



HAL
open science

Optimal bounds for inverse problems with Jacobi-type eigenfunctions

Thomas Willer

► **To cite this version:**

Thomas Willer. Optimal bounds for inverse problems with Jacobi-type eigenfunctions. *Statistica Sinica*, 2009, 19 (2), pp.785-800. hal-00133830v1

HAL Id: hal-00133830

<https://hal.science/hal-00133830v1>

Submitted on 27 Feb 2007 (v1), last revised 21 Oct 2013 (v2)

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.

NEW MINIMAX RATES FOR INVERSE PROBLEMS

Thomas Willer

Laboratoire de Probabilités et Modèles Aléatoires, Paris, France

Abstract: We consider inverse problems where one wishes to recover an unknown function from the observation of a transformation of it by a linear operator, corrupted by an additive white noise perturbation. We assume that the operator admits a singular value decomposition where the eigenvalues decay in a polynomial way, and where Jacobi polynomials appear as eigenfunctions. This includes, as an application, the well known Wicksell's problem. We determine the asymptotic rate of the minimax risk for this model in a wide framework, considering $(L^p)_{1 < p < \infty}$ losses, and Besov-like regularity spaces. We draw a comparison with the minimax rates of the deconvolution problem, which appears as a critical case of the Jacobi-type rates. We also establish some new results on the needlets introduced by Petrushev and Xu (2005) which appear as essential tools in this setting.

Key words and phrases: statistical inverse problems, minimax estimation, second-generation wavelets.

1. Motivation

We consider the problem of recovering a function f from a blurred (by a linear operator) and noisy version Y :

$$\forall v \in V, \quad Y(v) = (Kf, v)_V + \epsilon \xi(v),$$

where K is a linear operator between two Hilbert spaces: $K : U \mapsto V$, ξ is a V -white noise, and for H a Hilbert space and $h_1, h_2 \in H$, $(h_1, h_2)_H$ denotes the scalar product in H between h_1 and h_2 . We assume that f belongs to $U = L^2([-1, 1], \mu(x)dx)$, with $\mu(x) = (1-x)^\alpha(1+x)^\beta$, $\alpha, \beta > -1/2$, and that K admits a singular value decomposition (SVD), i.e. there exists an orthonormal basis (called SVD basis) formed by the eigenfunctions of the self-adjoint operator K^*K (where K^* is the adjoint of K). Moreover we assume that this SVD basis consists of the classical Jacobi polynomials of type (α, β) (see Szegő (1975)), and that the corresponding sequence of eigenvalues tend to zero at a polynomial rate. We will call such problems "Jacobi-type inverse problems".

The main motivation of this article is to establish the asymptotic minimax rates in a wide framework, considering $L^p([-1, 1], \mu)$ losses, for all $1 < p < \infty$, and a Besov like regularity space. We find rates which are new in the literature, and we give applications and draw parallels among well known inverse problems in practice, such as Wicksell's corpuscule problem or the deconvolution problem.

1.1 Interest of the results for inverse problems theory

The most popular technique for the treatment of inverse problems is probably singular value decomposition estimation, where the unknown function is expanded in the SVD basis, and the coefficients are estimated thanks to Y . Such techniques are very attractive theoretically and can be shown to be asymptotically minimax in many situations (see e.g. Mathé and Pereverzev (2003), Cavalier and Tsybakov (2002), Cavalier, Golubev, Picard, and Tsybakov (2002), Tsybakov (2000), Goldenshluger and Pereverzev (2003), Efromovich and Koltchinskii (2001))). However there are limitations in the minimax framework, in particular such estimators generally cannot estimate functions exhibiting inhomogeneous regularity. To avoid this problem, several wavelets methods have been introduced during the last decade (Donoho (1995) and Abramovich and Silverman (1998)), which are minimax over wide sets of target functions, for example Besov spaces. However such methods apply only to a category of inverse problems where the operator is well adapted to the structure of first generation wavelets (for example homogeneous operators).

Then the main interest of our results is to grapple with problems with a polynomial structure, which cannot be analysed through the perspective of classical wavelets (which are more adapted to operators linked to Fourier analysis). The main idea is to use new wavelets built upon polynomials (termed needlets, and introduced by Petrushev and Xu (2005)), and new spaces, which appear as an adaptation of the classical Besov spaces. Thanks to these new tools, we derive the minimax rates of all the Jacobi-type inverse problems. So this paper is closely linked to Kerkyacharian, Picard, Petrushev, and Willer (2006), where an estimation procedure, called NeedVD and based on needlets, was developed for several inverse problems, including Jacobi-type models. Here we prove that NeedVD is nearly optimal, since the lower bounds established for the minimax risk turn to match with the convergence rates of the procedure.

Note also that the results are established for all $L^p([-1, 1], \mu)$, whereas in most other works only the case $p = 2$ is considered, except for the deconvolution problem: this problem is well adapted to classical wavelet treatments (Pensky and Vidakovic (1999), Fan and Koo (2002), Kalifa and Mallat (2003)) and minimax rates for $L^p([0, 1], dx)$ losses and over Besov spaces were established in Johnstone, Kerkyacharian, Picard, and Raimondo (2004) and Willer (2005). We will draw a parallel between those rates and the ones obtained here: we exhibit elbow effects, and we show that the rates in the deconvolution model appear as a critical case of the rates in the Jacobi-type model. Moreover, we also give an application of our results to Wicksell's problem (Wicksell (1925)), which satisfies the required assumptions on the operator. This problem concerns the recovery of the density of the radii of spherical particles, when a sample of planar cuts is given, and has many applications in medicine and in biology.

1.2 Mains tools used to establish the results

The principle of NeedVD is to decompose the inverse problem in a needlet frame. Such functions are adapted at the same time adapted to K (as they are closely linked to the SVD basis of the problem, i.e. to Jacobi polynomials), and to f (as they enable sparse representations of various functions, including spatially inhomogeneous ones). In this paper our main concern is to lower bound the minimax risk. However the same kind of problematic arises, as the problem is to find a family of hypotheses $\{f_\lambda, \lambda \in \Lambda\} \subset \mathcal{U}$, representative of the difficulties of estimation inside the regularity space considered for the risk. This means that the functions f_λ must be chosen such that:

- they are distant from one another in $L^p(\mu)$ norm,
- and in the same time the distributions of the associated processes Y are close to one another (in a Kullback sense, for example).

A natural way to build such hypotheses is to use functions which enjoy localization properties, and whose images by K can be easily studied, and thus here again needlets are an essential tool. The hypotheses are built as linear combinations of such functions, with some parameters left free, which we adjust optimally with respect to the two constraints cited above. Then the $L^p(\mu)$

distance between the hypotheses yields the lower bound on the whole regularity space. In this context several additional difficulties must be treated in comparison to classical-wavelet problems such as deconvolution, which stem from the non orthogonality of the needlets and from the heterogeneity of their $L^p(\mu)$ norms. Thus we will make a brief list of the properties of needlets, and we will establish some new properties specific to the lower bound problem.

The paper is organized as follows. In section 2 we describe the model and state the main result, in section 3 we describe the needlets and we give their basic properties used to prove the main result. The fourth section gives the proofs of the minimax rates and the fifth section gives the proofs of two preliminary results on needlets from section 3.

2. Main result

2.1 Model and assumptions

We are interested in nonparametric inverse problems in white noise, with a polynomial structure of the operator. We define this framework as follows. Let f be an unknown function belonging to the Hilbert space $\mathbf{U} = L^2([-1, 1], \mu(x)dx)$, with $\mu(x) = (1-x)^\alpha(1+x)^\beta$, $\alpha, \beta > -1/2$. The estimation problem consists in recovering a good approximation of the function f from the observation of the random variable Y corresponding to a blurred and noisy version of f :

$$\forall v \in \mathbf{V}, \quad Y(v) = (Kf, v)_{\mathbf{V}} + \epsilon \xi(v). \quad (1)$$

Blurring effect: Let $I = [a, b]$ or $I = [a, b[$, with $-\infty < a < b \leq \infty$, and $\lambda : I \mapsto \mathbb{R}_+^*$ a continuous function. We set $\mathbf{V} = L^2(I, \lambda(x)dx)$. Let $K : \mathbf{U} \mapsto \mathbf{V}$ be a linear operator satisfying the two following conditions. First assume K^*K (where K^* denotes the adjoint of K) is diagonalizable, with a countable set of eigenvalues (denoted $(b_k^2)_{k \in \mathbb{N}}$) which are strictly positive and decrease at a polynomial rate for some ill posedness coefficient $\nu > 0$ (for two positive sequences (u_k) and (v_k) , the notation $u_k \asymp v_k$ means that there exist $0 < c_1 \leq c_2 < \infty$ such that $c_1 v_k \leq u_k \leq c_2 v_k$):

$$\forall k \in \mathbb{N}^*, \quad b_k \asymp k^{-\nu}.$$

Secondly, assume that the classical Jacobi polynomials normalized in \mathbf{U} (we denote by P_k the polynomial of degree k) appear as an orthonormal basis of eigen-

functions of K^*K . So P_k is the polynomial of degree k such that $\int_{-1}^1 P_k P_l d\mu = \delta_{k,l}$, and:

$$\forall k \in \mathbb{N}, \quad K^*K P_k = b_k^2 P_k.$$

Noise effect: $\epsilon > 0$ is deterministic, and ξ is a gaussian white noise on V , i.e.:

$$\forall v, w \in V, \quad \begin{cases} \xi(v) \sim \mathcal{N}(0, \|v\|_V^2), \\ \xi(v), \xi(w) \text{ are independent if } v \perp w. \end{cases}$$

2.2 Minimax rates

The aim of the paper is to establish the asymptotic minimax rates (when $\epsilon \rightarrow 0$) for inverse problems described above, in a wide framework, i.e. for numerous choices of functions f and of measures of estimation errors. For the latter, we consider all $\mathbb{L}^p(\mu)$ losses (for any $1 < p < +\infty$) defined by: $\forall u \in \mathcal{U}$, $\|u\|_{\mathbb{L}^p(\mu)} = [\int_{-1}^1 |u(x)|^p d\mu(x)]^{\frac{1}{p}}$.

