \in A_\delta$,*

$$\sum_{l=N+1}^{N+q} \|la\|^{-2} \le c_2\ q^2\,(\log q)^{2+2\delta}. \tag{40}$$

By definition of the Meyer wavelet, $C_j \subset \{l : 2^j \le |l| \le 2^{j+2}\}$. To simplify the exposition we use the symmetry of $x \to x^2$ about 0, and consider that $C_j \subset \{l : 2^j \le l \le 2^{j+2}\}$. Let m be the smallest integer such that $q_m \ge 2^j$. From the geometric growth of the convergents $q_{m+2r} \ge 2^r q_m$ and $C_j \subset \mathbb{N} \cap [q_{m-1}, q_{m+4})$. Introduce sets $D_\tau = \mathbb{N} \cap [q_{m+\tau-1}, q_{m+\tau})$, $\tau = 0, \ldots, 4$, whose union covers C_j. For all $a \in A_\delta$ we have that $q_{m+\tau} \le q_{m+\tau-1}\,(\log q_{m+\tau-1})^{1+\delta}$. Hence, there are at most $K_\tau = (\log q_{m+\tau-1})^{1+\delta}$ blocks of length $q_{m+\tau-1}$ to cover D_τ; we apply Lemma 1 within each block:

$$\sum_{D_\tau} \|la\|^{-2} = \sum_{blocks}\ \sum_{l\in block} \|la\|^{-2} \le K_\tau\, c_1\,(q_{m+\tau-1})^2\ K_\tau^2 = c_1\,(q_{m+\tau-1})^2\ K_\tau^3.$$

Taking logs in (27), we see that there exists a constant $c_2 = C(\delta, \tau)$ such that

$$K_\tau \le c_2 K_0, \quad q_{m+\tau-1} \le c_2\, q_{m-1}\, K_0^\tau, \quad \tau = 1, \ldots, 4. \tag{41}$$

It follows that

$$\begin{aligned}\sum_{C_j} \|la\|^{-2} &\le \sum_{\tau=0}^{4} \sum_{D_\tau} \|la\|^{-2}\\ &\le c_1\big(q_{m-1}^2\ K_0^3 + q_m^2\ K_1^3 + q_{m+1}^2\ K_1^3 + q_{m+2}^2\ K_3^3 + q_{m+3}^2\ K_\tau^4\big)\\ &\le c_2\, K_0^3\, q_{m-1}^2 (1 + K_0^2 + K_0^4 + K_0^6 + K_0^8) \le c_3 K_0^{11}\, q_{m-1}^2.\end{aligned}$$

By construction $q_{m-1} \le 2^j$ and so $K_0 = (\log q_{m-1})^{1+\delta} \le c_4\, j^{1+\delta}$ which, combined with (26), proves (30):

$$\tau_j^2 = \ |C_j|^{-1} \sum_{l\in C_j} |b_l|^{-2} \le C\, 2^j \sum_{C_j} \|la\|^{-2} \le C\, 2^{3j}\, j^{11(1+\delta)}.$$

Proof of Proposition 2. We consider the WaveD estimator (7) with the Minimax tuning (31) and (32) and we apply Theorem 1. In this setting, arguments similar to those in [JKPR] (Appendix A) are used to prove that the following claims hold (C1) inequalities (15) and (16) hold with $\eta \ge 2\sqrt{8\pi(p \vee 2)}$; (C2) the basis $(\sigma_j \Psi_{jk})$ satisfies (51); (C3) (14) is satisfied with our choice of parameters. Hence, Theorem 1 applies to (7) which gives $\mathbb{E}\|\hat{f}_1 - f\|_p^p \le C\, c_n^{p-q}$, if and only if

$$f \in l_{q,\infty}(\mu) = \left\{ f,\ \sup_{\lambda>0} \lambda^q \mu\{(j,k) :\ |\beta_{j,k}| > \sigma_j \lambda\} < \infty \right\},\ \text{and,} \tag{42}$$

$$\sup_n c_n^{q-p} \|f - \sum_{\kappa\in\Lambda_n} \beta_\kappa \psi_\kappa\|_p^p < \infty, \tag{43}$$

with

$$\mu\{(j,k)\} = 2^{j(2p-1)}. \tag{44}$$

We take $p - q = 2\beta$ and derive the rate results (33) for standard Besov spaces. Condition (43) essentially uses the same arguments as in [JKPR] (Appendix B.1).

