Localized spherical deconvolution

Gérard Kerkyacharian, Thanh Mai Pham Ngoc, Dominique Picard

LOCALIZED DECONVOLUTION ON THE SPHERE

By Gérard Kerkyacharian, Thanh Mai Pham
Ngoc and Dominique Picard

CNRS, LPMA and Université Paris VI and Université Paris VII

We provide a new algorithm for the treatment of deconvolution
on the sphere which combines the traditional SVD inversion with
an appropriate thresholding technique in a well chosen new basis.
We establish upper bounds for the behaviour of our procedure for
any $\mathbb{L}_p$ loss. It is important to emphasize the adaptation properties
of our procedures with respect to the regularity (sparsity) of the
object to recover as well as to inhomogeneous smoothness. We also
perform a numerical study which proves that the procedure shows
very promising properties in practice as well.

1. Introduction. The spherical deconvolution problem was first proposed by Rooij and Ruymgaart (1991) [14] and subsequently solved in Healy et al. (1998) [3]. Kim and Koo (2002) [7] established minimaxity for the $\mathbb{L}_2$-rate of convergence. The optimal procedures obtained there are using orthogonal series methods associated with spherical harmonics. One important problem arising with these procedures is their poor local performances due to the fact that spherical harmonics are spread all over the sphere. This explains for instance the fact that although they are optimal in the $\mathbb{L}_2$ sense, they cease to be optimal for other losses, such as $\mathbb{L}_p$ losses for instance.

In our approach we focus on two important points. We aim at a procedure of estimation which is efficient from a $\mathbb{L}_2$ point of view, as well as it performs satisfactorily from a local point of view (for other $\mathbb{L}_p$ losses for instance).

Deconvolution is an inverse problem and in such a problem there is a notable conflict between the inversion part which in presence of noise creates an instability reasonably handled by a Singular Value Decomposition (SVD) approach and the fact that the SVD basis very rarely is localized and capable of representing local features of images, which are especially important to recover. Our strategy is to follow the approach started in Kerkyacharian et al. (2007) [6] for the Wicksell case , Kerkyacharian et al. (2009) [5] for the Radon transform, which utilizes the construction borrowed from Narkowich Petrushev and Wald (2006) [9], [8] of a tight frame (i.e. a redundant family)

---

staying sufficiently close to the SVD decomposition but which enjoys at the same time enough localisation properties to be successfully used for statistical estimation (see for instance Baldi et al. (2009) Pietrobon et al. (2008) [1, 2, 10] for other types of applications). The construction [9] produces a family of functions which very much resemble to wavelets, the needlets.

To achieve the goals presented above, and especially adaptation to different regularities and local inhomogeneous smoothness, we essentially use a projection method on the needlets (which enables a stable inversion of the devonvolution, due to the closeness to the SVD basis) with a fine tuning subsequent thresholding process.

This provides a reasonably simple algorithm with very good performances, both from a theoretical point of view and a numerical point of view. In effect, this new algorithm provides a much better spatial adaptation, as well as adaptation to wider classes of regularity. We give here upper bounds obtained by the procedure over a large class of Besov spaces and any $\mathbb{L}_p$ losses.

It is important to notice that especially because we consider different $\mathbb{L}_p$ losses, we provide rates of convergence of new types attained by our procedure, which, of course, coincide with the usual ones for $\mathbb{L}_2$ losses.

Again, the problem of choosing appropriated spaces of regularity on the sphere in a serious question, and we decided to consider the spaces which may be the closest to our natural intuition: those which generalize to the sphere case the classical approximation properties of usual regularity spaces such as Holder spaces and include at the same time the Sobolev regularity spaces used in Kim and Koo (2002) [7].

Sphere deconvolution has a vast domain of application; our results are especially motivated by many recent developments in the area of observational astrophysics.

It is a common problem in astrophysics to analyse data sets consisting of a number of objects (such as galaxies of a particular type) or of events (such as cosmic rays or gamma ray bursts) distributed on the celestial sphere. In many cases, such objects trace an underlying probability distribution f on the sphere, which itself depends on the physics which governs the production of the objects and events.

