数理解析研究所講究録
第 1620 巻 2009 年 119-146



# New generation wavelets associated with statistical problems

KERKYACHARIAN, Gérard and PICARD, Dominique

**Abstract**

In this paper, we give a strong motivation, based on new statistical problems mostly concerned with high frequency data, for the construction of second generation wavelets. These new wavelets basically differ from the classical ones in the fact that, instead of being constructed on the Fourier basis, they are associated with different orthonormal bases such as bases of polynomials. We give in the introduction three statistical problems where these new wavelets are clearly helpful. These examples are revisited in the core of the paper, where the use of the wavelets are enlightened. The construction of these new wavelets is given as well as their important concentration properties in spectral and space domains. Spaces of regularity associated with these new wavelets are studied, as well as minimax rates of convergence for nonparametric estimation over these spaces.

## 1 Introduction

The last decade has witnessed a growing interest for so-called high resolution data, i.e. observations collected on processes observed on a domain of fixed amplitude but sampled at higher and higher frequency (equivalently, at smaller and smaller scales). As a consequence, the asymptotic point of view (where we let the number of observations go to infinity) has a rather different status than usual large sample theory in mathematical statistics. Such a strong interest has been motivated by the rich mathematical structure that these issues involve, and at the same time by a huge variety of strong motivations arising in applications. Concerning the latter, we mention first the analysis of financial data ; in particular, we recall the problems linked with microstructure noise, corresponding in fact to observations of different behaviours at different scales of time. A second, extremely important example is provided by the cosmological and astrophysical literature. Particularly relevant here are the problems related to Cosmic Microwave Background (CMB) radiation analysis, an extraordinarily active domain which is currently at the frontier of theoretical and experimental physics and has already led to the Nobel Prize in Physics for G. Smoot and J. Mather in 2006. As a third important area from the applied sciences we mention medical imaging. Here, PET scan tomography and more generally image analysis issued from the Radon transform is also a domain in which if the technology is very active, a huge amount of mathematical and algorithmical questions are still opened for research and many new ones are emerging here driven by the technological progress.

During the last decade, wavelet analysis has emerged as a major tool in various disciplines, including most branches of pure and applied mathematics and statistics. However, the study of high frequency phenomena by means of wavelets has only been developed quite recently. There are important technical reasons to explain why the use of wavelets in a high frequency environment is considerably more challenging than in a traditional framework. Loosely speaking, high resolution analysis fundamentally requires the construction of wavelet bases which are specifically adapted to the problem at hand. For instance, in finance, cosmology or tomography, data are observed either on very particular domains, not necessarily well adapted to wavelets (i.e. the sphere in astrophysics) or they are observed after being blurred by a linear operator (Radon transform in the case of tomography, differential operators in finance). More frequently, a combination of these difficulties may arise in the same problem. In any of these examples, the usual wavelet tools (based on standard Fourier transforms) are not necessarily well adapted, so that wavelet techniques do not enjoy the optimality properties established for instance in [4].

We will precise the arguments above by giving three particular examples where quite obviously the need for the construction of a 'basis' with concentration properties at the same time in frequency and in space domain arises.

**Testing for Gaussiannity and/or isotropy** Let us first consider the case where we observe $(T_\xi)_{\xi\in\mathbb{S}}$ where $\mathbb{S}$ can be the torus or the unit sphere $\mathbb{S}^2$ of $\mathbb{R}^3$ (the unit sphere corresponds to the CMB case). An example of important question is: Is $\xi \mapsto T_\xi$ an 'isotropic' field ? where by isotropic, we mean stationary in the torus case and

$$\forall \rho \in \mathbb{O}(3), \quad \forall \xi, \eta \in \mathbb{S}^2 \;\; \mathbb{E}(T_\xi, T_\eta) = \mathbb{E}(T_{\rho(\xi)}, T_{\rho(\eta)}), \quad \mathbb{E}T_\xi = c$$

in the sphere case. If $T$ is a centered Gaussian field that is mean square continuous and isotropic, the covariance kernel

$$\mathbb{E}[T(x)T(y)] = K(x,y)$$

is only depending on the distance between $x$ and $y$ ($K(x,y) = K(d(x,y))$ and has the following spectral decomposition. The spaces $\mathcal{H}_k$ are the eigenspaces of the covariance operator

$$f(x) \in \mathbb{L}_2(\mathbb{S}^2) \mapsto Kf(x) = \int_{\mathbb{S}^2} K(d(x,y))f(y)dy$$

and

$$\forall f \in \mathcal{H}_k, \quad Kf = C_k f$$

where $\mathcal{H}_k$ is the space spanned by $\{e_k, e_{-k}\}$ in the torus case and the space of spherical harmonics of order $k$ (i.e. the restriction to $\mathbb{S}^2$ of homogeneous polynomials on $\mathbb{R}^3$, of degree $k$, which are harmonic: i.e. $\Delta P = 0$, where $\Delta$ is the Laplacian on $\mathbb{R}^3$-; this space is of dimension $2k+1$ and spanned by the so called spherical harmonics basis $Y_{k,m} \ldots, m = -k, \ldots, k$.

The sequence $C_k$ is generally called the angular power spectrum of the field $T_\xi$. This decomposition yields the so called Karhunen-Loève expansion of the process, in the torus case :

$$T_\xi = \sum_{k\in\mathbb{N}} \sum_{m=-1,+1} (\int T_u e_{km}(u)du) e_{km}(u)(\xi) = \sum_{k\in\mathbb{N}} \sum_{m=-1,+1} Z^k_m e_{km}(\xi).$$

in the sphere case :

$$T_\xi = \sum_{k\in\mathbb{N}} \sum_{m=-k,\dots,k} (\int_{\mathbb{S}^2} T_u Y_{km}(u) d\sigma(u)) Y_{km}(\xi) = \sum_{k\in\mathbb{N}} \sum_{m=-k,\dots,k} Z^k_m Y_{km}(\xi).$$

and the variables $Z^k_m = \int_{\mathbb{S}^2} T_u Y_{km}(u) d\sigma(u)$ are independent and with variance $C_k$. Usual tests for isotropy regularly take advantage of this independence of the $Z^k_m$'s for instance using $\chi^2$ or Kolmogorov-Smirnov tests.

However, main objections occur on these tests when the field is only partially observed, as it is the case for the CMB or even for financial data, when some part of the data might be missing. As the trigonometric basis as well as the spherical harmonics basis are very poorly concentrated in the space domain, any corruption on the data may yield a corruption of the $Z^k_m$'s as well. We see here the need for the construction of a basis which is at the same time concentrated in the spectral domain (here the $k$ indices) -to build 'atoms' which have chances to remain asymptotically independent-, but also in the time domain, - to avoid the spots where the data is masked-.

**Estimating the density of probability of a distribution on the sphere** We consider the problem of estimating the density f of an independent sample of points $X_1, \dots, X_n$ observed on the $d$-dimensional sphere $S^d$ of $\mathbb{R}^{d+1}$.

This study is especially motivated by many recent developments in the area of observational astrophysics. As an example, we refer to experiments measuring incoming directions of Ultra High Energy Cosmic Rays, such as the AUGER Observatory (http://www.auger.org). Here, efficient estimation of the density function of these directional data may yield crucial insights into the physical mechanisms generating the observations. More precisely, a uniform density would suggest the High Energy Cosmic Rays are generated by cosmological effects, such as the decay of massive particles generated during the Big Bang; on the other hand, if these Cosmic Rays are generated by astrophysical phenomena (such as acceleration into Active Galactic Nuclei), then we should observe a density function which is highly non-uniform and tightly correlated with the local distribution of nearby Galaxies. Massive amount of data in this area are expected to be available in the next few years.
There is an abundant literature about this type of problems. In particular, minimax $\mathbb{L}_2$ results have been obtained (see [12], [13]). These procedures are generally obtained using either kernel methods (but in this case the manifold structure of the sphere is not well taken into account), or using orthogonal series methods associated with spherical harmonics (and in this case the 'local performances of the estimator are quite poor, since spherical harmonics are spread all over the sphere).

Here, and precisely in view of the Auger application, we need to add the following requirements to the standard properties: we aim at a procedure which is minimax from $\mathbb{L}_2$ point of view but also performs satisfactorily from a local point of view (in infinity norm, for instance). In addition, we require this procedure to be simple to implement, as well as adaptive to inhomogeneous smoothness. This type of requirements is generally well handled using thresholding estimates associated to wavelets. The problem requires a special construction adapted to the sphere, since usual tensorized wavelets will never reflect the manifold structure of the sphere and will necessarily create unwanted artifacts. We will need in this case a basis mimicking the good performances of the wavelet basis but adapted to the sphere case: the fundamental properties of wavelets are their concentration in the Fourier domain as well as in the space domain. Here, obviously the 'space' domain is the sphere itself whereas the spectral domain is now obtained by replacing the 'Fourier' basis by the basis of Spherical Harmonics which plays an analogous role on the sphere.

Again, the problem of choosing appropriated spaces of regularity on the sphere in a serious question, and it is important to consider the spaces which may be the closest to our natural intuition: those which generalize to the sphere case the approximation properties shared by standard Besov and Sobolev spaces.

**Inverse problems** This problem deals with recovering a function $f$, when we receive a blurred (by a linear operator) and noisy version : $Y_\varepsilon = Kf + \varepsilon\dot{W}$. There is an abundant litterature on these problems and important examples such as the deconvolution problem (on a interval of $\mathbb{R}$ or on the sphere), the Wicksell problem, or the Radon problem. The direct problem ($K$ is the identity) isolates the denoising operation. It can't be solved unless accepting to estimate a smoothed version of $f$: for instance, if $f$ has an expansion on a basis, this smoothing might correspond to stopping the expansion at some stage $m$. Then a crucial problem lies in finding an equilibrium for $m$ considering the fact that for $m$ large, the difference between $f$ and its smoothed version is small, whereas the random effect introduces an error which is increasing with $m$. In the true inverse problem, in addition to denoising we have to 'inverse the operator' $K$, which operation not only creates the usual difficulties, but also introduces the necessity to control the additional instability due to the inversion of the random noise. Our purpose here is to emphasize the fact that in such a problem, there generally exists a basis which is fully adapted to the problem, where for instance the inversion remains very stable : this is the Singular Value Decomposition basis -i.e. in fact two orthonormal bases $(e_i)$ and $(g_i)$ and a set of coefficient $(b_i)$ such that $Ke_i = b_i g_i$, $K^* g_i = b_i e_i$ for all $i$. On the other hand, the SVD basis might not be appropriate for the accurate description of the solution with a small number of parameters. Also in many practical situations, the signal provides inhomogeneous regularity, and its local features are especially interesting to recover. In such cases, other bases (in particular localised bases such as wavelet bases) may be much more appropriate to give a good representation of the object at hand. For instance, in the deconvolution problem (on a compact subset of $\mathbb{R}$), the SVD basis is the trigonometric basis. In the case of deconvolution on the sphere, the problem is more tedious, but one can prove (see see [10] and [11]) that the SVD basis also involves the spherical harmonic basis. In the wicksell case, following [6], we have the following SVD :

$$e_k(x) = 4(k+1)^{1/2} x^2 P_k^{0,1}(2x^2 - 1)$$
$$g_k(y) = U_{2k+1}(y)$$

$P_k^{0,1}$ is the Jacobi polynomial of type $(0,1)$ with degree $k$ $U_k$ is the second type Chebishev polynomial with degree $k$. In the Radon problem, the following bases form the SVD bases:

$$e_{k,l,i}(x) = (2k+d)^{1/2} P_j^{(0,\,l+d/2-1)}(2|x|^2 - 1) Y_{l,i}(x), \quad 0 \le l \le k, \; k-l = 2j, \; 1 \le i \le N_{d-1}(l),$$

$$g_{k,l,i}(\theta, s) = [h_k^{(d/2)}]^{-1/2} (1-s^2)^{(d-1)/2} C_k^{d/2}(s) Y_{l,i}(\theta), \quad k \ge 0, \; l \ge 0, \; 1 \le i \le N_{d-1}(l),$$

where $P_j^{(0,\,l+d/2-1)}$ and $C_k^{d/2}$ are respectively Jacobi and Gegenbauer polynomials.