For condition (46), as in [JKPR], we suppose that $f \in B^s_{\pi,r}(T)$ and distinguish between the cases $s+2 \geq 2p/\pi$ and $s+2 < 2p/\pi$. In the first case, we take $q = 2p/(s+2)$ and have (all the inequalities below are true up to obvious absolute constants)

$$\begin{aligned}\mu\{(j,k)\ :|\beta_{jk}| > 2^{\frac{3j}{2}}\lambda\} &= \sum_{jk} 2^{j(2p-1)} I\{|\beta_{jk}| > 2^{\frac{3j}{2}}\lambda\}\\ &\leq \sum_j 2^{2pj} \wedge 2^{j(2p-1)} \sum_k [\frac{|\beta_{jk}|}{2^{\frac{3j}{2}}\lambda}]^\pi\\ &\leq \sum_j 2^{2pj} \wedge [\frac{1}{\lambda}]^\pi 2^{-j(s\pi+2(\pi-p))}\varepsilon_j^\pi, \end{aligned}\tag{45}$$

where (ε_j) is a sequence belonging to l_r. If J is such that $2^{-J(s+2)} \sim \lambda$, we get

$$\begin{aligned}\mu\{(j,k)\ :|\beta_{jk}| > 2^{\frac{3j}{2}}\lambda\} &\leq \sum_{j\leq J} 2^{2pj} + \sum_{j>J} [\frac{1}{\lambda}]^\pi 2^{-j(s\pi+2(\pi-p))}\varepsilon_j^\pi\\ &\leq \lambda^{\frac{-2p}{s+2}} + [\frac{1}{\lambda}]^\pi [\lambda^{\frac{\pi(s+2)-2p}{s+2}} I\{s+2 > \frac{2p}{\pi}\}\\ &\quad + \sum_j \varepsilon_j^\pi I\{s+2 = \frac{2p}{\pi}\}],\end{aligned}$$

completing the proof of this case. Now, if $s+2 < 2p/\pi$ we take $q = (2p-1)/(2+s-1/\pi)$ and use Sobolev embeddings $B^s_{\pi,r}(T) \subset B^{s'}_{q,r}(T)$ for $s' - 1/q = s - 1/\pi$. Using (45), with q instead of π, we get

$$\begin{aligned}\mu\{(j,k)\ :|\beta_{jk}| > 2^{\frac{3j}{2}}\lambda\} &\leq \lambda^{-q} \sum_j 2^{-j(s'q+2(q-p))}\varepsilon_j^q\\ &\leq \lambda^{-q} \sum_j 2^{-j\frac{2p-\pi(s+2)}{2+s-\frac{1}{\pi}}}.\end{aligned}$$

Proof of Proposition 3. We consider the WaveD estimator (7) with the Maxiset tuning (34) and (35) and apply Theorem 1. In this setting, arguments similar to those in Appendix A of [JKPR] are used to prove that the following claims hold: (C1) inequalities (15) and (16) hold with $\eta \geq 2\sqrt{8\pi(p\vee 2)}$; (C2) the basis $(\sigma_j \Psi_{jk})$ satisfies (51); (C3) (14) is satisfied with our choice of parameters. Hence, Theorem 1 applies to (7) which gives $\mathbb{E}\|\hat{f}_2 - f\|_p^p \le C\, c_n^{p-q}$ if and only if:

$$f \in l_{q,\infty}(\tilde{\mu}) = \left\{ f,\ \sup_{\lambda>0} \lambda^q \tilde{\mu}\{(j,k) :\ |\beta_{j,k}| > \sigma_j(2)\lambda\} < \infty \right\}, \quad \text{and,} \\ \sup_n c_n^{q-p} \|f - \sum_{\kappa\in\Lambda_n} \beta_\kappa \psi_\kappa\|_p^p < \infty, \tag{46}$$

with

$$\tilde{\mu}\{(j,k)\} = 2^{j(2p-1)} j^{\frac{zp}{2}}. \tag{47}$$

Using standard wavelet arguments, it is elementary, using the definitions of j_1 and c_n, to see that

$$\sup_n c_n^{q-p} \|f - \sum_{\kappa\in\Lambda_n} \beta_\kappa \psi_\kappa\|_p^p < \infty \iff \sup_n c_n^{q-p} \sum_{j\ge j_1} 2^{j(\frac{p}{2}-1)} \sum_k |\beta_{jk}|^p < \infty.$$

And this is obviously true if $f \in \tilde{B}_{p,\infty}^{2(p-q)/p, z(p-q)/2p}$. Now, again we take $\beta = p-q$ and prove rate result (36) over Besov sub-scales (Definition 2), using arguments quite analogous to those of Proposition 2. First we observe that the fact of inclusion in the space $B_{p,\infty}^{2(p-q)/p, z(p-q)/2p}$ is established similar to the proof given in [JKPR].

For the inclusion in $l_{q,\infty}(\tilde{\mu})$ we have, as in (45), in the case $s+2 \ge 2p/\pi$,

$$\begin{aligned}\tilde{\mu}\{(j,k)\ : |\beta_{jk}| > 2^{\frac{3j}{2}} j^{\frac{z}{2}} \lambda\} &= \sum_{jk} 2^{j(2p-1)} j^{\frac{zp}{2}} I\{|\beta_{jk}| > 2^{\frac{3j}{2}} j^{\frac{z}{2}} \lambda\} \\ &\le \sum_j 2^{2pj} j^{\frac{zp}{2}} \wedge 2^{j(2p-1)} j^{\frac{zp}{2}} \sum_k [\frac{|\beta_{jk}|}{2^{\frac{3j}{2}} j^{\frac{z}{2}} \lambda}]^\pi \\ &\le \sum_j 2^{2pj} j^{\frac{zp}{2}} \wedge [\frac{1}{\lambda}]^\pi 2^{-j(s\pi+2(\pi-p))} j^{\frac{z(p-\pi)}{2} - \tau\pi}.\end{aligned}$$

Let J such that $2^J \sim \lambda^{1/(s+2)} \log \frac{1}{\lambda}^{-z/4}$. We get

$$\begin{aligned}\tilde{\mu}\{(j,k)\ /|\beta_{jk}| > 2^{\frac{3j}{2}} j^{\frac{z}{2}} \lambda\} &\le \sum_{j\le J} 2^{2pj} j^{\frac{zp}{2}} + \sum_{j>J} [\frac{1}{\lambda}]^\pi 2^{-j(s\pi+2(\pi-p)} j^{\frac{z(p-\pi)}{2} - \tau\pi} \\ &\le \lambda^{\frac{-2p}{s+2}} + [\frac{1}{\lambda}]^\pi [\lambda^{\frac{\pi(s+2)-2p}{s+2}} \log \frac{1}{\lambda}^{\frac{zs\pi}{4} - \tau\pi} I\{s+2 > \frac{2p}{\pi}\} \\ &\quad + \sum_j j^{\frac{zs\pi}{4} - \tau\pi} I\{s+2 = \frac{2p}{\pi}\}],\end{aligned}$$

which completes the proof of this case. If $s+2 < 2p/\pi$ we take $q = (2p-1)/(2+s-1/\pi)$ and use Sobolev embeddings and (45) with q instead of π to find

$$\tilde{\mu}\{(j,k)\ :|\beta_{jk}| > 2^{\frac{3j}{2}} j^{\frac{z}{2}} \lambda\} \le \lambda^{-q} \sum_j 2^{-j(s'q+2(q-p))} j^{\frac{z(p-q)}{2}-\tau q}$$
$$\le \lambda^{-q} \sum_j 2^{-j\frac{2p-\pi(s+2)}{2+s-\frac{1}{\pi}}} j^{\frac{z(p-q)}{2}-\tau q}.$$