Concerning the target functions, we introduce spaces $B_{\pi,r}^s(\mathcal{M})$ below, which appear as an adaptation of the classical Besov spaces. Let $(\psi_{j,\eta})_{j \geq 0, \eta \in \mathbb{Z}_j}$ denote the tight frame of needlets described in the next section. For any $f \in \mathcal{U}$, we have the following decomposition:

$$f = \sum_{j \geq 0} \sum_{\eta \in \mathbb{Z}_j} \beta_{j\eta} \psi_{j\eta}, \quad \text{where } \beta_{j\eta} = (f, \psi_{j\eta})_{\mathcal{U}}.$$

Then for $\pi \geq 1$, $s \geq 1/\pi$, $r \geq 1$, $M > 0$ we define:

$$B_{\pi,r}^s(\mathcal{M}) = \{f \in \mathcal{U} \mid \| (2^{js} (\sum_{\eta \in \mathbb{Z}_j} |\beta_{j,\eta}|^\pi \|\psi_{j,\eta}\|_\pi^\pi)^{1/\pi})_{j \geq -1} \|_{l^r} \leq M\}.$$

If $\psi_{j,\eta}$ were a classical wavelet, then $B_{\pi,r}^s$ would be exactly a Besov space. Details on this notion can be found in Härdle, Kerkycharian, Picard, and Tsybakov (1998), we recall simply that they are very general regularity spaces including as particular cases Sobolev and Holder spaces, and which can be described very simply, thanks to any regular enough wavelet basis. They often appear in the study of inverse problems, in particular when the operator displays links with Fourier analysis (Donoho (1995)). However here $B_{\pi,r}^s$ correspond to new spaces, characterized by needlets, which appear as a natural alternative to Besov spaces in the case of polynomial type inverse problems. Details on the space in this case can be found in Narcowich, Petrushev, and Ward (2006).

We are interested in the minimax risk defined by:

$$R_\epsilon(B_{\pi,r}^s(\mathcal{M}), \mathbb{L}^p(\mu)) := \inf_{\hat{f}} \sup_{f \in B_{\pi,r}^s(\mathcal{M})} \mathbb{E}_f(\|\hat{f} - f\|_{\mathbb{L}^p(\mu)}^p),$$

where the infimum is taken over all $\sigma(Y(t))_{t \geq 0}$ -measurable estimators \hat{f} . The results of Kerkyacharian, Picard, Petrushev, and Willer (2006), concerning the rates of convergence of the NeedVD estimator, give immediately an upper bound for the risk. This is Theorem 1, where we recall that $\nu > 0$ is a rate of decay of the eigenvalues of the operator ($b_k \asymp k^{-\nu}$), and that $\alpha, \beta > -\frac{1}{2}$ are parameters characterizing \mathcal{U} .

Theorem 1. *For all $1 < p < \infty$, $\pi \geq 1$, $r \geq 1$ and $s > \max_{\gamma \in \{\alpha, \beta\}} \{\frac{1}{2} - 2(\gamma + 1)(\frac{1}{2} - \frac{1}{\pi}) \vee 2(\gamma + 1)(\frac{1}{\pi} - \frac{1}{p}) \vee 0\}$ there exists $C > 0$ such that:*

$$R_\epsilon(B_{\pi,r}^s(\mathcal{M}), \mathbb{L}^p(\mu)) \leq C[\log(1/\epsilon)]^{p+1} [\epsilon \sqrt{\log(1/\epsilon)}]^\zeta,$$

where $\zeta = \min\{\zeta(s), \zeta(s, \alpha), \zeta(s, \beta)\}$ with:

$$\zeta(s) = \frac{s}{s + \nu + \frac{1}{2}}, \quad \zeta(s, \gamma) = \frac{s - 2(1 + \gamma)(\frac{1}{\pi} - \frac{1}{p})}{s + \nu + 2(1 + \gamma)(\frac{1}{2} - \frac{1}{\pi})}.$$

The main purpose of the paper is to prove that these rates coincide with the rates of the minimax risk up to log factors. We will establish the following theorem:

Theorem 2. *For all $1 < p < \infty$, $\pi \geq 1$, $r \geq 1$ and $s \geq 1/\pi$ there exists $C > 0$ such that:*

$$R_\epsilon(B_{\pi,r}^s(\mathcal{M}), \mathbb{L}^p(\mu)) \geq C\epsilon^\zeta,$$

where $\zeta = \min\{\zeta(s), \zeta(s, \alpha), \zeta(s, \beta)\}$ with:

$$\zeta(s) = \frac{s}{s + \nu + \frac{1}{2}}, \quad \zeta(s, \gamma) = \frac{s - 2(1 + \gamma)(\frac{1}{\pi} - \frac{1}{p})}{s + \nu + 2(1 + \gamma)(\frac{1}{2} - \frac{1}{\pi})}.$$

Note that the exact logarithmic factors of the minimax risk are not established yet. In this paper we have focused only on the main rate ϵ^ζ , so our results prove that NeedVD is "quasi optimal" in the Jacobi-type models.

2.3 Comparison with other inverse problems

In the literature on statistical inverse problems, there are few results in a minimax framework as general as the one given above. Generally, only the L^2 case is considered, and under the polynomial decay assumption of the eigenvalues the rate $\zeta = \frac{s}{s+\nu+1/2}$ (named "regular" rate) appears frequently (see Cavalier and Tsybakov (2002)). For more general L^p losses, only the case of deconvolution in a periodic setting (up to our knowledge) has been studied in Johnstone, Kerkyacharian, Picard, and Raimondo (2004) and Willer (2005), and elbow effects appear, with a second rate named "sparse". It is interesting to draw a parallel between such a problem, where classical wavelets are widely used tools, and polynomial type problems, which require needlets.

For the deconvolution problem, minimax rates have been established for all $L^p([0, 1], dx)$ losses ($1 < p < \infty$) and over balls of a Besov space characterized by parameters $\pi \geq 1$, $s \geq 1/\pi$, $r \geq 1$ as above. Then the rates are given as in Theorem 1 and 2 (up to the logarithmic factors) with ζ replaced by:

$$\zeta = \min\left\{\zeta_{\text{regular}} = \frac{s}{s + \nu + 1/2}, \quad \zeta_{\text{sparse}} = \frac{s - 1/\pi + 1/p}{s + \nu + 1/2 - 1/\pi}\right\}.$$

Then the deconvolution setting appears as a critical case of the Jacobi setting, if we set $\alpha = \beta = -\frac{1}{2}$. More generally if we set $\alpha = \beta > -\frac{1}{2}$ we can draw the cartography of the regular and sparse zones with respect to (p, π) (see figure 1), as was done in Härdle, Kerkyacharian, Picard, and Tsybakov (1998) in the direct observation case. In the deconvolution case (i.e. the "wavelet scenario") the separation between the zones is linear, whereas in the "Jacobi scenario" the critical case is more complicated. So in that scenario we find new rates, and note that this novelty is not an artifact stemming from the weights on the space, since in the Lebesgue case the rates for the Jacobi scenario (i.e. $\alpha = \beta = 0$) do not coincide with those of the wavelet scenario. Thus the origin of the differences lies in the polynomial structure of the inverse problems, in opposition to the "Fourier" structure of the problems usually treated by first generation wavelet methods.

2.4 Application to the Wicksell's problem

The Jacobi-type inverse models considered in this paper find applications in

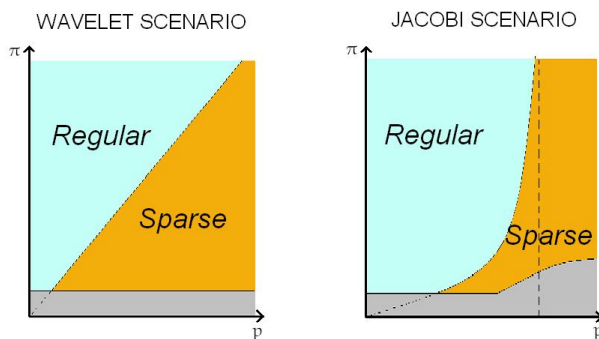


Figure 1: Cartography of the regular and sparse zones with respect to (p, π) in the deconvolution case (left) and in the Jacobi case if $\alpha = \beta$ (right)

practice, in particular with the well known Wicksell's problem (Wicksell (1925)), which corresponds to the following situation. Suppose a population of spheres is embedded in a medium. The spheres have radii that may be assumed to be drawn independently from a density f . A random plane slice is taken through the medium, and some spheres are intersected by it. They furnish circles, the radii of which yield the points of observation Y_1, \dots, Y_n , as illustrated in Figure 2. The unfolding problem is then to determine the density of the sphere radii from the observed circle radii. This problem arises in medicine, where the spheres might be tumors in an animal's liver (Nychka, Wahba, Goldfarb, and Pugh (1984)), as well as in numerous other contexts (biological, engineering, etc.) see for instance Cruz-Orive (1983).