All the examples of SVD bases listed above are poorly concentrated. To provide a procedure with basically the same requirements as in the density estimation case (simplicity and stability of the algorithm, minimax optimality with respect to various $\mathbb{L}_p$ norms, adaptation to signals which might present inhomogeneous smoothness...) we again need the construction of a 'basis' adapted to the previous SVD bases -i.e. concentrated in their spectral domain- as well as concentrated in the space domain.

We will also need to define spaces of regularity and see how these spaces can be expressed in terms of the 'wavelet coefficients.

Frames were introduced in the 1950's by Duffin and Schaeffer [5] as a means for studying nonharmonic Fourier series. These are redundant systems which behave like bases and allow for a lot of flexibility. Tight frame which are very close to orthonormal bases are particularly useful in signal and image processing.

In the sequel we discuss a general scheme for construction of frames due to Petrushev and his co-authors [17], [18], [19]. As will be shown this construction has the advantage of producing easily computable frame elements which are extremely well localized in all cases of interest. Following [17], [18], [19], we will term them "needlets".

## 2 A general framework

The following framework can be collected from several papers of Kyriazis, Narcowich, Petrushev, Ward, and Xu. See [16], [17], [18], [14], [19]

In a lot of mathematical situations we have $\mathcal{Y}$, a compact metric space, $\mu$ a finite borelian measure and the following decomposition :

$$\mathbb{L}_2(\mathcal{Y}, \mu) = \bigoplus_{k=0}^{\infty} \mathcal{H}_k, \quad \text{where } \mathcal{H}_0 = \{\lambda 1, \quad \lambda \in \mathbb{C}\}$$

and each $\mathcal{H}_k$ is a finite dimension eigenspace associated to the spectral decomposition of a natural operator on $(\mathcal{Y}, \mu)$, -say a Laplacian $\Delta$)-. Usually $f \in \mathcal{H}_k \implies \bar{f} \in \mathcal{H}_k$. Let

$L_k$ denote the orthogonal projection on $\mathcal{H}_k$:

$$\forall f \in \mathbb{L}_2(\mathcal{Y}, \mu), \quad L_k(f)(x) = \int_{\mathcal{Y}} f(y) L_k(x,y) d\mu(y)$$

with

$$L_k(x,y) = \sum_{i=1}^{l_k} e_i^k(x) \bar{e}_i^k(y)$$

if $l_k$ is the dimension of $\mathcal{H}_k$ and $(e_i^k)_{i=1,\dots l_k}$ any orthogonal basis of $\mathcal{H}_k$. Obviously,

$$\int L_k(x,y) L_m(y,z) d\mu(y) = \delta_{k,m} L_k(x,z) \tag{1}$$

*Remark* : In the sequel, if there is no ambiguity, $\mu(A)$, for a borel set $A$ could be replaced by $|A|$. As well we identify, with a slight abuse of notation and when there is no ambiguity, operators with the associated kernels (e.g. $L_k$ with $L_k(x,y)$). ⋆

## 2.1 Examples

The following examples will be used throughout the paper.

**Torus case** Let $\mathcal{Y} = \mathbb{S}^1$ be the torus equipped with the Lebesgue measure. The spectral decomposition of the Laplacian operator gives rise to the classical Fourier basis, with

$$\mathcal{H}_k = span\{e^{ikx}, e^{-ikx}\} = span\{\sin kx, \cos kx\}$$

Hence, for all $k > 1$, $dim(\mathcal{H}_k) = 2$, and

$$L_k(x,y) = 2\cos k(x-y)$$

**Jacobi case** Let us now take $\mathcal{Y} = [-1,1]$ equipped with the measure $\omega(x)dx$ with $\omega(x) = (1-x)^\alpha (1+x)^\beta$, $\alpha > -\frac{1}{2}$, $\beta > -\frac{1}{2}$.
If $\sigma(x) = (1-x)^2$, then $\tau := \frac{(\sigma\omega)'}{\omega}$ is a polynom of degree 1, and

$$D(f) = \frac{(\sigma\omega f')'}{\omega} = \sigma f'' + \tau f'$$

is a self-adjoint (in $\mathbb{L}_2(\omega(x)dx)$) second order differential operator (Here and in the sequel, $u'$ denote the derivative of $u$).
Using Gram Schmidt orthonormalization (again, in $\mathbb{L}_2(\omega(x)dx)$) of $x^k$ we get a family of orthonormal polynomials $P_k$, called Jacobi polynomials, which coincides with the spectral decomposition of $D$.

$$DP_k = [k(k-1)\frac{\sigma"}{2} + k\tau'] P_k$$

If we put for all $k \in \mathbb{N}$, $\mathcal{H}_k = span\{P_k\}$ we have $dim(\mathcal{H}_k) = 1$, $L_k(x,y) = P_k(x)P_k(y)$.

**Sphere case** Let now $\mathcal{Y} = \mathbb{S}^2 \subset \mathbb{R}^3$. Note that this construction has obvious an generalisation for $\mathbb{S}^k \subset \mathbb{R}^{k+1}$. The geodesic distance on $\mathbb{S}^2$ is given by

$$d(x,y) = \cos^{-1}(\langle x,y\rangle), \quad \langle x,y\rangle = \sum_{i=1}^{3} x_i y_i.$$

There is a natural measure $\sigma$ on $\mathbb{S}^2$ which is *rotation invariant*: (i.e. $\forall \rho \in \mathbb{O}(3)$, $\int_{\mathbb{S}^2} f(\rho(u))d\sigma(u) = \int_{\mathbb{S}^2} f(u)d\sigma(u)$.) Furthermore if $F$ is defined on $\mathbb{R}^3$,

$$\int_{\mathbb{R}^3} F(x)dx = \int_0^\infty r^2 \int_{\mathbb{S}^2} F(ru)d\sigma(u)dr.$$

There is a natural Laplacian on $\mathbb{S}^2$, $\Delta_{\mathbb{S}^2}$, which is a self-adjoint operator with the following spectral decomposition:
If $\mathcal{H}_k$ is the restriction to $\mathbb{S}^2$ of polynomials of degree k which are *homogeneous* ($P = \sum_{|\alpha|=k} a_\alpha x^\alpha$, $\alpha = (\alpha_1,\alpha_2,\alpha_3)$, $|\alpha| = \sum \alpha_i$, $\alpha_i \in \mathbb{N}$) and *harmonic* ($\Delta P = \sum_{i=1}^3 \frac{\partial^2 P}{\partial x_i^2} = 0$), we have :

$$\forall P \in \mathcal{H}_k, \quad \Delta_{\mathbb{S}^2} P = -k(k+1)P$$

$\mathcal{H}_k$ is called the space of spherical harmonics of order $k$, $dim(\mathcal{H}_k) = 2k+1$, and if $(Y_{ki})_{-k\le i\le k}$ is an orthonormal basis of $\mathcal{H}_k$, the projector writes

$$L_k(x,y) = \sum_{-k\le i\le k} Y_{ki}(x)\overline{Y_{ki}(y)}.$$

Furthermore, one can prove that

$$L_k(x,y) = L_k(\langle x,y\rangle)$$

where $L_k(u)$ is the Legendre polynomial of degree $k$ (a special case of Jacobi polynomial- with a different normalisation- corresponding to $\alpha = \beta = 0$) and:

$$\int_{-1}^{1} L_k(u)L_m(u)du = \delta_{k,m}\frac{2k+1}{8\pi^2}$$

**Ball case, Radon transform** Let $\mathcal{Y} = B^d$ be the unit ball of $\mathbb{R}^d$ equipped with the Lebesgue measure.
For $f \in \mathbb{L}^2(B^d, dx)$, $\theta \in \mathbb{S}^{d-1}$, (the unit sphere of $\mathbb{R}^d$), $t \in [-1,1]$, we define the Radon transform of $f$:

$$Rf(\theta,t) = \int_{\langle\theta,x\rangle = t} f(x)dx$$

($dx$ in the formula above denotes the Lebesgue measure on the hyperplan $\langle\theta,x\rangle = t$).
If $d\mu(\theta,t)$ is the measure $d\sigma(\theta)\frac{dt}{(\sqrt{1-t^2})^{d-1}}$ on $\mathbb{S}^{d-1}\times[-1,1]$, the Radon transform $R$ is

a continuous mapping from $\mathbb{L}^2(B^d, dx)$ to $\mathbb{L}^2(\mathbb{S}^{d-1} \times [-1,1], d\mu(\theta,t))$, with the following adjoint: if $g(\theta,t) \in \mathbb{L}^2(\mathbb{S}^{d-1} \times [-1,1], d\mu(\theta,t))$

$$R^*(g)(x) = \int_{\mathbb{S}^{d-1}} g(\theta, \langle x, \theta\rangle)\Big(\frac{1}{\sqrt{1-|\langle x,\theta\rangle|^2}}\Big)^{d-1} d\sigma(\theta)$$

Let $\Pi_k(B^d)$ be the space of polynomials of degree $\leq k$ on the unit ball of $\mathbb{R}^d$, the following decomposition is true:

$$\Pi_k(B^d) = \mathcal{V}_k(B^d) \bigoplus \Pi_{k-1}(B^d)$$
$$\mathbb{L}^2(B^d) = \bigoplus_{k=0}^{\infty} \mathcal{V}_k(B^d)$$

and the spaces $\mathcal{V}_k(B^d)$'s are the eigenspaces of $R^*R$. This provides a Singular Value Decomposition of $R$, with corresponding eigenvalues,

$$\mu_k^2 = \frac{\pi^{d-1}2^d}{(k+1)....(k+d)} \sim k^{-(d-1)}$$

The kernel projector on $\mathcal{V}_k$ is given by

$$L_k(x,y) = \frac{2k+d}{|\mathbb{S}^{d-1}|^2}\int_{\mathbb{S}^{d-1}} C_k^{\nu+1}(\langle x,\xi\rangle)C_k^{\nu+1}(\langle y,\xi\rangle)d\sigma(\xi),$$

where $\nu = \frac{d}{2} - 1$, and $C_k^{\nu+1}$ is the Gegenbauer polynomial.