Proof of Proposition 4. It is a consequence of Theorem 1 and Corollary 1.

Proof of Proposition 5. We start with a Lemma.

Lemma 2. *Let* $0 < q < \infty$, $-\infty < \alpha < \infty$. *Take* $\xi(t) = \xi_{q,\alpha}(t)$ *to be a continuous non-decreasing function such that* $\xi(0) = 0$ *and*

$$\xi(t) = \begin{cases} t^q(\log(\frac{1}{t}))^\alpha, & t \in [0,\zeta] \\ (\log(\frac{1}{\zeta}))^\alpha t^q & t > \zeta, \end{cases}$$

where $0 < \zeta \le \exp -\alpha/q$ *if* $\alpha \ge 0$, $0 < \zeta < 1$ *if* $\alpha < 0$. *Then*

$$\exists C_0, \quad \forall\, \lambda > 0, \qquad \sum_{j\ge 0} \frac{1}{\xi(2^j\lambda)} \le \frac{C_0}{\xi(\lambda)}, \tag{48}$$

$$\forall\, p > q, \quad \exists C_p, \quad \forall\, \lambda > 0, \qquad \sum_{j\ge 0} \frac{1}{2^{jp}\xi(2^{-j}\lambda)} \le \frac{C_p}{\xi(\lambda)}. \tag{49}$$

Proof. To prove (48), let $j_0 = \inf\{j \in \mathbb{N},\ 2^j\lambda > \zeta\}$. Then

$$\sum_{j\ge 0} \frac{1}{\xi(2^j\lambda)} = \sum_{j<j_0} \frac{1}{(2^j\lambda)^q(\log\frac{1}{2^j\lambda})^\alpha} + \sum_{j\ge j_0} \frac{1}{(\log(\frac{1}{\zeta}))^\alpha(2^j\lambda)^q}.$$

The result is clearly obvious if $j_0 = 0$. If $j_0 > 0$, and $0 \le j < j_0$, then $\lambda \le 2^j\lambda \le \zeta < 2^{j_0}\lambda < 2\zeta$, and

$$\sum_{j\ge 0} \frac{1}{\xi(2^j\lambda)} \le \frac{1}{(\log\frac{1}{\lambda})^\alpha\lambda^q} \sum_{j<j_0} 2^{-jq}\Big(\frac{\log\frac{1}{\lambda}}{\log\frac{1}{2^j\lambda}}\Big)^\alpha + c_q \frac{1}{(\log(\frac{1}{\zeta}))^\alpha(2^{j_0}\lambda)^q}.$$

As $(\log 1/\lambda)^\alpha\lambda^q$ is a nondecreasing function on $[0,\zeta]$,

$$\frac{1}{(\log(\frac{1}{\zeta}))^\alpha(2^{j_0}\lambda)^q} \le \frac{1}{(\log(\frac{1}{\zeta}))^\alpha\zeta^q} \le \frac{1}{(\log\frac{1}{\lambda})^\alpha\lambda^q}$$

and for $0 \le j < j_0$,

$$1 \le \frac{\log\frac{1}{\lambda}}{\log\frac{1}{2^j\lambda}} = \frac{\log\frac{1}{2^j\lambda} + j\log 2}{\log\frac{1}{2^j\lambda}} = 1 + j\frac{\log 2}{\log\frac{1}{2^j\lambda}} \le 1 + j\frac{\log 2}{\log\frac{1}{\zeta}}.$$