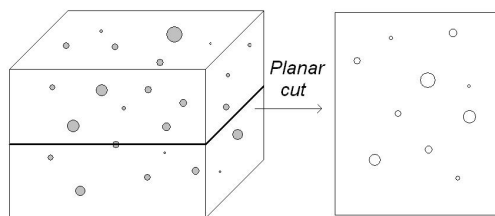


Figure 2: Wicksell's problem: observation of radii of disks after a planar cut of spheres

In this article we consider this problem in the white noise framework (considerations about the application to the density framework are made in Chapter 5 of Willer (2006)). We use the singular value decomposition established in Johnstone

and Silverman (1991), where the Wicksell's problem corresponds to the following operator: $U^* = L^2([0, 1], \mu^*(x)dx)$, $\mu^*(x) = (4x)^{-1}$, $V = L^2([0, 1], \lambda(y)dy)$, $\lambda(y) = 4\pi^{-1}(1 - y^2)^{1/2}$, and $Kf(y) = \frac{\pi}{4}y(1 - y^2)^{-1/2} \int_y^1 (x^2 - y^2)^{-1/2} f(x) d\mu^*(x)$. In this case K^*K has the following eigenvalues and eigenfunctions:

$$\begin{aligned} b_k &= \frac{\pi}{16}(1 + k)^{-1/2}, \\ e_k(x) &= 4(k + 1)^{1/2} x^2 P_k^{0,1}(2x^2 - 1), \end{aligned}$$

where $P_k^{0,1}$ is the k th degree Jacobi polynomial of type $(0, 1)$. Then we are (up to changes in the variables) in the framework considered in this paper. Thus Theorem 1 and 2 establish the rates for the minimax risk R_ϵ^{Wick} of the Wicksell problem, considered in the framework of Johnstone and Silverman (1991), with white noise perturbations. After taking into account the changes of variables (hence the notation \tilde{B} instead of B), and neglecting $\log(1/\epsilon)$ factors, we have: $R_\epsilon^{\text{Wick}}[\tilde{B}_{\pi,r}^s(M), \mathbb{L}^p([0, 1], x^{3-p}dx)] \asymp \epsilon^\zeta$, where:

$$\zeta = \min\left\{\frac{s}{s+1}, \frac{s - 2(\frac{1}{\pi} - \frac{1}{p})}{s + \frac{3}{2} - \frac{2}{\pi}}, \frac{s - 4(\frac{1}{\pi} - \frac{1}{p})}{s + \frac{5}{2} - \frac{4}{\pi}}\right\}.$$

Thus we find rates which are new in the literature, and we establish the quasi optimality of the NeedVD estimator. However, of course, other formulations of the Wicksell problem have been proposed, with some other results: a minimax study can be found in Golubev and Levit (1998) for the estimation of the distribution function associated to f , and in Antoniadis, Fan, and Gijbels (2001) convergence rates are established for the estimation of a probability distribution function closely related to f .

3. Construction and properties of needlets

3.1 Construction of Jacobi needlets

In this section we recall briefly the construction of Jacobi needlets introduced by Petrushev and Xu (2005), for more details we refer the reader to that paper. We denote by (P_k) the classical Jacobi polynomials of type (α, β) normalized in $U = L^2([-1, 1], \mu)$ where $d\mu(x) \asymp (1 - x)^\alpha(1 + x)^\beta dx$; $\alpha, \beta > -1/2$.

The first step of the construction consists in a Littlewood-Paley decomposition. Let $\alpha(\cdot)$ be a C^∞ function supported in $[-2, -\frac{1}{2}] \cup [\frac{1}{2}, 2]$ such that

$\sum_{j \geq 0} \alpha^2(x/2^j) = 1$, $\forall |x| \geq 1$. Moreover we add the condition: $\alpha(x) > c > 0$ for $3/4 \leq x \leq 7/4$ (so as to use results established in Kerkyacharian, Picard, Petrushev, and Willer (2006)). Then we introduce the following family of functions, which appear as the kernels of the operators performing the Littlewood-Paley decomposition:

$$\forall j \in \mathbb{N}, \Lambda_j(x, y) = \sum_{k \in \mathbb{N}} \alpha(k/2^j) P_k(x) P_k(y).$$

The second step is a quadrature formula. For $j \in \mathbb{N}$ we set $\mathbb{Z}_j = \{\eta_k : k = 1, 2, \dots, 2^j\}$, where $\eta_k = \cos \theta_{j,k}$ are the zeros of the Jacobi polynomial P_{2^j} ordered so that $\eta_1 > \eta_2 > \dots > \eta_{2^j}$, and hence $0 < \theta_{j,1} < \theta_{j,2} < \dots < \theta_{j,2^j} < \pi$. It is well known that $\theta_{j,k} \sim \frac{k\pi}{2^j}$ (cf Szegő (1975)). Let Π_n denote the space of all polynomials of degree inferior to n . Then the points of \mathbb{Z}_j serve as knots of the Gaussian quadrature which is exact for all polynomials from $\Pi_{2^{j+1}-1}$, that is,

$$\int_{[-1,1]} P d\mu = \sum_{\eta_k \in \mathbb{Z}_j} b_{j,\eta_k} P(\eta_k), \quad \forall P \in \Pi_{2^{j+1}-1},$$

where the coefficients $b_{j,\eta_k} > 0$ are the Christoffel numbers (Szegő (1975)) and $b_{j,\eta_k} \asymp 2^{-j} \omega_{\alpha,\beta}(2^j; \eta_k)$ with $\omega_{\alpha,\beta}(2^j; x) := (1-x+2^{-2j})^{\alpha+1/2} (1+x+2^{-2j})^{\beta+1/2}$.

We finally define the Jacobi needlets as

$$\forall j \in \mathbb{N}, k \in \{1, \dots, 2^j\}, \quad \psi_{j,\eta_k}(x) = \sqrt{b_{j,\eta_k}} \Lambda_{2^j}(x, \eta_k).$$

In view of the support of α , the needlets depend on the Jacobi polynomials in the following way: $\psi_{j,\eta}(x) = \sum_{l=2^{j-2}+1}^{2^j-1} c_{j,\eta,l} P_l(x)$, with coefficients $c_{j,\eta,l} = \alpha(1/2^{j-1}) P_l(\eta) \sqrt{b_{j,\eta}}$. Some examples of needlets are given in figure 3. Note that setting $\alpha \neq \beta$ introduces some dissymmetry in the function, which presents more variations in the interval corresponding to the highest parameter $\max(\alpha, \beta)$.

3.2 Properties of Jacobi needlets

In this section we give a list of some useful results on the needlets established in previous papers, and we give new properties of needlets that will be needed to establish Theorem 2.

Wavelet-like properties: First of all, the needlets form a tight frame:

$$\forall f \in \mathbb{H}, \quad f = \sum_{j \in \mathbb{N}} \sum_{\eta \in \mathbb{Z}_j} \langle f, \psi_{j,\eta} \rangle \psi_{j,\eta} \quad \text{and} \quad \|f\|^2 = \sum_{j \in \mathbb{N}} \sum_{\eta \in \mathbb{Z}_j} |\langle f, \psi_{j,\eta} \rangle|^2.$$

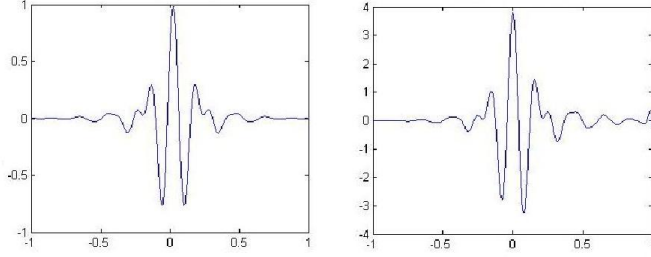


Figure 3: A Jacobi needlet of type $(\alpha, \beta) = (0, 0)$ (left) and $(\alpha, \beta) = (1, 0)$ (right)

Secondly each needlet ψ_{j, η_k} is concentrated on a small interval centered on η , as established in Petrushev and Xu (2005):

Theorem 3. *For any $l \geq 1$ there exists a constant $C_l > 0$ such that*

$$|\psi_{j, \eta_k}(\cos \theta)| \leq C_l \frac{1}{\sqrt{\omega_{\alpha, \beta}(2^j, \cos \theta)}} \frac{2^{j/2}}{(1 + 2^j |\theta - \frac{\pi k}{2^j}|)^l}, \quad 0 \leq \theta \leq \pi.$$

This almost exponential concentration property implies wavelet-like inequalities for the L^p norms of linear combinations of needlets. This is Theorem 4, established in Kerkyacharian, Picard, Petrushev, and Willer (2006):

Theorem 4. *Let $0 < p < \infty$. Then there exists a constant $C_p > 0$ such that for any collection of numbers $\{\lambda_k : k = 1, 2, \dots, 2^j\}$, $j \geq 0$,*

$$\left\| \sum_{k=1}^{2^j} \lambda_k \psi_{j, \eta_k} \right\|_{\mathbb{L}^p(\mu)}^p \leq C_p \sum_{k=1}^{2^j} |\lambda_k|^p \|\psi_{j, \eta_k}\|_{\mathbb{L}^p(\mu)}^p.$$

Differences with first generation wavelets: Needlets are not issued from a translation/dilatation scheme, hence major differences with classical wavelets. Let us for example describe the needlets at a given resolution level j . First they are not distributed uniformly on the interval, but around the η_k s. Second they behave quite differently depending on their locations η in the interval, which is reflected in Theorem 3 by the variations of the function $\omega_{\alpha, \beta}(2^j, \cdot)$. This is illustrated in figure 4: for a given resolution j , "edge" needlets have different shapes than "middle" needlets, and the L^p norms are not constant with respect to η (except arguably for $p = 2$). More precisely concerning L^p norms, the following bounds

have been established in Petrushev and Xu (2005) (for the upper bounds) and in Kerkyacharian, Picard, Petrushev, and Willer (2006) (for the lower bounds). They play an important role for the proofs of Theorem 1 and 2.