## 2.2 General construction of needlets

In the examples above as well as in many "physical" contexts, the eigenfunctions of a 'natural' operator (Laplacian, $D$, Radon,...) are of sinusoidal type and very badly concentrated in the space $\mathcal{Y}$

The consequence is that each time we compute the coefficient of a function $f$, $\int_{\mathcal{Y}} f(x)P_n(x)d\mu$, any local perturbation on $f$ (or lack of information) will severely corrupt the coefficient. In the sequel we describe the construction of a frame which will provide a "spectral" description of the function $f$ (in the sense that the atoms of the frame will be reasonably spectrally concentrated), without the drawback of the bases encountered above (since the atoms will be also well concentrated in the space domain). The following construction is based on three fundamental steps : Littlewood-Paley decomposition, splitting and discretization, which are summarized in the three following subsections.

### 2.2.1 Littlewood-Paley decomposition

Let $0 \leq a \leq 1$, be a $C_\infty$ non negative function defined on $[0,\infty)$. We impose $a$ to be identically 1 on $[0,1/2]$ and compactly supported on $[0,1]$. Let us now define the sequence of linear operators $A_j,\ j \geq 0$ and $B_j,\ j \geq 0$, with $A_0(f) = \frac{1}{|\mathcal{Y}|}\int_{\mathcal{Y}} f(x)d\mu(x)$ and for $j \geq 1$

$$\begin{aligned}
A_j f(x) &= \int_{\mathcal{Y}} A_j(x,y)f(y)d\mu(y), \\
A_j(x,y) &:= \sum_k a(\frac{k}{2^j})L_k(x,y) = \sum_{k<2^j} a(\frac{k}{2^j})L_k(x,y) \\
B_j &:= A_{j+1} - A_j = \sum_k b(\frac{k}{2^j})L_k \\
b(x) &:= a(x/2) - a(x)
\end{aligned}$$

Obviously,

$$\langle A_j f, f\rangle = \sum_k a(\frac{k}{2^j})\langle L_k f, f\rangle \leq \|f\|^2 \tag{2}$$

$$\lim_{j\to\infty} \|A_j(f) - f\|_2 = \lim_{j\to\infty} \|(A_0 + \sum_{m=0}^{j-1} B_m)(f) - f\|_2 = 0 \tag{3}$$

### 2.2.2 The splitting procedure.

Let us define

$$D_j(x,y) = \sum_{2^{j-1}<k<2^{j+1}} \sqrt{b}(\frac{k}{2^j})L_k(x,y).$$

Due to (1),

$$\int D_j(x,u)D_j(u,y)d\mu(u) = B_j(x,y)$$

And in the same way,

$$\begin{aligned}
C_j(x,y) &= \sum_{0\leq k<2^j} \sqrt{a}(\frac{k}{2^j})L_k(x,y) \\
A_j(x,y) &= \int C_j(x,u)C_j(u,y)d\mu(u)
\end{aligned}$$

### 2.2.3 Gauss quadrature formula

Let us now suppose, as important ingredient of this construction, that there is a *quadrature formula* for $\bigoplus_{l\leq 2^{2+j}} \mathcal{H}_l,\ j \in \mathbb{N}$. This means that there exists a finite set $\mathcal{X}_j$ of $\mathcal{Y}$, and for all $\xi \in \mathcal{X}_j$, there is an associated coefficient $\lambda_{j,\xi} > 0$, such that for all $f \in \bigoplus_{l\leq 2^{2+j}} \mathcal{H}_l$, we have the following interpolation formula:

$$\int_{\mathcal{Y}} f(y)d\mu(y) = \sum_{\xi\in\mathcal{X}_j} \lambda_{j,\xi} f(\xi).$$

Let us suppose, in addition that $f \in \bigoplus_{l\leq k}\mathcal{H}_l$ implies $\overline{f} \in \bigoplus_{l\leq k}\mathcal{H}_l$, as well as $f \in \bigoplus_{l\leq k}\mathcal{H}_l$, $g \in \bigoplus_{l\leq m}\mathcal{H}_l$ imply $fg \in \bigoplus_{l\leq k+m}\mathcal{H}_l$.

These assumptions ensure that $u \mapsto C_j(x,u)C_j(u,y)$ is in $\bigoplus_{l\leq 2^{2+j}}\mathcal{H}_l$ as well as $u \mapsto D_j(x,u)D_j(u,y)$. As a consequence we write,

$$A_j(x,y) = \int C_j(x,u)C_j(u,y)d\mu(u) = \sum_{\xi\in\mathcal{X}_j} \lambda_{j,\xi}C_j(x,\xi)C_j(\xi,y) \tag{4}$$

$$B_j(x,y) = \int D_j(x,u)D_j(u,y)d\mu(u) = \sum_{\xi\in\mathcal{X}_j} \lambda_{j,\xi}D_j(x,\xi)D_j(\xi,y) \tag{5}$$

which will directly induce the definition of the needlets.

**Example of Gauss quadrature formula** Let us first observe that the assumptions above are easy to handle, at least in the examples presented above

1. $\mathcal{Y} = \mathbb{S}^1$ It is easy to prove that for :

$$f = \sum_{|l|<N} a_l e^{ilx}, \quad \frac{1}{2\pi}\int_0^{2\pi} f(x)dx = \frac{1}{N}\sum_{0\leq k<N} f(\frac{2k\pi}{N})$$

   Indeed:

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)dx = \sum_{|l|<N} a_l\frac{1}{2\pi}\int_0^{2\pi} f(x)dx = a_0, \text{ and}, \forall\, |l| < N, \ \frac{1}{N}\sum_{0\leq k<N} e^{il\frac{2k\pi}{N}} = \delta_{0,l}.$$

2. Orthonormal polynomials. Let us recall that $P_N$ denotes the Jacobi polynomial of degree $N$. Let us consider $\{\xi_1, \xi_2, \dots, \xi_N\}$ the set of its roots. It is well known that this set is of cardinality $N$. Moreover, tehre exist $\lambda_1 > 0, \dots, \lambda_N > 0$ such that,

$$\forall P = \sum_{l=0}^{2N-1} a_l x^l, \quad \int_{-1}^{1} P(x)\omega(x)dx = \sum_{k=1}^{N}\lambda_k P(\xi_k).$$

   Indeed using the euclidean division we write,

$$P = QP_N + R, \quad d^o(Q) < N, d^o(R) < N, \text{ so clearly } \ P(\xi_k) = R(\xi_k)$$

   and by orthogonality

$$\int_{-1}^{1} P(x)\omega(x)dx = \int_{-1}^{1} Q(x)P_N(x)\omega(x)dx + \int_{-1}^{1} R(x)\omega(x)dx = \int_{-1}^{1} R(x)\omega(x)dx = \sum_{k=1}^{N}\lambda_k R(\xi_k)$$

where

$$\lambda_k = \int_{-1}^{1} \frac{\prod_{i\neq k}(x-\xi_i)}{\prod_{i\neq k}(\xi_k-\xi_i)}\omega(x)dx = \int_{-1}^{1} [\frac{\prod_{i\neq k}(x-\xi_i)}{\prod_{i\neq k}(\xi_k-\xi_i)}]^2\omega(x)dx > 0$$

by Lagrange interpolation formula (and $[\frac{\prod_{i\neq k}(x-\xi_i)}{\prod_{i\neq k}(\xi_k-\xi_i)}]^2 = P_N(x)Q(x) + \frac{\prod_{i\neq k}(x-\xi_i)}{\prod_{i\neq k}(\xi_k-\xi_i)}$).

3. For the sphere $\mathbb{S}^2$, one can prove (see [16] and [17]) that if $\mathcal{X}_N$ is an $\epsilon-$net with $\epsilon \sim \frac{1}{N}$, (hence $Card(\mathcal{X}_N) \sim N^2$), there exists a set of positive coefficients $\lambda_\xi$ -in fact $\lambda_\xi \sim \frac{1}{N^2}$-, such that

$$\forall\ f(x) = \sum_{|\alpha|\leq N} a_\alpha x^\alpha, \quad \int_{\mathbb{S}^2} f(x)d\sigma(x) = \sum_{\mathcal{X}_N} f(\xi)\lambda_\xi,$$

#### 2.2.4 Needlets frame.

Rewriting (4) and (5) in the following way,

$$A_j(x,y) = \sum_{\xi\in\mathcal{X}_j} \sqrt{\lambda_{j,\xi}}C_j(x,\xi)\sqrt{\lambda_{j,\xi}}\quad \overline{C_j(y,\xi)}$$

$$B_j(x,y) = \sum_{\xi\in\mathcal{X}_j} \sqrt{\lambda_{j,\xi}}D_j(x,\xi)\sqrt{\lambda_{j,\xi}}\quad \overline{D_j(y,\xi)},$$

we get

$$A_j(f)(x) = \sum_{\xi\in\mathcal{X}_j} \sqrt{\lambda_{j,\xi}}C_j(x,\xi)\int f(y)\overline{\sqrt{\lambda_{j,\xi}}C_j(y,\xi)}d\mu(y)$$

$$B_j(f)(x) = \sum_{\xi\in\mathcal{X}_j} \sqrt{\lambda_{j,\xi}}D_j(x,\xi)\int f(y)\overline{\sqrt{\lambda_{j,\xi}}D_j(y,\xi)}d\mu(y)$$

with as a consequence, the following definition,

**Definition 1.** *We define the 'mother' and 'father' needlet 'basis' as follows: for* $\xi \in \mathcal{X}_j$,

$$\phi_{j,\xi}(x) = \sqrt{\lambda_{j,\xi}}C_j(x,\xi) = \sqrt{\lambda_{j,\xi}} \sum_{2^{j-1}<k<2^{j+1}} \sqrt{a(\frac{k}{2^j})}L_k(x,\xi)$$

$$\psi_{j\xi}(x) = \sqrt{\lambda_{j,\xi}}D_j(x,\xi) = \sqrt{\lambda_{j,\xi}} \sum_{2^{j-1}<k<2^{j+1}} \sqrt{b(\frac{k}{2^j})}L_k(x,\xi).$$

We have $f = L_0(f) + \sum_j B_j f = \langle f,\phi_0\rangle_{\mathbf{L}_2(\mu)}\phi_0 + \sum_{j\in\mathbf{N}}\sum_{\xi\in\mathcal{X}_j}\langle f,\psi_{j\xi}\rangle_{\mathbf{L}_2(\mu)}\psi_{j\xi}$ (6)

and $A_j(f) = \sum_{\xi\in\mathcal{X}_j}\langle f,\phi_{j,\xi}\rangle_{\mathbf{L}_2(\mu)}\phi_{j,\xi}$ (7)