If $\alpha \geq 0$,

$$\sum_{j<j_0} 2^{-jq}\Big(\frac{\log\frac{1}{\lambda}}{\log\frac{1}{2^j\lambda}}\Big)^{\alpha} \leq \sum_{j<j_0} 2^{-jq}\Big(1+j\frac{\log 2}{\log\frac{1}{\zeta}}\Big)^{\alpha} \leq C\,,$$

with $\alpha < 0$,

$$\sum_{j<j_0} 2^{-jq}\Big(\frac{\log\frac{1}{\lambda}}{\log\frac{1}{2^j\lambda}}\Big)^{\alpha} \leq \sum_{j<j_0} 2^{-jq} \leq C\,.$$

To prove (49), let $j_0 = \inf\{j \in \mathbb{N},\ 2^{-j}\lambda \leq \zeta\}$, $j_0 = 0$ if $\lambda \leq \zeta$. Then

$$\begin{aligned}\sum_{j\geq 0}\frac{1}{2^{jp}\xi(2^{-j}\lambda)} &= \sum_{j\geq j_0}\frac{1}{2^{jp}(2^{-j}\lambda)^q(\log(\frac{1}{2^{-j}\lambda}))^{\alpha}} + \sum_{0\leq j<j_0}\frac{1}{(\log(\frac{1}{\zeta}))^{\alpha}2^{jp}(2^{-j}\lambda)^q}\\ &= \frac{1}{(\log\frac{1}{\lambda})^{\alpha}\lambda^q}\sum_{j\geq j_0}2^{-j(p-q)}\left(\frac{\log(\frac{1}{\lambda})}{\log(\frac{1}{2^{-j}\lambda})}\right)^{\alpha} + \frac{1}{(\log(\frac{1}{\zeta}))^{\alpha}\lambda^q}\sum_{0\leq j<j_0}2^{-j(p-q)}\,.\end{aligned}$$

If $\lambda \leq \zeta$ we have only the first term and

$$\begin{aligned}&\text{if}\quad \alpha \geq 0, \quad \Big(\frac{\log(\frac{1}{\lambda})}{\log(\frac{1}{2^{-j}\lambda})}\Big)^{\alpha} \leq 1\,,\\ &\text{if}\quad \alpha < 0, \quad \Big(\frac{\log(\frac{1}{\lambda})}{\log(\frac{1}{2^{-j}\lambda})}\Big)^{\alpha} \leq \Big(1+\frac{j\log 2}{\log(\frac{1}{\zeta})}\Big)^{|\alpha|}\,,\end{aligned}$$

and we get the result.
If $j_0 < \infty$, then for $2^{-j_0}\lambda \leq \zeta < 2^{-(j_0-1)}\lambda$:

$$\begin{aligned}&\text{if}\quad \alpha \geq 0,\ \frac{1}{\lambda^q}\sum_{j\geq j_0}\frac{2^{-j(p-q)}}{(\log(\frac{1}{2^{-j}\lambda}))^{\alpha}} \leq \Big(\frac{1}{\log\frac{1}{\zeta}}\Big)^{\alpha}\frac{2^{-j_0(p-q)}}{\lambda^q}\sum_{j\geq j_0}2^{-(j-j_0)(p-q)} \leq \frac{C}{\lambda^p}\,;\\ &\text{if}\quad \alpha < 0,\ \frac{1}{\lambda^q}\sum_{j\geq j_0}\frac{2^{-j(p-q)}}{(\log(\frac{1}{2^{-j}\lambda}))^{\alpha}} \leq \frac{2^{-j_0(p-q)}}{\lambda^q}\sum_{j\geq j_0}2^{-(j-j_0)(p-q)}\Big(\log\Big(\frac{1}{2^{-j_0}\lambda}\\ &\qquad\qquad +(j-j_0)\Big)\Big)^{|\alpha|} \leq \frac{C}{\lambda^p}\,.\end{aligned}$$

But as $\lambda \geq \zeta$, $1/\lambda^p = O(1/\lambda^q)$.