Theorem 5. $\forall 0 < p \leq \infty, \forall j \in \mathbb{N}$, we have up to scalars depending only on p :

$$\begin{aligned} \forall k = 1, \dots, 2^{j-1}, \quad & \|\psi_{j, \eta_k}\|_p \asymp \left(\frac{2^{j(\alpha+1)}}{k^{\alpha+1/2}} \right)^{1-2/p}, \\ \forall 2^{j-1} < k \leq 2^j, \quad & \|\psi_{j, \eta_k}\|_p \asymp \left(\frac{2^{j(\beta+1)}}{(1 + (2^j - k))^{\beta+1/2}} \right)^{1-2/p}. \end{aligned}$$

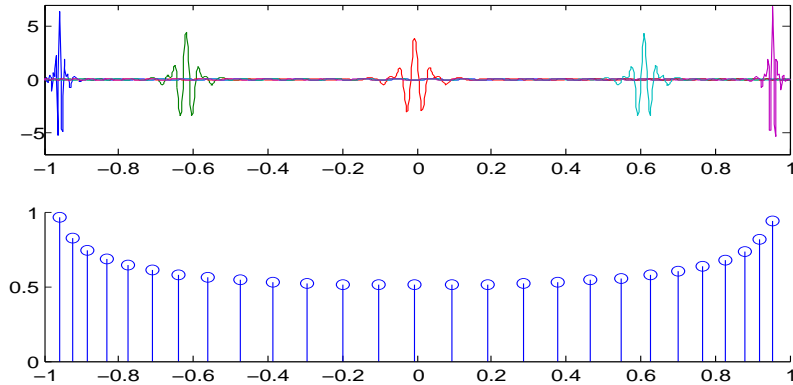


Figure 4: For a given resolution j : some of the needlets ψ_{j, η_k} (above), and the values of all the L^3 norms (below) when η_k varies

Moreover unlike first generation wavelets, needlets do not form an orthonormal basis, but only a redundant frame. This leads to some specific difficulties for the study of the lower bound of the minimax risk. So we need to add, to the previous list already used in Kerkyacharian, Picard, Petrushev, and Willer (2006), two new results which are proved in the appendix.

First we need to upper bound the scalar products between needlets. This is given by Lemma 1.

Lemma 1. *We have:*

1. $\forall j, j', k, l$ such that $|j' - j| \geq 2$, $\langle \psi_{j, \eta_k}, \psi_{j', \eta_l} \rangle = 0$.

2. $\forall \zeta > 0, \exists c_\zeta$ such that $\forall j, j', k, l$ with $|j' - j| \leq 1$: $|\langle \psi_{j, \eta_k}, \psi_{j', \eta_l} \rangle| \leq \frac{c_\zeta}{(1 + |k - 2^{j-j'} - l|)^\zeta}$.

Secondly we need to lower bound the L^p norm of linear combinations of needlets. Note that a result as general as the upper bound of Theorem 4 is impossible. Indeed, for instance with the non null coefficients $\sqrt{b_{j, \eta_k}}$ introduced in the definition of the needlets, one can check that: $\sum_{k=1}^{2^j} \sqrt{b_{j, \eta_k}} \psi_{j, \eta_k} = 0$. However we establish the following result for needlets with a large enough distance between the indexes of the η 's, in the case where p is an even integer:

Theorem 6. *Let $p \in 2\mathbb{N}^*$. Then there exists a constant $c_p > 0$ and an integer n_p such that for any collection of numbers $\{\lambda_k : k \in I_j\}$, $j \geq 0$, where $I_j \subset \{1, 2, \dots, 2^j\}$ and $k, l \in I_j, k \neq l \implies |k - l| \geq n_p$,*

$$\left\| \sum_{k \in I_j} \lambda_k \psi_{j, \eta_k} \right\|_{L^p(\mu)}^p \geq c_p \sum_{k \in I_j} |\lambda_k|^p \|\psi_{j, \eta_k}\|_{L^p(\mu)}^p.$$

4. Proof of the main result

4.1 Scheme of the proof

The proof of Theorem 2 requires well known methods for minimax lower bounds, as available in Tsybakov (2004), combined with new tools (i.e. needlets). We use Theorem 5.2 in Tsybakov (2004), which involves the Kullback-Leibler divergence $\mathcal{K}(P, Q)$ between two probability measures P and Q , defined by:

$$\mathcal{K}(P, Q) = \begin{cases} \int \ln\left(\frac{dP}{dQ}\right) dP, & \text{if } P \ll Q; \\ +\infty, & \text{otherwise.} \end{cases}$$

Changing the notations, and replacing slightly the conditions so as to include the case $m = 1$ (the result remains true using $\tau = 1/\sqrt{m+1}$ instead of $\tau = 1/\sqrt{m}$ in the proof), this theorem states that:

Theorem 7. *Assume there exist $m + 1$ functions f_0, \dots, f_m (with $m \geq 1$) satisfying the three following conditions:*

- **Condition (i):** for all $i \in \{0, 1, \dots, m\}$, $f_i \in B_{\pi, r}^s(M)$,
- **Condition (ii):** for all $i \neq j$, $\|f_i - f_j\|_p^p \geq 2\delta$ for some $\delta > 0$,

- **Condition (iii')**: for all $i \in \{1, \dots, m\}$, $P_{f_i} \ll P_{f_0}$ and $\frac{1}{m} \sum_{i \geq 1} \mathcal{K}(P_{f_i}, P_{f_0}) \leq \theta \log(M + 1)$, where $0 < \theta < \frac{1}{8}$ and P_f denotes the probability distribution of the process Y under the hypothesis f .

Then $\inf_{\hat{f}} \sup_{f \in \mathcal{B}_{\pi, r}^s(\mathcal{M})} P_f(\|\hat{f} - f\|_p^p \geq \delta) \geq \pi_0$, where π_0 is a positive universal constant.

We use this theorem by building several sets of hypotheses $\{f_i, i = 0, 1, \dots, m\}$ satisfying the three conditions. Then using Chebychev's inequality we have:

$$\inf_{\hat{f}} \sup_{f \in \mathcal{B}_{\pi, r}^s(\mathcal{M})} E_f \|\hat{f} - f\|_p^p \geq \pi_0 \delta.$$

With an appropriate choice of three sets $\{f_i, i = 0, 1, \dots, m\}$ depending on the level of noise ϵ , δ yields the three expected rates.

Let us precise condition (iii') in model 1. Let $I = [a, b]$ (the case $I = [a, b]$ is similar). If we define the variables $\tilde{Y}(w) = Y(w(\cdot - a)/\sqrt{\lambda(\cdot)})$ and $\tilde{\xi}(w) = \xi(w(\cdot - a)/\sqrt{\lambda(\cdot)})$ for all $w \in \tilde{V} = \mathbb{L}^2([0, b - a], dx)$ then model 1 is equivalent to: $\tilde{Y}(w) = (Kf(\cdot + a)\sqrt{\lambda(\cdot + a)}, w)_{\tilde{V}} + \epsilon \tilde{\xi}(w)$, which is equivalent to the stochastic equation: $\forall t \in [0, b - a]$, $d\tilde{Y}_t = Kf(\cdot + a)\sqrt{\lambda(\cdot + a)} + \epsilon dW_t$ where $(W_t)_{t \geq 0}$ denotes the standard Wiener process. Then using Girsanov's formula, or theorem 7.18 from Lipster and Shiryaev (1977), for all $f, g \in \mathcal{U}$ P_f is absolutely continuous with respect to P_g , and under the hypothesis g the likelihood ratio $\Lambda_\epsilon(f, g) := \frac{dP_f}{dP_g}(Y)$ is distributed as: $\log \Lambda_\epsilon(f, g) \sim \mathcal{N}(-\frac{1}{2} \|\frac{K(f-g)}{\epsilon}\|_{\tilde{V}}^2, \|\frac{K(f-g)}{\epsilon}\|_{\tilde{V}})$. Thus

$$\mathcal{K}(P_f, P_g) = E_f \ln(\Lambda_\epsilon(f, g)) = -E_f \log(\Lambda_\epsilon(g, f)) = \frac{1}{2} \|\frac{K(f-g)}{\epsilon}\|_{\tilde{V}}^2.$$

Then condition (iii') can be replaced by the sufficient condition (iii):

Condition (iii): $f_0 = 0$ and for all $i \in \{1, \dots, m\}$, $\|Kf_i\|_{\tilde{V}}^2 \leq \theta \log(M + 1)\epsilon^2$ where $0 < \theta < \frac{1}{4}$.