### 2.2.5 Frame property of the needlet basis.

Using (6), we get

$$\|f\|^2_{\mathbf{L}_2(\mu)} = \sum_{j\in\mathbb{N}} \sum_{\xi\in\mathcal{X}_j} |\langle f, \psi_{j\xi}\rangle_{\mathbf{L}_2(\mu)}|^2 \tag{8}$$

But

$$\sum_{\xi\in\mathcal{X}_j} \sqrt{\lambda_{j,\xi}}\psi_{j\xi}(x) = \sum_{\xi\in\mathcal{X}_j} \lambda_{j,\xi} B_j(x,\xi) = \int_{\mathcal{X}} B_j(x,u)d\mu(u) \equiv 0.$$

So the family $(\psi_{j\xi})_{j\in\mathbb{N},\xi\in\mathcal{X}_j}$ is a tight frame (with frame constant 1) but is not linearly independent, so *not* an orthonormal basis. By construction, the spectral localisation of $\psi_{j\xi}$ is between $2^{j-1}$ and $2^{j+1}$

**Proposition 1.**

$$\forall\, j,\ \xi \in \chi_j,\quad \|\psi_{j\xi}\|_2^2 \le 1,\quad \|\phi_{j,\xi}\|_2^2 \le 1$$

**Proof :** By (8)

$$\|\psi_{j\xi}\|^2_{\mathbf{L}_2(\mu)} = \sum_{j'\in\mathbb{N}} \sum_{\eta\in\mathcal{X}_{j'}} |\langle \psi_{j',\xi}, \psi_{j',\eta}\rangle_{\mathbf{L}_2(\mu)}|^2 \ge \|\psi_{j',\xi}\|^4_{\mathbf{L}_2(\mu)}$$

By (2)

$$\|\phi_{j,\xi}\|^4_{\mathbf{L}_2(\mu)} \le \sum_{\eta\in\mathcal{X}_j} |\langle \phi_{j,\xi}, \phi_{j,\eta}\rangle_{\mathbf{L}_2(\mu)}|^2 = \langle A_j(\phi_{j,\xi}), \phi_{j,\xi}\rangle \le \|\phi_{j,\xi}\|_2^2.$$

### 2.2.6 Spatial localisation of the needlet basis.

This construction has in general the beautiful property of producing a localised frame. The localisation results can be found in [17],[16], [18]. For instance, in the sphere case the following almost exponential concentration property is proved:

$$\forall M \in \mathbb{N},\ \exists\, C_M \text{ such that } |\psi_{j\xi}(x)| \le C_M \frac{2^j}{(1+2^j d(x,\xi))^M}.$$

The $\xi$'s corresponding to a level $j$, will also be denoted by $\xi_{jk}$, $\psi_{j,\eta}$ by $\psi_{j,k}$ by extension.

Let us verify this property in the simple case of the torus $\mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$, where all the calculations are very easy.
We have :

$$D_j(x,y) = D_j(x-y) = \sum_{2^{j-1}<k<2^{j+1}} \sqrt{b}(\frac{k}{2^j}) 2\cos k(x-y)$$

Indeed, using the Poisson summation formula, if $\mathcal{F}(h)$ denotes the Fourier transform of the function $h$

$$D_j(x) = \sum_{k\in\mathbb{Z}} \sqrt{b}(\frac{k}{2^j})e^{ikx} = \sum_{m\in\mathbb{Z}} 2^j\mathcal{F}(\sqrt{b})(2^j(x-2\pi m))$$

On the other part :

$$\forall M \in \mathbb{N},\ \exists\, C_M,\quad |\mathcal{F}(\sqrt{b}(x))| \leq \frac{C_M}{(1+|x|)^M}$$

$$|D_j(x)| \leq C_M \sum_{m\in\mathbb{Z}} 2^j \frac{1}{(1+|2^j(x-2\pi m)|)^M} \leq 5C_M \frac{2^j}{(1+|2^j x|)^M}$$

As seen above (see subsection 2.2.3) the cubature points are forming the set $\{\xi_l = \frac{2\pi l}{2^{j+2}},\ l = 0, 1 \ldots 2^{j+2}-1\}$ and $\lambda_{\xi_l} = 2^{-j-2}$,

$$|\psi_{j,\xi_l}(x)| = |\sqrt{\lambda_\xi} D_j(x-\xi_l)| = 2^{-j/2-1}|D_j(x-\frac{2\pi l}{2^{j+2}})|$$

$$\leq \frac{5}{2} C_M \frac{2^{j/2}}{(1+|2^j(x-\frac{2\pi l}{2^{j+2}}|)^M}$$

These localisation properties are very important in the applications. For instance, in the case where $T(x)$ is a gaussian isotropic field defined on the sphere, it is proved in [1] (see also[3], [1], [15]) under mild regularity conditions on the angular power spectrum of $T$, that the random spherical needlet coefficients defined as

$$\beta_{j,k} := \int_{\mathbb{S}^2} T(x)\psi_{j,\xi_{jk}}(x)\, d\sigma(x) = \sqrt{\lambda_{j,k}} \sum_l b(\tfrac{l}{2^j}) T_l(\xi_{j,k}) \ .$$

where

$$T_l(x) := \sum_{m=-l}^{l} \int T(u) Y_{lm}(u)\, d\sigma(u) Y_{lm}(x) \tag{9}$$

are such that

$$\left|Corr(\beta_{j,k}, \beta_{j,k'})\right| \leq \frac{C_M}{(1+2^j d(\xi_{j,k}, \xi_{j,k'}))^M} \tag{10}$$

where, as hinted above, $d(\xi_{j,k}, \xi_{j,k'}) = \arccos(\langle \xi_{j,k}, \xi_{j,k'}\rangle)$, and $Corr$ denotes the correlation. This almost exponential decreasing rate is a fondamental tool to prove central limit theorems, used in [1] to provide tests for Gaussianity or isotropy.

## 2.3 Key inequalities

In all the examples detailed above, the following inequalities are true (and proved in [17],[16], [18]) and, as will be obvious in the sequel, quite important in particular for the approximation and statistical properties. In the following lines , $g_{j,\xi}$ will stand either for $\phi_{j,\xi}$ or $\psi_{j\xi}$. There exist $c \leq C$ such that

$$\forall j \in \mathbb{N}, \forall \xi \in \chi_j, \qquad 0 < c \leq \|g_{j,\xi}\|_2^2 \leq 1 \tag{11}$$

$$\forall j \in \mathbb{N},\ \ \forall x \in \mathcal{Y},\quad \sum_{\xi\in\chi_j} \|g_{j,\xi}\|_1 |g_{j,\xi}(x)| \leq C < \infty \tag{12}$$

Let us give, as an example, a proof of the property (11), in the sphere case:
We have seen previously that we always have $\|g_{j,\xi}\|_2^2 \leq 1$.
Using Parseval equality,

$$\begin{aligned}
\|\phi_{j,\xi}\|_2^2 &= \lambda_\xi \sum_{0\leq k<2^j} a(\frac{k}{2^j})L_k(\xi,\xi) \\
&= \lambda_\xi \sum_{0\leq k\leq 2^{j-1}} L_k(\xi,\xi) + \lambda_\xi \sum_{2^{j-1}<k<2^j} a(\frac{k}{2^j})L_k(\xi,\xi) \\
\|\psi_{j\xi}\|_2^2 &= \lambda_\xi \sum_{2^{j-1}<k<2^{j+1}} b(\frac{k}{2^j})L_k(\xi,\xi).
\end{aligned}$$

But $L_k(\xi,\xi) = L_k(\langle\xi,\xi\rangle) = L_k(1)$ and

$$\begin{aligned}
L_k(1)\sigma(\mathbb{S}^2) &= \int_{\mathbb{S}^2} L_k(1)d\sigma(u) = \int_{\mathbb{S}^2} L_k(\langle u,u\rangle)d\sigma(u) \\
&= \int_{\mathbb{S}^2} \sum_{i=1}^{2k+1} |P_i^k(u)|^2 d\sigma(u) = 2k+1.
\end{aligned}$$

On the other hand, as $\lambda_\xi \sim 2^{-2j}$ for $\xi \in \chi_j$, we have,

$$\|\phi_{j,\xi}\|_2^2 \sim 2^{-2j} \sum_{0\leq k<2^j} a(\frac{k}{2^j})(2k+1) \geq 2^{-j} \sum_{0\leq k<2^j} a(\frac{k}{2^j})\frac{2k}{2^j} \sim \int_0^1 2xa(x)dx > 0.$$

With an analogous argument for $g = \psi_{j\xi}$.

# 3 Besov spaces

The problem of choosing appropriated spaces of regularity on the sphere or associated with one of the other examples presented above in a serious question, and it is important to consider the spaces which may be the closest to our natural intuition: those which generalize usual approximation properties. On the other hand, we are interested in spaces which can be characterised by their needlet coefficients. This will be our concern in this section. We will see that these 'new Besov' spaces share some properties with the standard ones. On the other hand some parts are slightly different as for instance the embeddings results (see the end of the section). For a more complete description see [17],[16], [18]. We begin with fundamental behaviors of the $\mathbb{L}_p$ norms.

## 3.1 $\mathbb{L}_p$ robustness of the $\psi_{j\xi}$ frame

**Theorem 1.** *Let us suppose that the frame $\{\psi_{j\xi}\}$ constructed above, verifies (11) and (12), where as above, $g_{j,\xi}$ stands either for $\phi_{j,\xi}$ or $\psi_{j\xi}$, then for all $1 \leq p \leq \infty$, (with the* *usual modification for $p = \infty$)*

$$\forall\, j \in \mathbb{N},\ f(\sum_{\xi\in\chi_j} |\langle f, g_{j,\xi}\rangle|^p \|g_{j,\xi}\|_p^p)^{1/p} \le C\|f\|_p \tag{13}$$

$$\forall\, j \in \mathbb{N},\ \lambda_\xi,\ \xi \in \chi_j,\ \|\sum_{\xi\in\chi_j} \lambda_\xi g_{j,\xi}(x)\|_p \le (\frac{C}{c})^2(\sum_{\xi\in\chi_j} \|\lambda_\xi g_{j,\xi}\|_p^p)^{1/p} \tag{14}$$

**Proof :**
From (11) , (12), and interpolation, we have

$$c \le \|g_{j,\xi}\|_2^2 \le \|g_{j,\xi}\|_1 \|g_{j,\xi}\|_\infty \le C. \tag{15}$$

(15) will be used all along this proof when dividing by $\|g_{j,\xi}\|_\infty$.