The case for instance of ultra high energy cosmic rays (UHECR) iillustrates well the type of applications of our results. Ultra high energy cosmic rays are particles of unknown nature which arrive at the earth from apparently random directions of the sky. They could originate from long-lived relic particles from the Big Bang, about 13 billion years old. Alternatively, they could be generated by the acceleration of standard particles, such as protons,

in extremely violent astrophysical phenomena, such as cluster shocks. They could also originate from Active Galactic Nuclei (AGN), or from neutron stars surrounded by extremely high magnetic fields.

Hence, in some hypotheses, the underlying probability distribution for the directions of incidences of observed UHECRs would be a finite sum of point-like sources – or near point like, taken into account the deflection of the cosmic rays by magnetic fields. In other hypotheses, the distribution could be uniform, or smooth and correlated with the local distribution of matter in the universe. The distribution could also be a superposition of the above. Identifying between these hypotheses is of primordial importance for understanding the origin and mechanism of production of UHECRs.

Of course, the observations of these events (X_i's in the sequel) are always most often perturbated by a secondary noise (ε_i) which leads to the deconvolution problem described below. Following Healy et al. (1998) [3], Kim and Koo (2002) [7], the spherical deconvolution problem can be described as follows. Consider the situation where we observe $Z_1, \ldots, Z_N$ N i.i.d. observation with

$$Z_i = \varepsilon_i X_i, \tag{1}$$

where the ε_i's are i.i.d. random elements in $SO(3)$ (the group of 3×3 rotation matrices) , and the Z_i's and X_i's are i.i.d. random elements of $\mathbb{S}^2$ (two-dimensional unit sphere of $\mathbb{R}^3$) random elements, with ε_i and X_i assumed to be independent. We suppose that the distribution of resp. X, Z and ε are absolutely continuous with the Haar measure of resp. $\mathbb{S}^2$, $\mathbb{S}^2$ and $SO(3)$ with the densities f_Z, f_ε, f_X.

Then,

$$f_Z = f_\varepsilon * f_X, \tag{2}$$

where $*$ denotes convolution and is defined below. In the sequel, f_X will be denoted by f to emphasize the fact that it is the object to recover.

The following paragraph recall the necessary definitions. It is largely inspired by Kim and Koo (2002) [7] and and Healy et al. (1998) [3].

2. Some preliminaries about harmonic analysis on $SO(3)$ and $\mathbb{S}^2$. We will provide a brief overview of Fourier analysis on $SO(3)$ and $\mathbb{S}^2$. Most of the material can be found in expanded form in Vilenkin (1969) [15], Talman (1968) [12], Terras (1985) [13], Kim and Koo (2002) [7], and Healy et al. (1998) [3]. Let

$$u(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad a(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix},$$

4

where , $\phi \in [0, 2\pi)$, $\theta \in [0, \pi)$. It is well known that any rotation matrix can be decomposed as a product of three elemental rotations, one around the z-axis first with angle ψ, followed by a rotation around the y-axis with angle θ, and finally another rotation again around the z-axis with angle ϕ. Indeed, the well known Euler angle decomposition says that any $g \in SO(3)$ can almost surely be uniquely represented by three angles (ϕ, θ, ψ), with the following formula (see Healy et al. (1998) [3] for details) :

$$(3) \qquad g = u(\phi)a(\theta)u(\psi),$$

where $\phi \in [0, 2\pi)$, $\theta \in [0, \pi)$, $\psi \in [0, 2\pi)$. Consider the functions, known as the rotational harmonics,

$$(4) \qquad D^l_{mn}(\phi, \theta, \psi) = e^{-i(m\phi + n\psi)} P^l_{mn}(\cos \theta),$$

where the generalized Legendre associated functions P^l_{mn} for $-l \le m$, $n \le l$, $l = 0, 1, \dots$ are fully described in Vilenkin (1969) [15]. The functions D^l_{mn} for $-l \le m$, $n \le l$, $l = 0, 1, \dots$ are the eigenfunctions of the Laplace Beltrami operator on $SO(3)$, hence, $\sqrt{2l+1} D^l_{mn}$, $-l \le m$, $n \le l$, $l = 0, 1, \dots$ is a complete orthonormal basis for $\mathbb{L}_2(SO(3))$ with respect to the probability Haar measure. In addition, if we define the $(2l+1) \times (2l+1)$ matrices by

$$(5) \qquad D^l(g) = [D^l_{mn}(g)],$$

where for $-l \le m$, $n \le l$, $l = 0, 1, \dots$ and $g \in SO(3)$, they constitute the collection of inequivalent irreducible representations of $SO(3)$ (for further details see Vilenkin (1969) [15]).