4.2 Sparse cases

The sparse rates $\mu(\alpha)$ and $\mu(\beta)$ are obtained respectively by applying Theorem 7 to the following set of functions: $\{f_0 = 0, f_1 = \gamma\psi_{j_0, \eta_1}\}$, and $\{f_0 = 0, f_1 = \gamma\psi_{j_1, \eta_{2j_1}}\}$, for some parameters γ, j_0 and j_1 chosen so as to satisfy conditions (i) to (iii). We detail only the proof for $\mu(\alpha)$ (the proof for $\mu(\beta)$ is similar).

Condition (i) is satisfied if $u_j := 2^{js}(\sum_{\eta \in \mathbb{Z}_j} |\langle f_1, \psi_{j, \eta} \rangle|^\pi \|\psi_{j, \eta}\|_\pi^\pi)^{1/\pi}$ belongs to $l^r(M)$, where $f_1 = \gamma\psi_{j_0, \eta_1}$. Using the first part of Lemma 1, $u_j = 0$ whenever $|j - j_0| \geq 2$. So in the sequel we assume that $j \in \{j_0 - 1, j_0, j_0 + 1\}$, and the l^r norm of (u_j) is bounded by a constant M (independent of $\gamma > 0$ and j_0) if for instance $u_j \leq 3^{-\frac{1}{r}}M$. We have: $u_j^\pi = 2^{j\pi s} \gamma^\pi \sum_{\eta \in \mathbb{Z}_j} |\langle \psi_{j_0, \eta_1}, \psi_{j, \eta} \rangle|^\pi \|\psi_{j, \eta}\|_\pi^\pi \leq c(I_1 + I_2)$, with, using the bound of Theorem 4:

$$I_1 = 2^{j[\pi s + (\pi-2)(\alpha+1)]} \gamma^\pi \sum_{k=1}^{2^{j-1}} |\langle \psi_{j_0, \eta_1}, \psi_{j, \eta} \rangle|^\pi k^{-(\pi-2)(\alpha+1/2)},$$

$$I_2 = 2^{j[\pi s + (\pi-2)(\beta+1)]} \gamma^\pi \sum_{k=2^{j-1}+1}^{2^j} |\langle \psi_{j_0, \eta_1}, \psi_{j, \eta} \rangle|^\pi (2^j - k + 1)^{-(\pi-2)(\beta+1/2)}.$$

Using the second part of Lemma 1, we have for any ζ : $|\langle \psi_{j_0, \eta_1}, \psi_{j, \eta_k} \rangle| \leq c \frac{1}{k^\zeta}$. Thus choosing any $\zeta > \frac{-(\pi-2)(\alpha+1/2)+1}{\pi}$, we obtain: $I_1 \leq c 2^{j[\pi s + (\pi-2)(\alpha+1)]} \gamma^\pi$. Moreover $\sum_{k=1}^{2^{j-1}} \frac{(2^j - k + 1)^{-(\pi-2)(\beta+1/2)}}{k^\zeta \pi} \leq c 2^{-\zeta \pi j} 2^{j[1 - (\pi-2)(\beta+1/2)]_+}$, so for a large enough ζ : $I_2 \leq c 2^{j[\pi s + (\pi-2)(\beta+1) - \zeta \pi + [1 - (\pi-2)(\beta+1/2)]_+]} \gamma^\pi \leq c I_1$. Thus we have for all $j \in \{j_0 - 1, j_0, j_0 + 1\}$: $u_j^\pi \leq c 2^{j_0[\pi s + (\pi-2)(\alpha+1)]} \gamma^\pi$, and condition (i) is satisfied if, for a small enough c depending on M :

$$\gamma \leq c 2^{-j_0[s + (1 - \frac{2}{\pi})(\alpha+1)]}.$$

Condition (ii), using theorem 5, is fulfilled with: $\delta \asymp \gamma^p 2^{j_0(p-2)(\alpha+1)}$.

Condition (iii) is satisfied if: $\int_I (\frac{K(\gamma\psi_{j_0, \eta_1})(t)}{\epsilon})^2 d\lambda(t) \leq C$. We have $\psi_{j_0, \eta}(x) = \sum_{l=2^{j-2}+1}^{2^j-1} c_{j, \eta, l} P_l(x)$, thus $K\psi_{j_0, \eta}(x) = \sum_l b_l c_{j, \eta, l} U_l(x)$, and:

$$\|K(\psi_{j_0, \eta_1})\|_V^2 = \sum_l [b_l c_{j, \eta, l}]^2 \asymp 2^{-2\nu j_0} \sum_l [c_{j, \eta, l}]^2 = 2^{-2\nu j_0} \|\psi_{j_0, \eta_1}\|_U^2 \leq C 2^{-2\nu j_0}.$$

So condition (iii) is satisfied if: $\frac{\gamma 2^{-\nu j_0}}{\epsilon} \leq c$.

In view of the three conditions, we set $\gamma = c\epsilon 2^{\nu j_0}$ with a small enough c , and $2^{j_0} \asymp \epsilon \frac{1}{s+\nu+(1-\frac{2}{\pi})(\alpha+1)}$. Then $\delta \asymp \epsilon \frac{p[s+2(\frac{1}{p}-\frac{1}{\pi})(\alpha+1)]}{s+\nu+(1-\frac{2}{\pi})(\alpha+1)}$ gives the sparse lower bound.

4.3 Regular case

Let m be an integer such that $2^m \geq n_2$, where n_2 is the integer from Theorem 6 in the case $p = 2$. For some parameters γ and $j_0 \geq m + 1$ chosen further, we consider for $\epsilon \in \{0, 1\}^{2^{j_0-m-1}}$ the $2^{2^{j_0-m-1}}$ functions:

$$f_\epsilon = \gamma \sum_{k=1}^{2^{j_0-m-1}} \epsilon_k k^\delta \psi_{j_0, \eta_{2^m k}},$$

for some δ satisfying: $\delta > \max[1, \alpha + 1/2, (1 - \frac{2}{\pi})(\alpha + \frac{1}{2}) - \frac{1}{\pi}]$. We only keep some of these functions. By Varshamov-Gilbert theorem (see for instance Tsybakov (2004)), there exists a subset $E_{j_0} = \{\epsilon^0, \dots, \epsilon^{T_{j_0}}\}$ of $\{0, 1\}^{2^{j_0-m-1}}$ and two constants $c > 0$, $\rho > 0$ such that $\forall 0 \leq u < v \leq T_{j_0}$:

$$\sum_{k=1}^{2^{j_0-m-1}} |\epsilon_k^u - \epsilon_k^v| \geq c 2^{j_0}, \quad T_{j_0} \geq \exp(\rho 2^{j_0}) \quad \text{and} \quad f_{\epsilon^0} = 0.$$

In the sequel we consider the set $\{f_\epsilon, \quad \epsilon \in E_{j_0}\}$.

Condition (i): for $\epsilon \in E_{j_0}$, let $u_j := 2^{js} (\sum_{\eta \in \mathbb{Z}_j} |\langle f_\epsilon, \psi_{j, \eta} \rangle|^\pi \|\psi_{j, \eta}\|^\pi)^\pi$. Once again $u_j = 0$ whenever $|j - j_0| \geq 2$. Now let $j \in \{j_0 - 1, j_0, j_0 + 1\}$. Then we have: $u_j^\pi \leq c(I_1 + I_2)$, with:

$$I_1 = 2^{j[\pi s + (\pi-2)(\alpha+1)]} \gamma^\pi \sum_{k=1}^{2^{j-1}} k^{-(\pi-2)(\alpha+1/2)} \left(\sum_{l=1}^{2^{j_0-1}} l^\delta |\langle \psi_{j_0, \eta_l}, \psi_{j, \eta_k} \rangle| \right)^\pi,$$

$$I_2 = 2^{j[\pi s + (\pi-2)(\beta+1)]} \gamma^\pi \sum_{k=2^{j-1}+1}^{2^j} (2^j - k + 1)^{-(\pi-2)(\beta+1/2)} \left(\sum_{l=1}^{2^{j_0-1}} l^\delta |\langle \psi_{j_0, \eta_l}, \psi_{j, \eta_k} \rangle| \right)^\pi.$$

Using Lemma 1 with some ζ given later, we have $|\langle \psi_{j_0, \eta_l}, \psi_{j, \eta_k} \rangle| \leq c \frac{1}{(1+|l-2^{j_0-j}k|)^\zeta}$. Then, for $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the largest integer smaller than x . We have:

$$\sum_{l \leq \lfloor 2^{j_0-j}k \rfloor} \frac{l^\delta}{(1+|l-2^{j_0-j}k|)^\zeta} \leq c k^\delta \sum_{l \leq \lfloor 2^{j_0-j}k \rfloor} \frac{1}{(1+\lfloor 2^{j_0-j}k \rfloor - l)^\zeta} \leq c k^\delta \sum_{l \geq 1} \frac{1}{l^\zeta} \leq c k^\delta,$$

for a large enough ζ . Moreover:

$$\begin{aligned} \sum_{l \geq \lfloor 2^{j_0-j} k \rfloor + 1} \frac{l^\delta}{(1 + |l - 2^{j_0-j} k|)^\zeta} &\leq \sum_{l \geq \lfloor 2^{j_0-j} k \rfloor + 1} \frac{l^\delta}{(l - \lfloor 2^{j_0-j} k \rfloor)^\zeta} = \sum_{l \geq 1} \frac{(l + \lfloor 2^{j_0-j} k \rfloor)^\delta}{l^\zeta} \\ &\leq c \sum_{l \geq 1} \frac{l^\delta + \lfloor 2^{j_0-j} k \rfloor^\delta}{l^\zeta} \leq Ck^\delta, \end{aligned}$$

for ζ large enough. To obtain the last line, we used the fact that $\delta \geq 1$. Thus $\sum_{l=1}^{2^{j_0-1}} \frac{l^\delta}{(1 + |l - 2^{j_0-j} k|)^\zeta} \leq ck^\delta$, and:

$$I_1 \leq c2^{j[\pi s + (\pi-2)(\alpha+1)]} \gamma^\pi \sum_{k=1}^{2^{j-1}} k^{-(\pi-2)(\alpha+1/2)} k^{\delta\pi} = c2^{j[s+\delta+\frac{1}{2}]} \gamma.$$