**Proof of (13)** Obviously,

$$\sup_{\xi\in\chi_j} |\langle f, g_{j,\xi}\rangle| \|g_{j,\xi}\|_\infty \le \sup_{\xi\in\chi_j} \int |f(x)||\phi_{j,\xi}(x)|d\mu(x)\|g_{j,\xi}\|_\infty \le \|f\|_\infty \sup_{\xi\in\chi_j} \|g_{j,\xi}\|_1\|g_{j,\xi}\|_\infty \le C\|f\|_\infty$$

and

$$\begin{aligned}\sum_{\xi\in\chi_j} |\langle f, g_{j,\xi}\rangle| \|g_{j,\xi}\|_1 &\le \sum_{\xi\in\chi_j} \int |f|(x)|g_{j,\xi}(x)|d\mu(x)\|g_{j,\xi}\|_1 \\ &= \int |f|(x) \sum_{\xi\in\chi_j} |g_{j,\xi}(x)| \|g_{j,\xi}\|_1 d\mu(x) \le C\|f\|_1\end{aligned}$$

Let now $1 < p < \infty$ and $1/p + 1/q = 1$, using Hölder inequality, and interpolation :

$$\begin{aligned}\sum_{\xi\in\chi_j} |\langle f, g_{j,\xi}\rangle|^p \|g_{j,\xi}\|_p^p &\le \sum_{\xi\in\chi_j} (\int |f(x)||g_{j,\xi}(x)|d\mu(x))^p \|g_{j,\xi}\|_p^p \\ &\le \sum_{\xi\in\chi_j} (\int |f(x)|^p |g_{j,\xi}(x)|d\mu(x))(\int |g_{j,\xi}(x)|d\mu(x))^{p/q} \|g_{j,\xi}\|_p^p \\ &= \sum_{\xi\in\chi_j} (\int |f(x)|^p g_{j,\xi}(x)|d\mu(x)) \|g_{j,\xi}\|_1^{p-1} \|g_{j,\xi}\|_p^p \\ &\le \int |f(x)|^p \left(\sum_{\xi\in\chi_j} |g_{j,\xi}(x)| \|g_{j,\xi}\|_1^{p-1} \|g_{j,\xi}\|_1 \|g_{j,\xi}\|_\infty^{p-1}\right) d\mu(x) \\ &\le C^{p-1} \int |f(x)|^p \left(\sum_{\xi\in\chi_j} |g_{j,\xi}(x)| \|g_{j,\xi}\|_1\right) d\mu(x) \le C^p \|f\|_p^p,\end{aligned}$$

using (12).

**Proof of (14)** For $p=1$ and even $0<p\le 1$,

$$\|\sum_{\xi\in\chi_j}\lambda_\xi g_{j,\xi}\|_p^p \le \sum_\xi \|\lambda_\xi g_{j,\xi}\|_p^p$$

and

$$\|\sum_{\xi\in\chi_j}\lambda_\xi g_{j,\xi}\|_\infty \le \|\sum_\xi |\lambda_\xi|\|g_{j,\xi}\|_\infty \frac{|g_{j,\xi}|}{\|g_{j,\xi}\|_\infty}\|_\infty \le \frac{C}{c}\sup_\xi |\lambda_\xi|\|g_{j,\xi}\|_\infty$$

as

$$\sum_\xi \frac{|g_{j,\xi}(x)|}{\|g_{j,\xi}\|_\infty} \le \sum_\xi |g_{j,\xi}(x)|\frac{\|g_{j,\xi}\|_1}{c} \le \frac{C}{c}. \tag{16}$$

Now for $1<p<\infty$, $\frac{1}{p}+\frac{1}{q}=1$,

$$\begin{aligned}
|\sum_{\xi\in\chi_j}\lambda_\xi g_{j,\xi}(x)|^p &\le \left(\sum_{\xi\in\chi_j}|\lambda_\xi|\|g_{j,\xi}\|_\infty\Big(\frac{|g_{j,\xi}(x)|}{\|g_{j,\xi}\|_\infty}\Big)^{1/p}\Big(\frac{|g_{j,\xi}(x)|}{\|g_{j,\xi}\|_\infty}\Big)^{1/q}\right)^p \\
&\le \left(\sum_{\xi\in\chi_j}|\lambda_\xi|^p\|g_{j,\xi}(x)\|_\infty^p\frac{|g_{j,\xi}(x)|}{\|g_{j,\xi}\|_\infty}\right)\left(\sum_{\xi\in\chi_j}\frac{|g_{j,\xi}|}{\|g_{j,\xi}\|_\infty}\right)^{p/q} \\
&\le \Big(\frac{C}{c}\Big)^{p-1}\sum_{\xi\in\chi_j}|\lambda_\xi|^p\|g_{j,\xi}\|_\infty^{p-1}|g_{j,\xi}(x)|
\end{aligned}$$

So

$$\int|\sum_{\xi\in\chi_j}\lambda_\xi g_{j,\xi}(x)|^p d\mu(x) \le \Big(\frac{C}{c}\Big)^{p-1}\sum_{\xi\in\chi_j}|\lambda_\xi|^p\|g_{j,\xi}\|_\infty^{p-1}\|g_{j,\xi}\|_1 \le \Big(\frac{C}{c}\Big)^{p-1}\Big(\frac{C}{c}\Big)^p\sum_{\xi\in\chi_j}|\lambda_\xi|^p\|g_{j,\xi}\|_p^p$$

using the following lemma : □

**Lemma 1.** *Under the condition (11), we have, if* $1\le p\le\infty$, $\frac{1}{p}+\frac{1}{q}=1$ :

$$\forall j\in\mathbb{N},\xi\in\chi_j,\quad \frac{c}{C}\|g_{j,\xi}\|_1^{1/p}\|g_{j,\xi}\|_\infty^{1/q} \le \|g_{j,\xi}\|_p \le \|g_{j,\xi}\|_1^{1/p}\|g_{j,\xi}\|_\infty^{1/q} \tag{17}$$

**Proof of the lemma** The right hand side inequality is always true by interpolation, as $\frac{1}{p}=\frac{1}{p}\frac{1}{1}+\frac{1}{\infty}\frac{1}{q}$.
For the left hand side inequality, again by interpolation, as $\frac{1}{2}=\frac{1}{2}\frac{1}{p}+\frac{1}{2}\frac{1}{q}$,

$$c\le\|g_{j,\xi}\|_2^2\le\|g_{j,\xi}\|_p\|g_{j,\xi}\|_q. \tag{18}$$

We have, using (18), the right hand side inequality with $q$ instead of $p$, and (18) again:

$$\begin{aligned}
c\|g_{j,\xi}\|_1^{1/p}\|g_{j,\xi}\|_\infty^{1/q} &\le \|g_{j,\xi}\|_p\|g_{j,\xi}\|_q\|g_{j,\xi}\|_1^{1/p}\|g_{j,\xi}\|_\infty^{1/q} \\
&\le \|g_{j,\xi}\|_p\|g_{j,\xi}\|_1^{1/q}\|g_{j,\xi}\|_\infty^{1/p}\|g_{j,\xi}\|_1^{1/p}\|g_{j,\xi}\|_\infty^{1/q} \\
&= \|g_{j,\xi}\|_p\|g_{j,\xi}\|_1\|g_{j,\xi}\|_\infty \le C\|g_{j,\xi}\|_p
\end{aligned}$$

If we recall the following 'projections on the multiresolution spaces',

$$A_j(f) = \sum_k a(\frac{k}{2^j})L_k(f) = \sum_{\xi\in\chi_j} \langle f, \phi_{j,\xi}\rangle \phi_{j,\xi} \tag{19}$$

$$B_j(f) = A_{j+1}(f) - A_j(f) = \sum_{\xi\in\chi_j} \langle f, \psi_{j\xi}\rangle \psi_{j\xi}, \tag{20}$$

a major consequence of Theorem 1 is the following corollary

**Corollary 1.** *For $f \in \mathbb{L}_p$, $1 \le p \le \infty$, (with the usual modification for $p = \infty$)*

$$\frac{1}{C}\|A_j(f)\|_p \le (\sum_{\xi\in\chi_j} |\langle f, \phi_{j,\xi}\rangle|^p \|\phi_{j,\xi}\|_p^p)^{1/p} \le C\|A_{j+1}(f)\|_p \tag{21}$$

$$\frac{1}{C}\|B_j(f)\|_p \le (\sum_{\xi\in\chi_j} |\langle f, \psi_{j\xi}\rangle|^p \|\psi_{j\xi}\|_p^p)^{1/p} \le C(\|B_{j-1}(f)\|_p + \|B_j(f)\|_p + \|B_{j+1}(f)\|_p) \tag{22}$$

**Proof of the corollary** We have $\langle f, \phi_{j,\xi}\rangle = \langle A_{j+1}f, \phi_{j,\xi}\rangle$; (21) is a direct consequence of theorem 1 .
In the same way, $\langle f, \psi_{j\xi}\rangle = \langle (B_{j-1}(f) + B_j(f) + B_{j+1}(f)), \psi_{j\xi}\rangle$; (22)is also a direct consequence of theorem 1 .

### 3.2 Definition of Besov Spaces

We define the Besov spaces in this context as spaces of approximation (as usual Besov spaces may also be defined). For a more complete description see [17],[16], [18]. Let $\Pi_n = \bigoplus_{k=0}^n \mathcal{H}_k$, be the space spanned by all the $\mathcal{H}_k$'s up to $n$. We define, for a function $f : \mathcal{Y} \mapsto \mathbb{R}$, its best $\mathbb{L}_p$-approximation by

$$E_n(f,p) = \inf_{P\in\Pi_n} \|f - P\|_p$$

and the associated Besov spaces by the following definition.

**Definition 2.** *For $0 < s < \infty$, $1 \le p \le \infty$ and $0 < q \le \infty$, we define the space $B_{p,q}^{s,0}$ as the space of functions $f \in \mathbb{L}_p$ such that :*

$$\left(\sum_{n\ge1} (n^s E_n(f,p))^q \frac{1}{n}\right)^{1/q} < \infty$$

*(if $q = \infty$, $\sup_{n\ge1} n^s E_n(f,p) < \infty$). This space is equipped with the following norm,*

$$\|f\|_{B_{p,q}^s} = \|f\|_p + \left(\sum_{n\ge1} (n^s E_n(f,p))^q \frac{1}{n}\right)^{1/q}$$

*(with the obvious modification for $q = \infty$.)*

For instance, if $p = q = \infty$, the space defined above is the space of functions which can be uniformly approximated using elements of $\Pi_n$ at the speed $n^{-s}$.

*Remarks* As $A_j f = \sum_{\xi \in \chi_j} \langle f, \phi_{j,\xi} \rangle \phi_{j,\xi} \in \Pi_{2^j}$,

$$E_{2^j}(f,p) \leq \|f - A_j(f)\|_p$$

and on the other hand $P \in \Pi_{2^{j-1}} \Longrightarrow A_j(P) = P$, hence, $\|f - A_j f\|_p = \|f - P + P - A_j f\|_p \leq \|f - P\|_p + \|A_j(P - f)\|_p \leq (1 + C_0 C_0')\|f - P\|_p = K\|f - P\|_p$. So we have

$$E_{2^j}(f,p) \leq \|f - A_j(f)\|_p \leq K E_{2^{j-1}}(f,p) \tag{23}$$

On the other hand, as $n \mapsto E_n(f,p)$ is obviously non increasing, we have, using the condensation principle :

$$\|f\|_{B^s_{p,q}} \sim \|f\|_p + \|(2^{js} E_{2^j}(f,p))_{j \in \mathbb{N}}\|_{l_q(\mathbb{N})}$$

This leads to the following characterisation of Besov spaces.