Corollary 2. *Let I be a set, $q < p$, and $\forall\, i$, $\mu(i) \geq 0$. For all $\gamma : I \mapsto \mathbb{R}$ we have the following equivalence :*

1. $\exists C > 0, \forall\, \lambda > 0, \displaystyle\sum_{\{i, |\gamma(i)|>\lambda\}} \mu(i) \leq \frac{C}{\xi(\lambda)}.$

2. $\exists C', \forall \lambda > 0, \displaystyle\sum_{\{i, |\gamma(i)| \le \lambda\}} |\gamma(i)|^p \mu(i) \le C' \frac{\lambda^p}{\xi(\lambda)}.$

Proof. $1 \Longrightarrow 2$. We use (49) to get

$$\begin{aligned}\sum_{\{i,|\gamma(i)|\le\lambda\}} |\gamma(i)|^p\mu(i) &= \sum_{j\in\mathbb{N}} \sum_{\{i,2^{-j-1}|\gamma(i)|\le 2^{-j}\lambda\}} |\gamma(i)|^p\mu(i)\\ &\le \lambda^p \sum_{j\in\mathbb{N}} 2^{-jp} \sum_{\{i,2^{-j-1}|\gamma(i)|\}} \mu(i) \le \lambda^p \sum_{j\in\mathbb{N}} 2^{-jp} \frac{C}{\xi(2^{-j-1}\lambda)}\\ &\le 2CC_p \frac{\lambda^p}{\xi(\lambda)}.\end{aligned}$$

$2 \Longrightarrow 1$. We use (48) to get

$$\begin{aligned}\sum_{\{i,|\gamma(i)|>\lambda\}} \mu(i) &= \sum_{j\in\mathbb{N}} \sum_{\{i,2^{j+1}\lambda\ge|\gamma(i)|>2^j\lambda\}} \mu(i) \le \sum_{j\in\mathbb{N}} \frac{1}{2^{jp}\lambda^p} \sum_{\{i,\ 2^{j+1}\lambda\ge|\gamma(i)|\}} \mu(i)|\gamma(i)|^p\\ &\le \sum_{j\in\mathbb{N}} \frac{1}{2^{jp}\lambda^p} C' \frac{2^{(j+1)p}\lambda^p}{\xi(2^{j+1}\lambda)} \le 2^p C' C_0 \frac{1}{\xi(\lambda)}\,.\end{aligned}$$

We apply the previous Lemma with $\mu(j,k) = 2^{j(p/2-1)}\tau_j^p$ and $\gamma(jk) = \beta_{jk}/\tau_j$.

Lemma 3. *If τ_j an arbitrary positive sequence, $p > q$, and β are positive real numbers, the following assertions are equivalent.*

1. *There exists C, such that:*

$$\forall \lambda > 0, \sum_{j\ge -1,k} 2^{j(\frac{p}{2}-1)} \tau_j^p I\{|\frac{\beta_{jk}}{\tau_j}| \ge \lambda\} \le C \frac{1}{\xi_{(q,\beta)}(\lambda)}\,.$$

2. *There exists C_1, such that:*

$$\forall \lambda > 0,\ \sum_{j,k} |\beta_{jk}|^p I\{|\frac{\beta_{jk}}{\tau_j}| \le \lambda\} 2^{j(\frac{p}{2}-1)} \le C_p \frac{\lambda^p}{\xi_{(q,\beta)}(\lambda)}\,.$$