For I_2 remark that for any $k \in \{2^{j-1} + 1, \dots, 2^j\}$ and any $l \in \{1, \dots, 2^{j_0-1}\}$, we have: $|\frac{k}{2^j} - \frac{l}{2^{j_0}}| = \frac{k}{2^j} - \frac{l}{2^{j_0}} \geq |\frac{2^j-k}{2^j} - \frac{l}{2^{j_0}}|$. So for such a k , as previously: $\sum_{l=1}^{2^{j_0-1}} \frac{l^\delta}{(1 + |l - 2^{j_0-j} k|)^\zeta} \leq \sum_{l=1}^{2^{j_0-1}} \frac{l^\delta}{(1 + |l - 2^{j_0-j} (2^j - k)|)^\zeta} \leq c(2^j - k)^\delta$, and:

$$I_2 \leq c2^{j[\pi s + (\pi-2)(\beta+1)]} \gamma^\pi \sum_{k=2^{j-1}+1}^{2^j} (2^j - k + 1)^{-(\pi-2)(\beta+1/2)} (2^j - k + 1)^{\delta\pi} = c2^{j[s+\delta+\frac{1}{2}]} \gamma.$$

Finally we have $u_j \leq c2^{j[s+\delta+\frac{1}{2}]} \gamma$ so f_ε belongs to $B_{\pi,r}^s(M)$ if, with a small enough c depending on M :

$$\gamma \leq c2^{-j_0[s+\delta+\frac{1}{2}]}.$$

Condition (ii): for all $u, v \in E_{j_0}$ with $u \neq v$, $f_u - f_v = \sum_{k=1}^{2^{j_0-m-1}} \gamma(\varepsilon_k^u - \varepsilon_k^v) k^\delta \psi_{j_0, n_{2^m k}}$. So by Theorem 6 and Theorem 5, we have:

$$\|f_u - f_v\|_{\mathbb{U}}^2 \geq c\gamma^2 \sum_{k=1}^{2^{j_0-m-1}} (\varepsilon_k^u - \varepsilon_k^v)^2 k^{2\delta} = c\gamma^2 \sum_{\{k \mid \varepsilon_k^u \neq \varepsilon_k^v\}} k^{2\delta}.$$

Let $N_{u,v}$ denote the cardinal of the set $\{k \in \{1, \dots, 2^{j_0-m-1}\} \mid \varepsilon_k^u \neq \varepsilon_k^v\}$, then we have $N_{u,v} \geq c2^{j_0}$ and, since $\delta > 0$:

$$\|f_u - f_v\|_{\mathbb{U}}^2 \geq c\gamma^2 \sum_{k=1}^{N_{u,v}} k^{2\delta} = \gamma^2 N_{u,v}^{1+2\delta} \geq c\gamma^2 2^{j_0(1+2\delta)}. \quad (2)$$

Let us distinguish two cases. *Suppose* $2 < p < \infty$ and let $1/p + 1/q = 1$. By (2) and Hölder's inequality we have:

$$c2^{j_0(1+2\delta)} \leq \|f_u - f_v\|_{\mathbb{L}^2(\mu)}^2 \leq \|f_u - f_v\|_{\mathbb{L}^p(\mu)} \|f_u - f_v\|_{\mathbb{L}^q(\mu)}.$$

Using Theorem 4 and the fact that, under our assumptions, $q\delta - (q-2)(\alpha+1/2) > -1$, we have:

$$\|f_u - f_v\|_{\mathbb{L}^q(\mu)} \leq c\gamma 2^{j\frac{(q-2)}{q}(\alpha+1)} \left(\sum_{k=1}^{2^{j_0-m-1}} k^{q\delta - (q-2)(\alpha+1/2)} \right)^{1/q} \leq c'\gamma 2^{j_0(\frac{1}{2}+\delta)},$$

therefore $\|f_u - f_v\|_{\mathbb{L}^p(\mu)}^p \geq c\gamma^p 2^{j_0 p(\frac{1}{2}+\delta)}$.

Suppose now $1 < p < 2$, we have using (2):

$$c2^{j_0(1+2\delta)} \leq \|f_u - f_v\|_{\mathbb{L}^2(\mu)}^2 \leq \|f_u - f_v\|_{\mathbb{L}^p(\mu)}^p \|f_u - f_v\|_{\mathbb{L}^\infty(\mu)}^{2-p}.$$

From Theorem 3 we infer for all $0 \leq \theta \leq \pi/2$:

$$|\psi_{j_0, \eta_k}(\cos \theta)| \leq C \frac{2^{j_0(1+\alpha)}}{(1 + 2^{j_0}|\theta - \frac{k\pi}{2^{j_0}}|)^l} \frac{1}{(2^{j_0}\theta + 1)^{\alpha+1/2}},$$

so for l large enough: $|\psi_{j_0, \eta_k}(\cos \theta)| \leq C \frac{2^{j_0(1+\alpha)}}{k^{\alpha+1/2}} \frac{1}{(1+2^{j_0}|\theta - \frac{k\pi}{2^{j_0}}|)^2}$ and, since $\delta - (\alpha + 1/2) \geq 0$:

$$|f_u(\cos \theta) - f_v(\cos \theta)| \leq c\gamma 2^{j_0(\alpha+1)} \sum_{k=1}^{2^{j_0-m-1}} k^{\delta - (\alpha+1/2)} \frac{1}{(1 + 2^{j_0}|\theta - \frac{k\pi}{2^{j_0}}|)^2} \leq c'\gamma 2^{j_0(\frac{1}{2}+\delta)},$$

where in the last line we used the fact that for any θ , $\sum_{k=1}^{2^{j_0-m-1}} \frac{1}{(1+2^{j_0}|\theta - \frac{k\pi}{2^{j_0}}|)^2} \leq c \sum_{l=1}^{+\infty} \frac{1}{l^2}$. Similarly the same bound holds for any $\pi/2 \leq \theta \leq \pi$, thus we have: $\|f_u - f_v\|_{\mathbb{L}^\infty(\mu)} \leq c2^{j_0(\frac{1}{2}+\delta)}$, and once again: $\|f_u - f_v\|_{\mathbb{L}^p(\mu)}^p \geq c\gamma^p 2^{j_0 p(\frac{1}{2}+\delta)}$.

Condition (iii): we have $\sqrt{T_{j_0}} \geq \exp(\frac{\rho}{2}2^{j_0})$, so (iii) is satisfied if for all $\varepsilon^u \in E_{j_0}$, $\int_{\mathbb{I}} (\frac{K(f_u)(t)}{\varepsilon})^2 d\lambda(t) \leq c2^{j_0}$ for a small enough constant c . We have: $f_u = \sum_{k=1}^{2^{j_0-m-1}} \beta_{j_0, k} \psi_{j_0, \eta_{2^m k}} = \sum_{k=1}^{2^{j_0-m-1}} \sum_{l \in \mathbb{N}} \beta_{j_0, k} c_{j_0, \eta_k, l} P_l(x)$, with $\beta_{j_0, k} = \gamma \varepsilon_k^u k^\delta$. Thus:

$$\begin{aligned} \|K(f_u)\|_{\mathbb{L}^2(I, \lambda)}^2 &= \sum_l \left[\sum_{k=1}^{2^{j_0-m-1}} \beta_{j_0, k} b_l c_{j_0, \eta_k, l} \right]^2 \asymp 2^{-2\nu j_0} \sum_l \left[\sum_{k=1}^{2^{j_0-m-1}} \beta_{j_0, k} c_{j_0, \eta_k, l} \right]^2 \\ &= 2^{-2\nu j_0} \left\| \sum_{k=1}^{2^{j_0-m-1}} \beta_{j_0, k} \psi_{j_0, \eta_{2^m k}} \right\|_{\mathbb{L}^2(I, \mu)}^2 \leq c2^{-2\nu j_0} \sum_{k=1}^{2^{j_0-m-1}} \beta_{j_0, k}^2 \\ &\leq c2^{-2\nu j_0} \gamma^2 \sum_{k=1}^{2^{j_0-m-1}} k^{2\delta} = c2^{-2\nu j_0} \gamma^2 2^{(2\delta+1)j_0}. \end{aligned}$$

So finally we need: $\frac{2^{-\nu j_0} \gamma 2^{(s+\frac{1}{2})j_0}}{\epsilon} \leq C 2^{j_0/2}$, i.e. $\frac{2^{(\delta-\nu)j_0} \gamma}{\epsilon} \leq C$ with a small enough constant C .

In view of the three conditions, we set $2^{j_0} \asymp \epsilon^{-\frac{1}{s+\nu+\frac{1}{2}}}$ and $\gamma \asymp \epsilon^{\frac{s+\delta+\frac{1}{2}}{s+\nu+\frac{1}{2}}}$, and we obtain the lower bound: $\delta \asymp \epsilon^{\frac{ps}{s+\nu+\frac{1}{2}}}$.