### 3.2.1 Characterisation of Besov spaces

If we recall the definition of $A_j(f)$ and $B_j(f)$, in (19), (20), we have the following characterisation.

**Theorem 2.** *For $f \in \mathbb{L}_p$, with the following decomposition*

$$f = \langle f, 1 \rangle + \sum_{j \geq 0} \sum_{\xi \in \chi_j} \langle f, \psi_{j\xi} \rangle \psi_{j\xi}$$

*the following properties are equivalent:*

1. $f \in B^s_{p,q}$.
2. $\|f\|_p + \|(2^{js}\|f - A_j(f)\|_p)_{j \in \mathbb{N}}\|_{l_q(j \in \mathbb{N})} < \infty$.
3. $\|f\|_p + \|(2^{js}\|B_j(f)\|_p)_{j \in \mathbb{N}}\|_{l_q(j \in \mathbb{N})} < \infty$.
4. $\|f\|_{B^s_{p,q}} \sim \|f\|_p + \|(2^{js}(\sum_{\xi \in \chi_j} |\langle f, \psi_{j\xi} \rangle|^p \|\psi_{j\xi}\|_p^p))_{j \in \mathbb{N}}\|_{l_q(j \in \mathbb{N})} < \infty$.
   *with equivalence of the induced norms.*

**Proof of the theorem** The equivalence of 1 and 2 comes from (23). As $B_j(f) = A_{j+1} - A_j(f) = (A_{j+1} - f) + (A_j(f) - f)$, we get easily $2 \Rightarrow 3$. 3 and 4 are equivalent from (22). It remains to prove that 3 implies 2. This is clear from :

$$\|f - A_j(f)\|_p = \|\sum_{l=j}^{\infty} B_l(f)\|_p \le \sum_{l=j}^{\infty} \|B_l(f)\|_p$$

But

$$\|B_l(f)\|_p = \alpha_l 2^{-ls}, \quad \alpha \in l_q.$$

So

$$\sum_{l=j}^{\infty} \|B_l(f)\|_p = \sum_{l=j}^{\infty} \alpha_l 2^{-ls} = 2^{-js}(\sum_{l=j}^{\infty} \alpha_l 2^{-(l-j)s}) = 2^{-js}\delta_j$$

and by convolution inequality $\delta \in l_q$.

As the needlet system is a frame and not a basis, we need additionally to prove the following theorem concerning a decomposition where the $\lambda_{j,\xi}$ are not necessarily the 'needlet coefficients' $\langle f, \psi_{j\xi}\rangle$.

**Theorem 3.** *For $s > 0$, $\beta \in \mathbb{C}$ and $(\lambda_{j,\xi})_{j\in\mathbb{N},\xi\in\chi_j} \in \mathbb{C}$, such that*

$$(\sum_{\xi\in\chi_j} |\lambda_{j,\xi}|^p \|\psi_{j\xi}\|_p^p)^{1/p} = \alpha_j 2^{-js}, \quad \alpha_. \in l_q(\mathbb{N}),$$

*then*

$$f = \beta + \sum_{j\in\mathbb{N}}(\sum_{\xi\in\chi_j} \lambda_{j,\xi}\psi_{j\xi}) \in B^s_{p,q}$$

*and*

$$\|f\|_{B^s_{p,q}} \le D(|\beta| + \|\alpha\|_{l_q})$$

**Proof of the theorem** Let $\Delta_j(f) = \sum_{\xi\in\chi_j} \lambda_{j,\xi}\psi_{j\xi}$. Clearly : $\|\Delta_j(f)\|_p \le C\alpha_j 2^{-js}$ and

$$\|f\|_p \le |\beta| + \sum_{j\ge 0} \|\Delta_j(f)\|_p \le |\beta| + C\sum_{j\ge 0} \alpha_j 2^{-js} \le |\beta| + C'\|\alpha\|_{l_q}$$

Moreover

$$\langle f, \psi_{j\xi}\rangle = \langle \Delta_{j-1}(f), \psi_{j\xi})\rangle + \langle \Delta_j(f), \psi_{j\xi})\rangle + \langle \Delta_{j+1}(f), \psi_{j\xi})\rangle$$

$$(\sum_{\xi\in\chi_j} |\langle f, \psi_{j\xi}\rangle|^p \|\psi_{j\xi}\|_p^p)^{1/p} \le C(\|\Delta_{j-1}(f)\|_p + \|\Delta_j(f)\|_p + \|\Delta_{j+1}(f)\|_p).$$

*Remark:* Of course it is not true that

$$\|f\|_{B^s_{pq}} \sim (|\beta| + \|\alpha\|_{l_q})$$

since it is possible to have $\sum_{\xi\in\chi_j} \lambda_{j,\xi}\psi_{j\xi} = 0$, and $\sum_{\xi\in\chi_j} |\lambda_{j,\xi}|^p \|\psi_{j\xi}\|_p^p > 0$.

### 3.3 Besov embeddings

It is a key point for approximation properties as well as statistical rates of convergence to clarify how the spaces defined above may be included in each others. As will be seen, the embeddings will parallel the standard embeddings of usual Besov spaces, but with important differences which yield new minimax rates of convergence for instance as detailed in the next section.

We begin with the following proposition which provides an important tool for comparison of the different $\mathbb{L}_p$ norms of the needlets.

**Proposition 2.** *Let $(\psi_{j\xi})_{j\in\mathbb{N},\xi\in\chi_j}$ a tight frame verifying condition (11) and*

$$\forall j \in \mathbb{N}, \xi \in \chi_j, \quad \|\psi_{j\xi}\|_1 \|\psi_{j\xi}\|_\infty \leq C < \infty. \text{ Then} \tag{24}$$

$$\forall\, 1 \leq r \leq 2, \;\; c^{1/r}\|\psi_{j\xi}\|_\infty^{1-2/r} \leq \|\psi_{j\xi}\|_r \leq C^{1/r}\|\psi_{j\xi}\|_\infty^{1-2/r} \tag{25}$$

$$\forall\, 2 \leq r \leq \infty, \;\; \frac{c}{C}\|\psi_{j\xi}\|_\infty^{1-2/r} \leq \|\psi_{j\xi}\|_r \leq \|\psi_{j\xi}\|_\infty^{1-2/r} \tag{26}$$

$$\forall\, 0 < r \leq 2, \;\; c^{1/r}\|\psi_{j\xi}\|_\infty^{1-2/r} \leq \|\psi_{j\xi}\|_r \tag{27}$$

**Proof of the proposition** By interpolation, if $1 \leq r \leq \infty$ ($\frac{1}{r} = \frac{\theta}{1} + \frac{\theta}{\infty}$, $(\theta = \frac{1}{r})$)and using (24),

$$\|\psi_{j\xi}\|_r \leq \|\psi_{j\xi}\|_1^{1/r}\|\psi_{j\xi}\|_\infty^{1-1/r} \leq C^{1/r}\|\psi_{j\xi}\|_\infty^{1-2/r}.$$

Again by interpolation, if $1 \leq q \leq 2 \leq p \leq \infty$, $\frac{1}{p} + \frac{1}{q} = 1$ ($\frac{1}{2} = \frac{1}{2}\frac{1}{p} + \frac{1}{2}\frac{1}{q}$,) and using (11) and (24),

$$c \leq \|\psi_{j\xi}\|_2^2 \leq \|\psi_{j\xi}\|_p\|\psi_{j\xi}\|_q \leq \|\psi_{j\xi}\|_1\|\psi_{j\xi}\|_\infty \leq C. \tag{28}$$

If $2 \leq p \leq \infty$, $\frac{1}{p} = \frac{\theta}{2} + \frac{1-\theta}{\infty}$, $(\theta = \frac{2}{p})$, by (11),

$$\|\psi_{j\xi}\|_p \leq \|\psi_{j\xi}\|_2^{2/p}\|\psi_{j\xi}\|_\infty^{1-2/p} \leq \|\psi_{j\xi}\|_\infty^{1-2/p}$$

Moreover, if $0 < q \leq 2$, $\frac{1}{2} = \frac{\theta}{q} + \frac{1-\theta}{\infty}$, $\theta = \frac{q}{2}$

$$c \leq \|\psi_{j\xi}\|_2^2 \leq \|\psi_{j\xi}\|_q^q\|\psi_{j\xi}\|_\infty^{2-q}$$

So

$$\forall\, 0 < q \leq 2, \quad c\|\psi_{j\xi}\|_\infty^{q-2} \leq \|\psi_{j\xi}\|_q^q. \tag{29}$$

On the other side,

$$c \leq \|\psi_{j\xi}\|_p\|\psi_{j\xi}\|_q \leq \|\psi_{j\xi}\|_p C^{1/q}\|\psi_{j\xi}\|_\infty^{1-2/q} = \|\psi_{j\xi}\|_p C^{1-1/p}\|\psi_{j\xi}\|_\infty^{-(1-2/p)}$$

$$\forall\, 2 \leq p \leq \infty, \quad \frac{c}{C}\|\psi_{j\xi}\|_\infty^{1-2/p} \leq C^{1/p}\frac{c}{C}\|\psi_{j\xi}\|_\infty^{1-2/p} \leq \|\psi_{j\xi}\|_p \leq \|\psi_{j\xi}\|_\infty^{1-2/p}.$$

Moreover by $\frac{1}{2} = \frac{\theta}{1} + \frac{1-\theta}{p}$, $\theta = \frac{p-2}{2(p-1)}$, $1 - \theta = \frac{p}{2(p-1)}$

$$c \leq \|\psi_{j\xi}\|_2^2 \leq \|\psi_{j\xi}\|_1^{2\theta}\|\psi_{j\xi}\|_p^{2(1-\theta)} \leq \Big(\frac{C}{\|\psi_{j\xi}\|_\infty}\Big)^{\frac{p-2}{p-1}}(\|\psi_{j\xi}\|_p^p)^{1/(p-1)}.$$

So

$$C(\frac{c}{C})^{p-1}\|\psi_{j\xi}\|_\infty^{p-2} \leq \|\psi_{j\xi}\|_p^p, \text{ hence}$$
$$\forall\, 2 \leq p \leq \infty, \quad (\frac{c}{C})^p\|\psi_{j\xi}\|_\infty^{p-2} \leq \|\psi_{j\xi}\|_p^p$$

□

We are now able to state the embeddings results.