Hence, for $f \in \mathbb{L}_2(SO(3))$, we define the rotational Fourier transform on $SO(3)$ by

$$(6) \qquad \hat{f}^l_{mn} = \int_{SO(3)} f(g) D^l_{mn}(g) dg,$$

where again we think of (5) as the matrix entries of the $(2l+1) \times (2l+1)$ matrix

$$\hat{f}^l = [\hat{f}^l_{mn}]_{-l \le m, \, n \le l}, \; l = 0, 1, \dots$$

and dg is the probability Haar measure on $SO(3)$. The rotational inversion can be obtained by

$$
\begin{aligned}
f(g) &= \sum_l \sum_{-l \le m, \, n \le l} \hat{f}^l_{mn} \overline{D^l_{mn}(g)} \\
(7) \qquad &= \sum_l \sum_{-l \le m, \, n \le l} \hat{f}^l_{mn} D^l_{mn}(g^{-1})
\end{aligned}
$$

(7) is to be understood in $\mathbb{L}_2$-sense although with additional smoothness conditions, it can hold pointwise.

A parallel spherical Fourier analysis is available on $\mathbb{S}^2$. Any point on $\mathbb{S}^2$ can be represented by

$$\omega = (cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta)^t,$$

with $, \phi \in [0, 2\pi), \theta \in [0, \pi)$. We also define the functions :

$$(8) \qquad Y_m^l(\omega) = Y_m^l(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P_m^l(\cos\theta) e^{im\varphi},$$

for $-l \leq m \leq l, \ l = 0, \ 1, \ldots, \ \phi \in [0, 2\pi), \ \theta \in [0, \pi)$ and where $P_m^l(\cos\theta)$ are the associated Legendre functions.

The functions Y_m^l obey

$$(9) \qquad\qquad\qquad Y_{-m}^l(\theta, \phi) = (-1)^m \bar{Y}_m^l(\theta, \phi).$$

The set $\{Y_m^l, \ -l \leq m \leq l, \ l = 0, \ 1, \ldots\}$ is forming an orthonormal basis of $\mathbb{L}_2(\mathbb{S}^2)$, generally referred to as the spherical harmonic basis.

Again, as above, for $f \in \mathbb{L}_2(\mathbb{S}^2)$, we define the spherical Fourier transform on $\mathbb{S}^2$ by

$$(10) \qquad\qquad\qquad \hat{f}_m^l = \int_{\mathbb{S}^2} f(\omega) \overline{Y_m^l(\omega)} d\omega,$$

where $d\omega$ is the probability Haar measure on the sphere $\mathbb{S}^2$. The spherical inversion can be obtained by

$$(11) \qquad\qquad\qquad f(\omega) = \sum_l \sum_{-l \leq m \leq l} \hat{f}_m^l Y_m^l(\omega).$$

The bases detailed above are important because they realize a singular value decomposition of the convolution operator created by our model. In effect, we define for $f_\varepsilon \in \mathbb{L}_2(SO(3)), \ f \in \mathbb{L}_2(\mathbb{S}^2)$ the convolution by the following formula:

$$f_\varepsilon * f(\omega) = \int_{SO(3)} f_\varepsilon(u) f(u^{-1}\omega) du$$

and we have for all $-l \leq m \leq l, \ l = 0, \ 1, \ldots,$

$$(12) \qquad\qquad\qquad \widehat{(f_\varepsilon * f)}_m^l = \sum_{n=-l}^{l} \hat{f}_{\varepsilon,mn}^l \hat{f}_n^l := (\hat{f}_\varepsilon^l \hat{f}^l)_m.$$