5. Appendix

Proof of Theorem 6. Let $\mathbf{p} \in 2\mathbb{N}^*$ and $I_j \subset \{1, 2, \dots, 2^j\}$. We have the following upper bound: $\|(\sum_{k \in I_j} \lambda_k \psi_{j, \eta_k})\|_{\mathbb{L}^p(\mu)}^p = A + B$, where:

$$A = \sum_{k \in I_j} \lambda_k^p \|\psi_{j, \eta_k}\|_{\mathbb{L}^p(\mu)}^p,$$

$$B = \sum_{(\mathbf{p}_k)_{k \in I_j} \in \Lambda} \frac{p! \prod_{k \in I_j} \lambda_k^{p_k}}{\prod_{k \in I_j} p_k!} \int_{-1}^1 \left(\prod_{k \in I_j} \psi_{j, \eta_k}^{p_k}(x) \right) \mu(x) dx,$$

and $\Lambda = \{(\mathbf{p}_k)_{k \in I_j} \mid p_k \in \mathbb{N}, \sum_{k \in I_j} p_k = p \text{ and } \exists u \neq v \text{ such that } p_u > 0 \text{ and } p_v > 0\}$.

Let us introduce the functions $\varphi_{j,k}(x) = \frac{1}{\sqrt{\omega_{\alpha,\beta}(2^j, x)}} \frac{2^{j/2}}{(1+2^j|\arccos x - \frac{\pi k}{2^j}|)^{\frac{2}{s}}}$, for some $0 < s < \min\{1, \frac{p}{\alpha\sqrt{\beta+1}}\}$. For $(\mathbf{p}_k)_{k \in I_j} \in \Lambda$, we use Theorem 3 with $l = \frac{2}{s} + 1$ for every ψ_{j, η_k} , $k \in I_j$. There exists C such that:

$$\prod_{k \in I_j} |\psi_{j, \eta_k}(\cos \theta)|^{p_k} \leq C \prod_{k \in I_j} \varphi_{j,k}(\cos \theta)^{p_k} \prod_{k \in I_j} \frac{1}{(1+2^j|\theta - \frac{\pi k}{2^j}|)^{p_k}}.$$

Let $\mathbf{u}, \mathbf{v} \in I_j$, $\mathbf{u} \neq \mathbf{v}$ such that $p_u > 0$ and $p_v > 0$, and let $n_{\text{inf}} = \min_{k, l \in I_j, k \neq l} |k - l|$. We have:

$$\prod_{k \in I_j} (1+2^j|\theta - \frac{\pi k}{2^j}|)^{p_k} \geq (1+2^j|\theta - \frac{\pi \mathbf{u}}{2^j}|)(1+2^j|\theta - \frac{\pi \mathbf{v}}{2^j}|) \geq c|\mathbf{u} - \mathbf{v}| \geq c n_{\text{inf}}.$$

Thus we obtain:

$$\begin{aligned} \sum_{(\mathbf{p}_k)_{k \in I_j} \in \Lambda} \frac{p! \prod_{k \in I_j} |\lambda_k^{p_k}|}{\prod_{k \in I_j} p_k!} \prod_{k \in I_j} |\psi_{j, \eta_k}|^{p_k} &\leq \frac{C}{n_{\text{inf}}} \sum_{(\mathbf{p}_k)_{k \in I_j} \in \Lambda} \frac{p! \prod_{k \in I_j} |\lambda_k|^{p_k}}{\prod_{k \in I_j} p_k!} \prod_{k \in I_j} \varphi_{j, \eta_k}^{p_k} \\ &\leq C \frac{(\sum_{k \in I_j} |\lambda_k| \varphi_{j, \eta_k})^p}{n_{\text{inf}}}. \end{aligned}$$

Now let us proceed similarly to the sketch of the proof of theorem 4 available in Kerkyacharian, Picard, Petrushev, and Willer (2006). Let us recall the two main tools.

First, consider the maximal operator $(M_s f)(x) = \sup_{J \ni x} \left(\frac{1}{|J|} \int_J |f(u)|^s du \right)^{1/s}$, where the supremum is taken over all intervals $J \subset [-1, 1]$ which contain x , $s > 0$, and $|J|$ denotes the length of J . Then one can infer the following bound from the Fefferman-Stein maximal inequality (see Fefferman and Stein (1971) and Andersen and John (1980/81)). If $0 < p, r < \infty$ and $0 < s < \min\{p, r, \frac{p}{\alpha\sqrt{\beta+1}}\}$, then for any sequence of functions (f_k) on $[-1, 1]$

$$\left\| \left(\sum_k (M_s f_k)^r \right)^{1/r} \right\|_{\mathbb{L}^p(\mu)} \leq C \left\| \left(\sum_k |f_k|^r \right)^{1/r} \right\|_{\mathbb{L}^p(\mu)}.$$

Secondly set $\eta_0 = 1$, $\eta_{2^j+1} = -1$ and $\theta_{j,0} = 0$, $\theta_{j,2^j+1} = \pi$, respectively. Denote $I_k = [\frac{\eta_k + \eta_{k+1}}{2}, \frac{\eta_k + \eta_{k-1}}{2}]$ and put $H_k = h_k 1_{I_k}$ with $h_k = \left(\frac{2^j}{\omega_{\alpha,\beta}(2^j; \eta_k)} \right)^{1/2}$, where 1_{I_k} is the indicator function of I_k . Then $\|H_k\|_{\mathbb{L}^p(\mu)} \sim \|\psi_{j,\eta_k}\|_{\mathbb{L}^p(\mu)}$, and one shows in Kerkyacharian, Picard, Petrushev, and Willer (2006) that for any $s > 0$

$$\varphi_{j,\eta_k}(x) \leq c(M_s H_k)(x), \quad x \in [-1, 1], \quad \forall k = 1, 2, \dots, 2^j, j \geq 0.$$

We use these two results, with $f_k = H_k$ and $r = 1$. Noticing that the (H_k) have disjoint supports, we obtain:

$$\begin{aligned} \left\| \sum_{k=1}^{2^j} |\lambda_k| \varphi_{j,\eta_k} \right\|_{\mathbb{L}^p(\mu)}^p &\leq C \left\| \sum_{k=1}^{2^j} |\lambda_k| H_k \right\|_{\mathbb{L}^p(\mu)}^p = C \sum_{k=1}^{2^j} |\lambda_k|^p \|H_k\|_{\mathbb{L}^p(\mu)}^p \\ &\leq C' \sum_{k=1}^{2^j} |\lambda_k|^p \|\psi_{j,\eta_k}\|_{\mathbb{L}^p(\mu)}^p. \end{aligned}$$

So finally there exists $C > 0$ such that $|B| \leq C \frac{A}{n_{\text{inf}}}$, and if we impose the following condition on I_j : $n_{\text{inf}} \geq 2C$, then we obtain $|B| \leq \frac{1}{2}A$, and thus:

$$\left\| \left(\sum_{k \in I_j} \lambda_k \psi_{j,\eta_k} \right) \right\|_{\mathbb{L}^p(\mu)}^p \geq \frac{1}{2} \sum_{k \in I_j} \lambda_k^p \|\psi_{j,\eta_k}\|_{\mathbb{L}^p(\mu)}^p.$$

□

Proof of Lemma 1. As indicated previously, the needlets are defined as: $\psi_{j,\eta} = \sum_{l=2^j-2+1}^{2^j-1} c_{j,\eta,l} P_l(x)$, with coefficients $c_{j,\eta,l} = a(l/2^{j-1}) P_l(\eta) \sqrt{b_{j,\eta}}$. So if $|j' - j| \geq 2$ then $\{2^{j-2} + 1, \dots, 2^j - 1\} \cap \{2^{j'-2} + 1, \dots, 2^{j'} - 1\} = \emptyset$, and $\langle \psi_{j,\eta_k}, \psi_{j',\eta_l} \rangle = 0$, $\forall (k, l)$.

For the second part of the lemma we use Theorem 3. For any δ there exists c_δ such that for all j, k :

$$|\psi_{j,\eta_k}(\cos \theta)| \leq c_\delta \frac{1}{\sqrt{\omega_{\alpha,\beta}(2^j, \cos \theta)}} \frac{2^{j/2}}{(1 + 2^j |\theta - \frac{\pi k}{2^j}|)^\delta}, \quad 0 \leq \theta \leq \pi.$$

We recall that $\omega_{\alpha,\beta}(x) = (1-x)^\alpha(1+x)^\beta$, and $\omega_{\alpha,\beta}(2^j; x) = (1-x+2^{-2j})^{\alpha+1/2}(1+x+2^{-2j})^{\beta+1/2}$. For a given $\zeta > 0$ and j, j', k, l such that $|j' - j| \leq 1$, we use this inequality for $|\psi_{j,\eta_k}|$ with $\delta = \zeta + 2$ and for $|\psi_{j',\eta_l}|$ with $\delta = \zeta$. Noticing that $\omega_{\alpha,\beta}(2^j, \cos \theta) \asymp \omega_{\alpha,\beta}(2^{j'}, \cos \theta)$ we obtain:

$$\begin{aligned} |\langle \psi_{j,\eta_k}, \psi_{j',\eta_l} \rangle| &\leq c 2^j \int_0^\pi \frac{\omega_{\alpha,\beta}(\cos \theta)}{\omega_{\alpha,\beta}(2^j, \cos \theta)} \frac{\sin \theta d\theta}{(1 + 2^j |\theta - \frac{\pi k}{2^j}|)^{\zeta+2} (1 + 2^{j'} |\theta - \frac{\pi l}{2^{j'}}|)^\zeta} \\ &\leq c \frac{I_{j,k,\alpha,\beta}}{(\min_{0 \leq \theta \leq \pi} f_{j,j',k,l}(\theta))^\zeta}, \end{aligned}$$

with $f_{j,j',k,l}(\theta) = (1 + 2^j |\theta - \frac{\pi k}{2^j}|)(1 + 2^{j'} |\theta - \frac{\pi l}{2^{j'}}|)$, $0 \leq \theta \leq \pi$, and $I_{j,k,\alpha,\beta} = 2^j \int_0^\pi \frac{\omega_{\alpha,\beta}(\cos \theta)}{\omega_{\alpha,\beta}(2^j, \cos \theta)} \frac{\sin \theta d\theta}{(1 + 2^j |\theta - \frac{\pi k}{2^j}|)^2}$.