**Theorem 4.** *Under the conditions of Proposition 2*

1. $1 \leq p \leq \pi \leq \infty \Rightarrow B^s_{\pi,r} \subseteq B^s_{p,r}$.
2. *If* $\sup_{\xi\in\mathcal{X}_j} \|\psi_{j\xi}\|_\infty^2 \leq C2^{j\mu}$ *then*

$$\infty \geq p \geq \pi > 0, \quad s > \mu(1/\pi - 1/p), \quad \Rightarrow B^s_{\pi,r} \subseteq B^{s-\mu(1/\pi-1/p)}_{p,r}$$

**Proof of the theorem** Recall that :

$$f \in B^s_{\pi,r} \Leftrightarrow \left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_\pi)^\pi\right)^{1/\pi} \leq \delta_j 2^{-js}, \quad \delta_. \in l_r(\mathbb{N}).$$

Writing, $\left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_p)^p\right)^{1/p} = \left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_\pi)^p(\frac{\|\psi_{j\xi}\|_p}{\|\psi_{j\xi}\|_\pi})^p\right)^{1/p}$, we have the following cases,

1. $1 \leq p \leq \pi \leq \infty$

$$\left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_p)^p\right)^{1/p} \leq \left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_\pi)^\pi\right)^{1/\pi}\left(\sum_{\xi\in\mathcal{X}_j}\left(\frac{\|\psi_{j\xi}\|_p}{\|\psi_{j\xi}\|_\pi}\right)^{\frac{1}{1/p-1/\pi}}\right)^{(1/p-1/\pi)}.$$

Using (17),

$$\frac{\|\psi_{j\xi}\|_p}{\|\psi_{j\xi}\|_\pi} \leq \frac{\|\phi_{j,\xi}\|_1^{1/p}\|\phi_{j,\xi}\|_\infty^{1-1/p}}{\frac{c}{C}\|\phi_{j,\xi}\|_1^{1/\pi}\|\phi_{j,\xi}\|_\infty^{1-1/\pi}} = \frac{C}{c}(\frac{\|\psi_{j\xi}\|_1}{\|\psi_{j\xi}\|_\infty})^{1/p-1/\pi}$$

$$\left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_p)^p\right)^{1/p} \leq \frac{C}{c}\left(\sum_{\xi\in\mathcal{X}_j}\frac{\|\psi_{j\xi}\|_1}{\|\psi_{j\xi}\|_\infty}\right)^{(1/p-1/\pi)}\left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_\pi)^\pi\right)^{1/\pi}.$$

But using (16)

$$\left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_p)^p\right)^{1/p} \leq \frac{C}{c}(\frac{C}{c}\mu(\mathcal{X}))^{1/p-1/\pi}\left(\sum_{\xi\in\mathcal{X}_j}(|\beta_{j,\xi}|\|\psi_{j\xi}\|_\pi)^\pi\right)^{1/\pi} \leq C(p,\pi)\delta_j 2^{-js}$$

(As $\mu(\mathcal{Y})) < \infty$).

2. $\infty \geq p \geq \pi > 0$; $\infty \geq p \geq 1$

$$\left(\sum_{\xi\in\chi_j} (|\beta_{j,\xi}|\|\psi_{j\xi}\|_p)^p\right)^{1/p} \leq \sup_{\xi\in\chi_j} \frac{\|\psi_{j\xi}\|_p}{\|\psi_{j\xi}\|_\pi} \left(\sum_{\xi\in\chi_j} (|\beta_{j,\xi}|\|\psi_{j\xi}\|_\pi)^\pi\right)^{1/\pi}$$

Using Proposition 2,

$$\frac{\|\psi_{j\xi}\|_p}{\|\psi_{j\xi}\|_\pi} \leq K(p,\pi)\|\psi_{j\xi}\|_\infty^{2(1/\pi-1/p)}$$

where $K(p,\pi) = \frac{C}{c}$, if $2 \leq \pi$, $K(p,\pi) = \frac{C^{1/p}}{c^{1/\pi}}$, if $0 < \pi < 2$.

$$\left(\sum_{\xi\in\chi_j} (|\beta_{j,\xi}|\|\psi_{j\xi}\|_p)^p\right)^{1/p} \leq K(p,\pi)(\sup_{\xi\in\chi_j}(\|\psi_{j\xi}\|_\infty^2))^{1/\pi-1/p} \left(\sum_{\xi\in\chi_j} (|\beta_{j,\xi}|\|\psi_{j\xi}\|_\pi)^\pi\right)^{1/\pi}$$

$$\leq K'(p,\pi)\delta_j 2^{-js}2^{j\mu(1/\pi-1/p)}.$$

using Proposition (2), and the hypothesis.

# 4 Minimax rates of convergence

**Density estimation on the sphere** Let us come back to the statistical examples presented in the introduction and consider the problem of estimating the density $f$ of an independent sample of points $X_1, \ldots, X_n$ observed on the $d$-dimensional sphere $\mathbb{S}^d$ of $\mathbb{R}^{d+1}$. Taking now advantage of the construction of needlets based on the spherical harmonic basis, we can build the following needlet estimator, using a hard thresholding of a needlet expansion as follows. We start by letting:

$$\hat{\beta}_{j\xi} := \frac{1}{n}\sum_{i=1}^{n} \psi_{j\xi}(X_i) \tag{30}$$

$$\hat{f} = \frac{1}{|\mathbb{S}^d|} + \sum_{j=0}^{J}\sum_{\xi\in\chi_j} \hat{\beta}_{j\xi}\,\psi_{j\xi}\,1_{\{|\hat{\beta}_{j\xi}|\geq \kappa t_n\}} \tag{31}$$

The tuning parameters of the needlet estimator are the following:

- The range of resolution levels (frequencies) where the approximation (31) is used:

$$2^J = \left(\frac{n}{\log n}\right)^{\frac{1}{d}}.$$

- The threshold constant $\kappa$ which is precised in the following theorem.
- $t_n$ is a sample size-dependent scaling factor.

$$t_n = \left(\frac{\log n}{n}\right)^{1/2}.$$

The following theorem is proved in [2], and gives the rates of convergence of the needlet estimator. It is also proved in [2] that the rates given here are optimal from the minimax point of view, up to logarithmic terms. As announced in the introduction, this procedure is minimax from $\mathbb{L}_2$ point of view but also performs satisfactorily from a local point of view (in infinity norm, for instance) and more generally for any $\mathbb{L}_p$ norm. In addition, the procedure is simple to implement, as well as adaptive to the regularity class (its construction does not involve the knowledge of the regularity) and also to inhomogeneous smoothness (the optimality of the rate stands for a very large class of Besov spaces). Of course, the Besov spaces used here are those defined in the previous section.

**Theorem 5.** *For $0 < \pi \leq \infty$, $p \geq 1$, $s > \frac{d}{r}$, so that $B^s_{\pi,q} \subset L^\infty$.*

*a) For any $z > 0$, there exist some constants $c_\infty = c_\infty(s,p,\pi,A,M)$ such that if $\kappa > \frac{z+1}{6}$,*

$$\sup_{f\in B^s_{\pi,q}(M)\cap\|f\|_\infty\leq A} \mathbb{E}\|\hat{f} - f\|^z_\infty \leq c_\infty(\log n)^{z-1}\left[\frac{n}{\log n}\right]^{\frac{-(s-\frac{d}{\pi})}{2\left(s-d(\frac{1}{\pi}-\frac{1}{2})\right)}} \tag{32}$$

*b) For $1 \leq p < \infty$ there exist some constant $c_p = c_p(s,\pi,p,A,M)$ such that if $\kappa > \frac{p}{12}$,*

$$\sup_{f\in B^s_{\pi,q}(M)\cap\|f\|_\infty\leq A} \mathbb{E}\|\hat{f} - f\|^p_p \leq c_p(\log n)^{p-1}(\log n)^{I\{\pi=\frac{dp}{2s+d}\}}\left[\frac{n}{\log n}\right]^{-\frac{(s-d(\frac{1}{\pi}-\frac{1}{p}))p}{2\left(s-d(\frac{1}{\pi}-\frac{1}{2})\right)}}, \quad \text{if } \pi \leq \frac{dp}{2s+d} \tag{33}$$

$$\sup_{f\in B^s_{\pi,q}(M)\cap\|f\|_\infty\leq A} \mathbb{E}\|\hat{f} - f\|^p_p \leq c_p(\log n)^{p-1}\left[\frac{n}{\log n}\right]^{\frac{-sp}{2s+d}}, \qquad \text{if } \pi > \frac{dp}{2s+d} \tag{34}$$

**Denoising-deblurring in a Jacobi context** This part is inspired by [9], where we consider recovering a function $f$, when receiving a blurred (by a linear operator) and noisy version : $Y_\varepsilon = Kf + \varepsilon\dot{W}$. We consider the particular case where the SVD basis is composed of Jacobi polynomials (with coefficients $\alpha \geq \beta > -\frac{1}{2}$) and the eigenvalues have a polynomial decreasing rate ($b_i \sim i^{-\nu}$, this notation meaning that there exist two constants $c$, $C$ such that for all $i$, $ci^{-\nu} \leq b_i \leq Ci^{-\nu}$, where $\nu$ is a known constant);

the Wicksell problem is a special example of this case. Taking again advantage of the construction of needlets based on the Jacobi basis, we can also provide here a needlet estimator, using a hard thresholding of a needlet expansion as follows. Let us denote for simplicity by $(e_i)$ and $(g_i)$ the SVD bases.

The needlet decomposition of any $f$ takes the form

$$f = \sum_j \sum_{\xi \in \chi_j} \langle f, \psi_{j\xi} \rangle \psi_{j\xi}.$$

Using Parseval's identity, we have $\beta_{j\xi} = \langle f, \psi_{j\xi} \rangle = \sum_i f^i \psi^i_{j\xi}$ with $f^i = \langle f, e_i \rangle$ and $\psi^i_{j\xi} = \langle \psi_{j\xi}, e_i \rangle$. If we put $Y^i = \langle Y_\varepsilon, g_i \rangle$, then

$$Y^i = \langle Kf, g_i \rangle + \varepsilon\eta^i = \langle f, K^* g_i \rangle + \varepsilon\eta^i = \langle \sum_j f^j e_j, K^* g_i \rangle + \varepsilon\eta^i = b_i f^i + \varepsilon\eta^i,$$

where the $\eta^i = \langle \dot{W}, g_i \rangle$ form a sequence of independent centered Gaussian variables with variance 1. Thus

$$\hat{\beta}_{j\xi} = \sum_i \frac{Y^i}{b_i} \psi^i_{j\xi}$$

is an unbiased estimate for $\beta_{j\xi}$. Notice that from the needlet construction (see the previous section) it follows that the sum above is finite. More precisely, $\psi^i_{j\xi} \neq 0$ only for $2^{j-1} < i < 2^{j+1}$.