2.1. *The SVD Method.* The singular value method (see Healy et al. (1998) [3] and Kim Koo (2002) [7]) consists in expanding f in the spherical harmonics basis Y_m^l and estimating the spherical Fourier coefficients using the formula above (12). We get the following estimator of the spherical Fourier transform of f:

$$
\begin{aligned}
(13) \qquad \hat{f}_m^{l,N} &:= \frac{1}{N} \sum_{j=1}^{N} \sum_{n=-l}^{l} \hat{f}_{\varepsilon^{-1},mn}^{l} \bar{Y}_n^l(Z_j) \\
\hat{f}_{\varepsilon^{-1}}^{l} &:= (\hat{f}_{\varepsilon}^{l})^{-1},
\end{aligned}
$$

provided, of course, that these inverse matrices exist, and then the estimator of the distribution f is

$$
(14) \qquad f^N(\omega) = \sum_{l=0}^{\tilde{N}} \sum_{m=-l}^{l} \hat{f}_m^{l,N} Y_m^l(\omega),
$$

where $\tilde{N}$ depending on the number of observations has to be properly selected.

3. Needlet construction. This construction is due to Narcowich et al. (2006) [8]. Its aim is essentially to build a very well localized tight frame constructed using spherical harmonics, as discussed below. It was recently extended to more general euclidean settings with fruitful statistical applications (see Kerkyacharian et al. (2007) [6]), Baldi et al. (2009) Pietrobon et al. (2008) [1, 2, 10].As described above, we have the following decomposition:

$$
(15) \qquad \mathbb{L}_2(dx) = \bigoplus_{l=0}^{\infty} \mathbb{H}_l \ ,
$$

where $\mathbb{H}_l$ is the space of spherical harmonics of $\mathbb{S}^2$, of degree l (which dimension is $2l+1$).

The orthogonal projector on $\mathbb{H}_l$ can be written using the following the kernel operator

$$
(16) \qquad \forall f \ \in \mathbb{L}_2(dx), \ P_{\mathbb{H}_l} f(x) = \int_{\mathbb{S}^d} L_l(\langle x, y \rangle) f(y)\, dy
$$

where $\langle x, y \rangle$ is the standard scalar product of $\mathbb{R}^3$, and L_l is the Gegenbauer polynomial with parameter $\frac{1}{2}$ of degree l, defined on $[-1, +1]$ and normalized so that

$$(17) \qquad \int_{-1}^{1} L_l(t)L_k(t) = \frac{2l+1}{8\pi^2}$$

Let us point out the following reproducing property of the projection operators:

$$(18) \qquad \int_{\mathbb{S}^d} L_l(\langle x, y \rangle)L_k(\langle y, z \rangle) \, dy = \delta_{l,k} L_l(\langle x, z \rangle) \ .$$

The following construction is based on two fundamental steps: Littlewood-Paley decomposition and discretization, which are summarized in the two following subsections.

3.1. *Littlewood-Paley decomposition.* Let ϕ be a C^∞ function on $\mathbb{R}$, symmetric and decreasing on $\mathbb{R}^+$ supported in $|\xi| \leq 1$, such that $1 \geq \phi(\xi) \geq 0$ and $\phi(\xi) = 1$ if $|\xi| \leq \frac{1}{2}$.

$$b^2(\xi) = \phi(\tfrac{\xi}{2}) - \phi(\xi) \geq 0$$

so that

$$(19) \qquad \forall |\xi| \geq 1, \qquad \sum_{j \geq 0} b^2(\tfrac{\xi}{2^j}) = 1 \ .$$

Remark that $b(\xi) \neq 0$ only if $\frac{1}{2} \leq |\xi| \leq 2$. Let us now define the operator $\Lambda_j = \sum_{l \geq 0} b^2(\frac{l}{2^j})L_l$ and the associated kernel

$$\Lambda_j(x, y) = \sum_{l \geq 0} b^2(\tfrac{l}{2^j})L_l(\langle x, y \rangle) = \sum_{2^{j-1} < l < 2^{j+1}} b^2(\tfrac{l}{2^j})L_l(\langle x, y \rangle) \ .$$

We obviously have:

$$(20) \qquad \forall f \in \mathbb{L}_2(\mathbb{S}^2), \qquad f = \lim_{J \to \infty} L_0(f) + \sum_{j=0}^{J} \Lambda_j(f) \ .$$

and if $M_j(x, y) = \sum_{l \geq 0} b(\frac{l}{2^j})L_l(\langle x, y \rangle)$, then

$$(21) \qquad \Lambda_j(x, y) = \int M_j(x, z)M_j(z, y) \, dz \ .$$

3.2. *Discretization and localization properties.* Let us define

$$\mathscr{P}_l = \bigoplus_{m=0}^{l} \mathbb{H}_m \ ,$$

the space of the restrictions to S^2 of the polynomials of degree less than l.