To prove Proposition 5, let us suppose that $f \in MAX(1)$, and prove that $f \in l_{q,\infty}(\tilde{\mu})$. Let C be a generic constant which may change from line to line. Observe first that

$$\forall n\in\mathbb{N},\ n^{-1}\log(n)]^{\frac{q-p}{2}} \|f - \sum_{\kappa, j\le j_1} \beta_\kappa \psi_\kappa\|_p^p \le C < \infty, \quad \text{for} \quad 2^{j_1} = n^{\frac{1}{4}} (\log n)^{-\frac{1+z}{4}},$$

which implies that

$$\forall\, j_1 \in \mathbb{N}, \quad \Big\| f - \sum_{\kappa, j\le j_1} \beta_\kappa \psi_\kappa \Big\|_p \le C 2^{\frac{-2(p-q)j_1}{p}} j_1^{\frac{-z(p-q)}{p}}.$$

$$\forall\, j \ \in \mathbb{N}, \quad \Big(\sum_k |\beta_{jk}|^p 2^{j(\frac12 - \frac1p)} \Big)^{\frac1p} \le C 2^{\frac{-2(p-q)j}{p}} j^{\frac{-z(p-q)}{p}}.$$

More generally

$$\forall\, J > j \ \in \mathbb{N}, \qquad \sum_{j<j'<J,k} |\beta_{j'k}|^p 2^{j'(\frac{p}{2}-1)} \le C 2^{-2(p-q)j} j^{-z(p-q)}. \tag{50}$$

To apply Lemma 3 we have to check that

$$\sum_{jk} |\beta_{jk}|^p 2^{j(\frac{p}{2}-1)} 1_{|\frac{\beta_{jk}}{\sigma_j(2)}| \le \lambda} \le \lambda^{p-q}.$$

Obviously, from (50) we have only to check this for λ in a neighbourhood of 0, $[0,\kappa]$ say. Let j_0 be such that $j_0 \sim \log 1/\lambda$. Now

$$\begin{aligned}
&\sum_{jk} |\beta_{jk}|^p 2^{j(\frac{p}{2}-1)} 1_{|\frac{\beta_{jk}}{\sigma_j(2)}| \le \lambda} \\
&\quad = [\sum_{j\le j_0} + \sum_{j>j_0}] \sum_k |\beta_{jk}|^p 2^{j(\frac{p}{2}-1)} 1_{|\frac{\beta_{jk}}{\sigma_j(2)}| \le \lambda} \\
&\quad \le \sum_{j\le j_0} \sum_k |\beta_{jk}|^p 2^{j(\frac{p}{2}-1)} 1_{|\frac{\beta_{jk}}{\sigma_j(1)}| \le \lambda (\log \frac{1}{\lambda})^{\frac{z}{2}}} + \sum_{j>j_0} \sum_k |\beta_{jk}|^p 2^{j(\frac{p}{2}-1)} \\
&\quad \le C[\lambda (\log \frac{1}{\lambda})^{\frac{z}{2}}]^{p-q} (\log \frac{1}{\lambda})^{\frac{-z(p-q)}{2}} + C' 2^{-j_0 2(p-q)} \le C" \lambda^{p-q}.
\end{aligned}$$

The first term of the last inequality is obtained by another application of Lemma 3 using the assumption $f \in l_{\xi_{q,z(p-q)/2)},\infty}(\mu)$. The second term is obtained using (50).

Appendix. Temlyakov Inequalities

Let us recall the Temlyakov property for a basis $e_n(x)$ in L^p: there exists absolute constants c, C such that for all $\Lambda \subset \mathbb{N}$,

$$c \sum_{n\in\Lambda} \int |e_n(x)|^p dx \le \int (\sum_{n\in\Lambda} |e_n(x)|^2)^{p/2} dx \le C \sum_{n\in\Lambda} \int |e_n(x)|^p dx$$

or, equivalently,

$$c' \|(\sum_{n\in\Lambda} |e_n(x)|^p)^{1/p}\|_p \le \|(\sum_{n\in\Lambda} |e_n(x)|^2)^{1/2}\|_p \le C' \|(\sum_{n\in\Lambda} |e_n(x)|^p)^{1/p}\|_p . \tag{51}$$

Acknowledgement

We are grateful to a referee and an associate editor for their constructive comments. This project began when Marc Raimondo was invited to the Laboratoire de Probabilités et Modèles Aléatoires de l'Université de Paris VII in December 2003.