First we have: $\min_{0 \leq \theta \leq \pi} f_{j,j',k,l}(\theta) = \min\{f_{j,j',k,l}(\frac{\pi k}{2^j}), f_{j,j',k,l}(\frac{\pi l}{2^{j'}})\} \geq 1 + \frac{\pi}{2^{j-j'} \Gamma} |k - 2^{j-j'} l| \geq c(1 + |k - 2^{j-j'} l|)$. Secondly let us divide $I_{j,k,\alpha,\beta}$ into two terms: $I_{j,k,\alpha,\beta} = I_{j,k,\alpha,\beta}^1 + I_{j,k,\alpha,\beta}^2$, with:

$$\begin{aligned} I_{j,k,\alpha,\beta}^1 &= 2^j \int_0^{\frac{\pi}{2}} \frac{\omega_{\alpha,\beta}(\cos \theta)}{\omega_{\alpha,\beta}(2^j, \cos \theta)} \frac{\sin \theta d\theta}{(1 + 2^j |\theta - \frac{\pi k}{2^j}|)^2}, \\ I_{j,k,\alpha,\beta}^2 &= 2^j \int_{\frac{\pi}{2}}^\pi \frac{\omega_{\alpha,\beta}(\cos \theta)}{\omega_{\alpha,\beta}(2^j, \cos \theta)} \frac{\sin \theta d\theta}{(1 + 2^j |\theta - \frac{\pi k}{2^j}|)^2} \\ &= 2^j \int_0^{\frac{\pi}{2}} \frac{\omega_{\alpha,\beta}(-\cos \theta)}{\omega_{\alpha,\beta}(2^j, -\cos \theta)} \frac{\sin \theta d\theta}{(1 + 2^j |\pi - \theta - \frac{\pi k}{2^j}|)^2} \\ &= 2^j \int_0^{\frac{\pi}{2}} \frac{\omega_{\beta,\alpha}(\cos \theta)}{\omega_{\beta,\alpha}(2^j, \cos \theta)} \frac{\sin \theta d\theta}{(1 + 2^j |\theta - \frac{\pi(2^j - k)}{2^j}|)^2} \\ &= I_{j,2^j - k, \beta, \alpha}^1. \end{aligned}$$

We have: $\sin \theta \omega_{\alpha, \beta}(\cos \theta) = \sin \theta (2 \sin^2(\theta/2))^\alpha (2 \cos^2(\theta/2))^\beta \leq c_1 \theta^{2\alpha+1}$, for all $0 \leq \theta \leq \frac{\pi}{2}$, and:

$$\omega_{\alpha, \beta}(2^j; \cos \theta) = (2 \sin^2(\theta/2) + 2^{-2j})^{\alpha+1/2} (2 \cos^2(\theta/2) + 2^{-2j})^{\beta+1/2} \geq c_2 \theta^{2\alpha+1}.$$

Thus $I_{j, k, \alpha, \beta}^1 \leq c 2^j \int_0^{\frac{\pi}{2}} \frac{d\theta}{(1+2^j|\theta-\frac{\pi k}{2^j}|)^2} \leq c \int_0^{\frac{\pi 2^j}{2}} \frac{d\theta}{(1+|\theta-\pi k|)^2} \leq C$, since $\int_{-\infty}^{+\infty} \frac{d\theta}{(1+\theta)^2}$ is finite, and the same goes for $I_{j, k, \alpha, \beta}^2$.

Thus there exists $C(\alpha, \beta) > 0$ such that for all (j, k) : $I_{j, k, \alpha, \beta} \leq C(\alpha, \beta)$, which completes the proof of the lemma. \square

Acknowledgment

The author would like to thank Professor Dominique Picard for her helpful comments on the paper.

References

- F. Abramovich and B. W. Silverman. Wavelet decomposition approaches to statistical inverse problems. *Biometrika*, 85(1):115–129, 1998. ISSN 0006-3444.
- K. F. Andersen and R. T. John. Weighted inequalities for vector-valued maximal functions and singular integrals. *Studia Math.*, 69(1):19–31, 1980/81. ISSN 0039-3223.
- A. Antoniadis, J. Fan, and I. Gijbels. A wavelet method for unfolding sphere size distributions. *The Canadian Journal of Statistics*, 29:265–290, 2001.
- L. Cavalier and A. Tsybakov. Sharp adaptation for inverse problems with random noise. *Probab. Theory Related Fields*, 123(3):323–354, 2002. ISSN 0178-8051.
- L. Cavalier, G. K. Golubev, D. Picard, and A. B. Tsybakov. Oracle inequalities for inverse problems. *Ann. Statist.*, 30(3):843–874, 2002. ISSN 0090-5364.
- L. M. Cruz-Orive. Distribution-free estimation of sphere size distributions from slabs showing overprojections and truncations, with a review of previous methods. *J. Microscopy*, 131:265–290, 1983.

- D. L. Donoho. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmon. Anal.*, 2(2):101–126, 1995. ISSN 1063-5203.
- S. Efromovich and V. Koltchinskii. On inverse problems with unknown operators. *IEEE Trans. Inform. Theory*, 47(7):2876–2894, 2001. ISSN 0018-9448.
- J. Fan and J.Y. Koo. Wavelet deconvolution. *IEEE Trans. Inform. Theory*, 48(3):734–747, 2002. ISSN 0018-9448.
- C. Fefferman and E. M. Stein. Some maximal inequalities. *Amer. J. Math.*, 93:107–115, 1971. ISSN 0002-9327.
- A. Goldenshluger and S. V. Pereverzev. On adaptive inverse estimation of linear functionals in Hilbert scales. *Bernoulli*, 9(5):783–807, 2003. ISSN 1350-7265.
- G. K. Golubev and B. Y. Levit. Asymptotically efficient estimation in the Wicksell problem. *Ann. Statist.*, 26(6):2407–2419, 1998. ISSN 0090-5364.
- W. Härdle, G. Kerkyacharian, D. Picard, and A. Tsybakov. *Wavelets, Approximation and Statistical Applications*. Springer-Verlag, 1998.
- I. M. Johnstone and B. W. Silverman. Discretization effects in statistical inverse problems. *J. Complexity*, 7(1):1–34, 1991. ISSN 0885-064X.
- I. M. Johnstone, G. Kerkyacharian, D Picard, and M. Raimondo. Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society*, 66(3):1–27, 2004.
- J. Kalifa and S. Mallat. Thresholding estimators for linear inverse problems and deconvolutions. *Ann. Statist.*, 31(1):58–109, 2003. ISSN 0090-5364.
- G. Kerkyacharian, D. Picard, P. Petrushev, and T. Willer. Needvd: a second generation wavelet procedure for estimation in inverse problems. 2006. LPMA 2006.
- R. Lipster and A. N. Shiryaev. *Statistics of random processes*, volume I. 1st edition, 1977.

- P. Mathé and S. V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003. ISSN 0266-5611.
- F. J. Narcowich, P. Petrushev, and J.M. Ward. Decomposition of besov and triebel-lizorkin spaces on the sphere. *J. Funct. Anal.*, 2006.
- D. Nychka, G. Wahba, S. Goldfarb, and T. Pugh. Cross validated spline methods for the estimation of three-dimensional tumor size distributions from observations on two-dimensional cross sections. *J. Amer. Statist. Assoc.*, 79:832–846, 1984.
- M. Pensky and B. Vidakovic. Adaptive wavelet estimator for nonparametric density deconvolution. *Annals of Statistics*, 27:2033–2053, 1999.
- P. Petrushev and Y. Xu. Localized polynomial frames on the interval with Jacobi weights. *J. Fourier Anal. Appl.*, 11(5):557–575, 2005. ISSN 1069-5869.
- G. Szegő. *Orthogonal polynomials*. American Mathematical Society, Providence, R.I., 1975.
- A. Tsybakov. On the best rate of adaptive estimation in some inverse problems. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(9):835–840, 2000. ISSN 0764-4442.
- A. B. Tsybakov. *Introduction à l'estimation non-paramétrique*, volume 41 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2004. ISBN 3-540-40592-5.
- S. D. Wicksell. The corpuscle problem: a mathematical study of a biometric problem. *Biometrika*, 17:84–99, 1925.
- T. Willer. Deconvolution in white noise with a random blurring effect. 2005. LPMA 2005.
- T. Willer. Non parametric estimation and inverse problems. *Ph.D. thesis, Université Paris 7*, 2006.