The following quadrature formula is true: for all $l \in \mathbb{N}$ there exists a finite subset $\mathscr{X}_l$ of S^2 and positive real numbers $\lambda_\eta > 0$, indexed by the elements η of $\mathscr{X}_l$, such that

(22) $$\forall f \in \mathscr{P}_l, \qquad \int_{\mathbb{S}^2} f(x)\, dx = \sum_{\eta \in \mathscr{X}_l} \lambda_\eta f(\eta) \ .$$

Then the operator M_j defined in the subsection above is such that:

$$z \mapsto M_j(x, z) \in \mathscr{P}_{[2^{j+1}]} \ ,$$

so that

$$z \mapsto M_j(x, z) M_j(z, y) \in \mathscr{P}_{[2^{j+2}]} \ ,$$

and we can write:

$$\Lambda_j(x, y) = \int M_j(x, z) M_j(z, y) dz = \sum_{\eta \in \mathscr{X}_{[2^{j+2}]}} \lambda_\eta M_j(x, \eta) M_j(\eta, y) \ .$$

This implies:

$$\Lambda_j f(x) = \int \Lambda_j(x, y) f(y) dy = \int \sum_{\eta \in \mathscr{X}_{[2^{j+2}]}} \lambda_\eta M_j(x, \eta) M_j(\eta, y) f(y) dy$$

$$= \sum_{\eta \in \mathscr{X}_{[2^{j+2}]}} \sqrt{\lambda_\eta} M_j(x, \eta) \int \sqrt{\lambda_\eta} M_j(y, \eta) f(y) dy \ .$$

We denote

$$\mathscr{X}_{[2^{j+2}]} = \mathscr{Z}_j, \qquad \psi_{j,\eta}(x) := \sqrt{\lambda_\eta}\, M_j(x, \eta) \text{ for } \eta \in \mathscr{Z}_j \ ,$$

It can also be proved that the set of cubature points $\mathscr{X}_l$ can be chosen so that:

(23) $$\tfrac{1}{c}\, 2^{2j} \leq \#\mathscr{Z}_j \leq c\, 2^{2j}$$

for some $c > 0$. Actually in the simulations of §5 we make use of some sets of cubature points such that $\#\mathcal{Z}_j = 12.2^{2j}$ exactly. It holds, using (20)

$$f = L_0(f) + \sum_j \sum_{\eta \in \mathcal{Z}_j} \langle f, \psi_{j,\eta} \rangle_{\mathbb{L}_2(\mathbb{S}^d)} \psi_{j,\eta} \ .$$

The main result of Narcowich et al. (2006) [8] is the following localization property of the $\psi_{j,\eta}$, called needlets: for any $k \in \mathbb{N}$ there exists a constant c_k such that, for every $\xi \in \mathbb{S}^2$:

(24)
$$|\psi_{j,\eta}(\xi)| \le \frac{c_k 2^j}{(1 + 2^j d(\eta, \xi))^k} \ .$$

where d is the natural geodesic distance on the sphere $(d(\xi, \eta) = \arccos\langle \eta, \xi \rangle$). In other words needlets are almost exponentially localized around their associated cubature point, which motivates their name.

A major consequence of this localization property can be summarized in the following properties which will play an essential role in the sequel.