Let us consider the following estimate of $f$:

$$\hat{f} = \sum_{j=-1}^{J} \sum_{\xi \in \chi_j} t(\hat{\beta}_{j\xi}) \psi_{j\xi},$$

where $t$ is a thresholding operator defined by

$$t(\hat{\beta}_{j\xi}) = \hat{\beta}_{j\xi} I\{|\hat{\beta}_{j\xi}| \geq \kappa t_\varepsilon \sigma_j\} \quad \text{with} \tag{35}$$

$$t_\varepsilon = \varepsilon \sqrt{\log \frac{1}{\varepsilon}} \tag{36}$$

$$\sigma_j^2 = \sum_i [\frac{\psi^i_{j\xi}}{b_i}]^2 \quad \forall\, j \geq 0 \tag{37}$$

**Theorem 6.** *Let us suppose that the properties above are verified with fixed $\alpha \geq \beta > -\frac{1}{2}$ and that $b_i \sim i^{-\nu}$, $\nu > -\frac{1}{2}$. We put,*

$$2^J = t_\epsilon^{-\frac{2}{1+2\nu}},$$

*and choose $\kappa^2 \geq 16p[1 + 4\{(\frac{\alpha}{2} - \frac{\alpha+1}{p})_+ \vee (\frac{\beta}{2} - \frac{\beta+1}{p})_+\}]$.*
*Then if $f \in B^s_{\pi,r}(M)$ , $s > sup_{\gamma \in \{\alpha,\beta\}}\{\frac{1}{2} - 2(\gamma+1)(\frac{1}{2} - \frac{1}{\pi}) \vee 2(\gamma+1)(\frac{1}{\pi} - \frac{1}{p}) \vee 0\}$*

$$\mathbb{E}\|\hat{f}-f\|_p^p \leq C[\log(1/\varepsilon)]^{p-1+a}[\varepsilon\sqrt{\log(1/\varepsilon)}]^{\mu p},$$

*with :*

$$\mu = inf\{\mu(s),\ \mu(s,\alpha),\ \mu(s,\beta)\} \tag{38}$$
$$\mu(s) = \frac{s}{s+\nu+\frac{1}{2}} \tag{39}$$
$$\mu(s,u) = \frac{s-2(1+u)(\frac{1}{\pi}-\frac{1}{p})}{s+\nu+2(1+u)(\frac{1}{2}-\frac{1}{\pi})} \tag{40}$$

$$a = \begin{cases} I\{\delta_p = 0\} & \textit{if } [p-\pi][1-(p-2)(\alpha+1/2)] \geq 0, \\ \frac{(\alpha+\frac{1}{2})(\pi-p)}{(\pi-2)(\alpha+1/2)-1} + I\{\delta_s = 0\} & \textit{if } [p-\pi][1-(p-2)(\alpha+1/2)] < 0, \end{cases}$$

*with* $\delta_p = 1-(p-2)(\alpha+1/2)$ *and* $\delta_s = s[1-(p-2)(\alpha+1/2)] - p(2\nu+1)(\alpha+1)(\frac{1}{\pi}-\frac{1}{p})$.

These results are proved in [9]. It is also proved there that the rates given here are optimal from the minimax point of view, up to logarithmic terms. This procedure has the same minimax and adaptation properties as the previous one. The rates obtained here are quite new. The novelty is coming from the behavior of the needlets -varying with the place of the needlet center (near the interval bounds or in the center- which depends on the coefficients $\alpha$ and $\beta$ of the Jacobi polynomials, andyields contrary to the standard Besov case different bounds for the $\mathbb{L}_p$ norms.

**Denoising the Radon transform** As before, we consider observations of the form

$$dY(\theta,s) = Rf(\theta,s)d\mu(\theta,s) + \varepsilon dW(\theta,s), \tag{41}$$

where $R$ here is the Radon transform. Following [7] and [8], we can use, as in the previous paragraph the needlets built on the SVD of the Radon tranform, to produce linear or nonlinear procedures. Let us discuss there the nonlinear procedure ([8]) as it is an adaptation to the Radon case of the thresholding procedure presented in the Jacobi case in the previous paragraph. Our estimator is based on an appropriate thresholding of a needlet expansion as before. $f$ can be decomposed using the frame above :

$$f = \sum_j \sum_{\xi\in\chi_j} \langle f, \psi_{j\xi}\rangle \psi_{j\xi},$$

Following [8], (recall that $g_{k,l,i}$ is defined at the end of section 1) we define in this case

$$\widehat{\beta}_{j\xi} = \sum_{k,l,i} \gamma_{k,l,i}^{j,\xi} \frac{1}{\lambda_k} \int g_{k,l,i} dY, \tag{42}$$

$$\gamma_{k,l,i}^{j,\xi} = \langle g_{k,l,i}, \psi_{j\xi} \rangle$$

$$\hat{f} = \sum_{j=-1}^{J} \sum_{\xi \in \chi_j} \widehat{\beta}_{j\xi} I_{\{|\widehat{\beta}_{j\xi}| \geq \kappa 2^{jd} t_\varepsilon\}} \psi_{j\xi} \tag{43}$$

Again, the tuning parameters of this estimator are

- The range $J = J_\varepsilon$ of resolution levels will be taken into account such that

$$2^{Jd} = [\varepsilon \sqrt{\log 1/\varepsilon}]^{-1}$$

- The threshold constant $\kappa$.
- $t_\varepsilon$ is a constant depending on the noise level. We shall see that the following choice is appropriate

$$t_\varepsilon = \varepsilon \sqrt{\log 1/\varepsilon}.$$

- Notice that the threshold function for each coefficient contains $2^{jd}$. This is due to the inversion of the Radon operator, and the concentration relative to the $g_{k,l,i}$'s of the needlets.

We will consider the minimax properties of this estimator on the Besov bodies as contructed above.

**Theorem 7.** *For* $0 < r \leq \infty$, $\pi \geq 1$, $1 \leq p < \infty$ *there exist some constant* $c_p = c_p(s, r, p, M)$, $\kappa_0 = \kappa_0(p)$ *such that if* $\kappa > \kappa_0$, $s > (d+1)(\frac{1}{\pi} - \frac{1}{p})_+$, *in addition with if* $\pi < p$, $s > \frac{d+1}{\pi} - \frac{1}{2}$

*1. If* $\frac{1}{p} < \frac{d}{d+1}$

$$\sup_{f \in B^s_{\pi,r}(M)} [\mathbb{E}\|\hat{f} - f\|_p^p]^{\frac{1}{p}} \leq c_p [\log 1/\epsilon]^p [(\epsilon \sqrt{\log 1/\epsilon})^{\frac{s-(d+1)(1/\pi - 1/p)}{s+d-(d+1)/\pi}} \vee (\epsilon \sqrt{\log 1/\epsilon})^{\frac{s}{s+d-1/2}}$$
$$\vee (\epsilon \sqrt{\log 1/\epsilon})^{\frac{s-2(1/\pi - 1/p)}{s+d-2/\pi}}]$$

*2. If* $\frac{d}{d+1} \leq \frac{1}{p} < \frac{5d-1}{4d+1}$

$$\sup_{f \in B^s_{\pi,r}(M)} [\mathbb{E}\|\hat{f} - f\|_p^p]^{\frac{1}{p}} \leq c_p [\log 1/\epsilon]^p [(\epsilon \sqrt{\log 1/\epsilon})^{\frac{s}{s+d-1/2}} \vee (\epsilon \sqrt{\log 1/\epsilon})^{\frac{s-2(1/\pi - 1/p)}{s+d-2/\pi}}]$$

*3. If* $\frac{5d-1}{4d+1} \leq \frac{1}{p}$

$$\sup_{f \in B^s_{\pi,r}(M)} [\mathbb{E}\|\hat{f} - f\|_p^p]^{\frac{1}{p}} \leq c_p[\log 1/\epsilon]^p(\epsilon\sqrt{\log 1/\epsilon})^{\frac{s}{s+d-1/2}}$$

This theorem is proved in [8]. It is also proved there that, up to logarithmic terms, the rates observed here are minimax. A close look to these rates of convergence and to the proof of the theorem reveals that, as in the case of Jacobi-type inverse problems, the forms of the rates are coming from the behavior of the needlets together with the coefficient of ill-posedness of the problem $\nu = \frac{d}{2} - 1$.

## References


[1] P. Baldi, G. Kerkyacharian, D. Marinucci, and D. Picard. Asymptotics for spherical needlets. *arXiv:math/0606599*, 2006.

[2] P. Baldi, G. Kerkyacharian, D. Marinucci, and D. Picard. Adaptive density estimation for directional data using needlets. *arXiv:0807.5059*, 2008.

[3] Paolo Baldi, Gérard Kerkyacharian, Domenico Marinucci, and Dominique Picard. High frequency asymptotics for wavelet-based tests for Gaussianity and isotropy on the torus. *J. Multivariate Anal.*, 99(4):606–636, 2008.

[4] D. Donoho, I. Johnstone, G. Kerkyacharian, and D. Picard. Wavelet shrinkage: asymptopia? *J. Roy. Statist. Soc. Ser. B*, 57(2):301–369, 1995.

[5] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952.

[6] I. Johnstone and B. Silverman. Discretization effects in statistical inverse problems. *J. Complexity*, 7(1):1–34, 1991.

[7] G. Kerkyacharian, G. Kyriazis, E. Le Pennec, P. Petrushev, and D. Picard. Inversion of noisy radon transform by svd based needlet. *arXiv:0809.3332v1 [math.ST]*, 2008.

[8] G. Kerkyacharian, E. Le Pennec, and D. Picard. Radon needlet thresholding. *technical report LPMA*, 2008.

[9] G. Kerkyacharian, P. Petrushev, D. Picard, and T. Willer. Needlet algorithms for estimation in inverse problems. *Electron. J. Stat.*, 1:30–76 (electronic), 2007.

[10] Peter T. Kim and Ja-Yong Koo. Optimal spherical deconvolution. *J. Multivariate Anal.*, 80(1):21–42, 2002.

[11] Peter T. Kim, Ja-Yong Koo, and Heon Jin Park. Sharp minimaxity and spherical deconvolution for super-smooth error distributions. *J. Multivariate Anal.*, 90(2):384–392, 2004.

[12] Jussi Klemelä. Estimation of densities and derivatives of densities with directional data. *J. Multivariate Anal.*, 73(1):18–40, 2000.

[13] Jussi Klemelä. Lower bounds for the asymptotic minimax risk with spherical data. *J. Statist. Plann. Inference*, 113(1):113–136, 2003.

[14] G. Kyriazis, P. Petrushev, and Y. Xu. Jacobi decomposition of weighted triebel-lizorkin and besov space. *Studia Math.*, 2008. to appear (arXiv:math/0610624).

[15] D. Marinucci, D. Pietrobon, P. Balbi, P. Baldi, G. Cabella, G. Kerkyacharian, P. Natoli, D. Picard, and N. Vittorio. Spherical needlets for cmb data analysis. *arXiv:0707.0844*, 2007.

[16] F. Narcowich, P. Petrushev, and J. Ward. Local tight frames on spheres. *SIAM J. Math. Anal.*, 2006.

[17] F. J. Narcowich, P. Petrushev, and J.M. Ward. Decomposition of besov and triebel-lizorkin spaces on the sphere. *J. Funct. Anal.*, 2006.

[18] P. Petrushev and Y. Xu. Localized polynomial frames on the interval with Jacobi weights. *J. Fourier Anal. Appl.*, 11(5):557–575, 2005.

[19] P. Petrushev and Y. Xu. Localized polynomials frames on the ball. *Constr. Approx.*, 27:121–148, 2008.