References

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

Bertero, M. and Boccacci, P. (1998). *Introduction to Inverse Problems in Imaging*. Institute of Physics, Bristol and Philadelphia.

Cavalier, L. and Raimondo, M. (2007). Wavelet deconvolution with noisy eigen values. *IEEE Trans. Signal Process.* To appear.

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

Donoho, D. L. (1995). Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmon. Anal.* **2**, 101-126.

Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. and Picard, D. (1995). Wavelet shrinkage: asymptopia? (with discussion). *J. Roy. Statist. Soc. Ser. B* **57**, 301-369.

Donoho, D. L. and Raimondo, M. E. (2005). A fast wavelet algorithm for image deblurring. In *Proc. of* 12*th Computational Techniques and Applications Conference*, CTAC-2004 46, C29-C46. http://anziamj.austms.org.au/V46/CTAC2004/Dono.

Donoho, D. L. and Raimondo, M. E. (2004). Translation invariant deconvolution in a periodic setting. *Int. J. Wavelets, Multiresolut. Inf. Process.* **14**, 415.

Fan, J. and Koo, J. (2002). Wavelet deconvolution. *IEEE Trans. Inform. Theory* **48**, 734-747.

Hall, P., Ruymgaart, F., van Gaans, O. and van Rooij, A. (2001). Inverting noisy integral equations using wavelet expansions: A class of irregular convolutions. In *State of the Art in Probability and Statistics: Festschrift for Willem R. van Zwet*. Lecture Notes-Monograph Series **36**, 533-546. Institute of Mathematical Statistics.

Hernàndez, E. and Weiss, G. (1996) . *A First Course on Wavelets*. CRC Press.

Johnstone, I., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004). Wavelet deconvolution in a periodic setting. (with discussion). *J. Roy. Statist. Soc. Ser. B* **66**, 547-573.

Johnstone, I. M. (1999). Wavelet shrinkage for correlated data and inverse problems: adaptivity results. *Statist. Sinica* **9**, 51-83.

Johnstone, I. M. and Raimondo, M. (2004). Periodic boxcar deconvolution and diophantine approximation. *Ann. Statist.* **32**, 1781-1804.

Kalifa, J. and Mallat, S. (2003). Thresholding estimators for linear inverse problems and deconvolutions. *Ann. Statist.* **31**, 58-109.

Kerkyacharian, G. and Picard, D. (2000). Thresholding algorithms and well-concentrated bases. *Test* **9**.

Khintchine, A. (1963). *Continued Fractions*. Noordhoff, Groningen.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. 2nd edition. Academic Press Inc., San Diego, CA.

Meyer, Y. (1990). *Ondelettes et Opérateurs-I.* Hermann.

Neelamani, R., Choi, H. and Baraniuk, R. (2004). Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Trans. Signal Process.* **52**, 418-433.

Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Statist.* **27**, 2033-2053.

Pensky, M. and Zayed, A. (2002). Density deconvolution of different conditional distributions. *Ann. Inst. Statist. Math.* **54**, 701-712.

Walter, G. and Shen, X. (1999). Deconvolution using the Meyer wavelet. *J. Integral Equations Appl.* **11**, 515-534.

Université de Paris X, 200 Avenue de la République 92001 Nanterre Cedex, France.

E-mail: kerk@math.jussieu.fr

Laboratoire de Probabilités et Modèles aléatoires, Universite Paris VII, 4, Place Jussieu, 75252 PARIS Cedex 05, France.

E-mail: picard@math.jussieu.fr

School of Mathematics and Statistics F07, University of Sydney NSW 2006, Australia

E-mail: marcr@maths.usyd.edu.au

(Received September 2004; accepted September 2005)