For any $1 \le p < \infty$, there exist positive constants c_p, C_p, c, C and D_p such that

(25)
$$c_p 2^{2j(\frac{p}{2}-1)} \le \|\psi_{j\eta}\|_p^p \le C_p 2^{2j(\frac{p}{2}-1)},$$

(26)
$$c2^j \le \|\psi_{j\eta}\|_\infty \le C2^j,$$

(27)
$$\Big\| \sum_{\eta \in \mathcal{Z}_j} \lambda_\eta \psi_{j,\eta} \Big\|_\pi \le c2^{2j(\frac{1}{2}-\frac{1}{\pi})} \Big(\sum_{\eta \in \mathcal{Z}_j} |\lambda_\eta|^\pi \Big)^{1/\pi}$$

(28)
$$\Big\| \sum_{\eta \in \mathcal{Z}_j} \lambda_\eta \psi_{j\eta} \Big\|_p^p \le D_p \sum_{\eta \in \mathcal{Z}_j} |\lambda_\eta|^p \|\psi_{j\eta}\|_p^p$$

(29)
$$\Big\| \sum_{\eta \in \mathcal{Z}_j} u_\eta \psi_{j\eta} \Big\|_\infty \le C \sup_{\eta \in \mathcal{Z}_j} |u_\eta| 2^j$$

To conclude this section, let us give a graphic representation of a spherical needlet in the spherical coordinates in order to illustrate the above theory. In the following graphic, we chose $j = 3$ and $\eta = 250$.
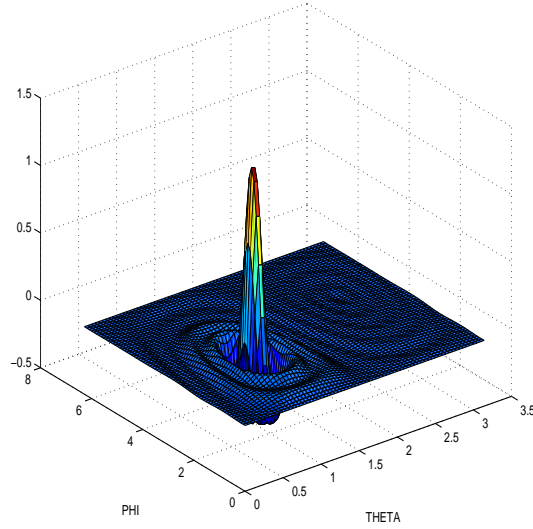
FIG 1. *A spherical needlet.*

3.3. *Besov spaces on the sphere.* The problem of choosing appropriated
spaces of regularity on the sphere in a serious question, and we decided to
consider the spaces which may be the closest to our natural intuition: those
which generalize to the sphere case the classical approximation properties
used to define for instance Sobolev spaces. In this section we summarize
the main properties of Besov spaces which will be used in the sequel, as
established in [8].
Let $f : \mathbb{S}^2 \to \mathbb{R}$ be a measurable function. We define

$$E_k(f, \pi) = \inf_{P \in \mathscr{P}_k} \|f - P\|_\pi,$$

the $\mathbb{L}_\pi$ distance between f and the space of polynomials of degree k. The
Besov space $B^s_{\pi,r}$ is defined as the space of functions such that

$$f \in \mathbb{L}_\pi \text{ and } \left(\sum_{k=0}^{\infty} (k^s E_k(f, \pi))^r \frac{1}{k} \right)^{1/r} < +\infty.$$

Remarking that $k \to E_k(f, \pi)$ is decreasing, by a standard condensation
argument this is equivalent to

$$f \in \mathbb{L}_\pi \text{ and } \left(\sum_{j=0}^{\infty} (2^{js} E_{2^j}(f, \pi))^r \right)^{1/r} < +\infty.$$

and the following theorem states that as it is the case for Besov spaces
in $\mathbb{R}^d$, the needlet coefficients are good indicators of the regularity and in

fact Besov spaces of $\mathbb{S}^2$ are Besov bodies, when expressed using the needlet expansion.

THEOREM 1. *Let* $1 \leq \pi \leq +\infty$, $s > 0$, $0 \leq r \leq +\infty$. *Let* f *a measurable function and define*

$$\langle f, \psi_{j,\eta} \rangle = \int_{\mathbb{S}^d} f(x)\psi_{j,\eta}(x)\, dx \stackrel{def}{=} \beta_{j,\eta},$$

provided the integrals exist. Then f *belongs to* $B^s_{\pi,r}$ *if and only if, for every* $j = 1, 2, \ldots,$

$$\Big(\sum_{\eta \in \mathscr{X}_j} (|\beta_{j,\eta}| \|\psi_{j,\eta}\|_\pi)^\pi \Big)^{1/\pi} = 2^{-js}\delta_j,$$

where $(\delta_j)_j \in \ell_r$.

As has been seen above,

$$c2^{2j(\frac{1}{2}-\frac{1}{\pi})} \leq \|\psi_{j,\eta}\|_\pi \leq C2^{2j(\frac{1}{2}-\frac{1}{\pi})},$$

for some positive constants c, C, the Besov space $B^s_{\pi,r}$ turns out to be a Banach space associate to the norm

(30) $\qquad \|f\|_{B^s_{\pi,r}} := \|(2^{j[s+2(\frac{1}{2}-\frac{1}{\pi})]} \|(\beta_{j\eta})_{\eta \in \mathscr{X}_j}\|_{\ell_\pi})_{j\geq 0}\|_{\ell_r} < \infty, \quad$ and

Using standard arguments (reducing to comparisons of l_q norms), it is easy to prove the following embeddings:

$$B^s_{\pi,r} \quad \subset \quad B^s_{p,r} \quad \text{for } p \leq \pi$$

(31) $\qquad B^s_{\pi,r} \quad \subset \quad B^{s-2(\frac{1}{\pi}-\frac{1}{p})}_{p,r} \quad \text{for } \pi \leq p \text{ and } s > 2(\frac{1}{\pi} - \frac{1}{p}).$

Moreover, it is also true that for $s > \frac{2}{\pi}$, if f belongs to $B^s_{\pi,r}$, then it is continuous, and as a consequence bounded.

In the sequel we shall denote by $B^s_{\pi,r}(M)$ the ball of radius M of the Besov space $B^s_{\pi,r}$.

4. Needlet algorithm: thresholding needlet coefficients. The first step is to construct a needlet system (frame) $\{\psi_{j\eta} : \eta \in \mathscr{Z}_j, j \geq -1\}$ as described in section 3.

The needlet decomposition of any $f \in \mathbb{L}_2(\mathbb{S}^2)$ takes the form

$$f = \sum_j \sum_{\eta \in \mathscr{Z}_j} (f, \psi_{j\eta})_{\mathbb{L}_2(\mathbb{S}^2)} \psi_{j\eta}.$$

Using Parseval's identity, we have $\beta_{j\eta} = (f, \psi_{j\eta})_{\mathbb{L}_2(\mathbb{S}^2)} = \sum_{lm} \hat{f}_m^l \psi_{j\eta}^{lm}$ with $\hat{f}_m^l = (f, Y_m^l)$ and $\psi_{j\eta}^{lm} = (\psi_{j\eta}, Y_m^l)$.

Thus

$$\hat{\beta}_{j\eta} = \sum_{lm} \hat{f}_m^{l,N} \psi_{j\eta}^{lm}, \tag{32}$$

is an unbiased estimate of $\beta_{j\eta}$. We recall that $\hat{f}_m^{l,N}$ has been defined in (13).

Notice that from the needlet construction (see the previous section) it follows that the sum above is finite. More precisely, $\psi_{j\eta}^{lm} \neq 0$ only for $2^{j-1} < l < 2^{j+1}$.

Let us consider the following estimate of f:

$$\hat{f} = \sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} t(\hat{\beta}_{j\eta}) \psi_{j\eta},$$

where t is a thresholding operator defined by

$$t(\hat{\beta}_{j\eta}) = \hat{\beta}_{j\eta} I\{|\hat{\beta}_{j\eta}| \geq \kappa t_N |\sigma_j|\} \quad \text{with} \tag{33}$$

$$t_N = \sqrt{\frac{\log N}{N}}, \tag{34}$$

$$\sigma_j^2 = M^2 \sum_{ln} |\sum_m \psi_{j\eta}^{lm} \hat{f}_{\varepsilon^{-1}mn}^l|^2. \tag{35}$$

Here κ is a tuning parameter of the method which will be properly selected later on. M is such that $\|f\|_\infty \leq M$. Notice that the thresholding depends on the resolution level j through the constant σ_j which will also be specified later on, and the same with regard to the upper level of details J.

4.1. Performances of the procedure.

THEOREM 2. *Let $1 \leq p < \infty$, $\nu > 0$, and let us assume that*

$$\sigma_j^2 := M^2 \sum_{ln} |\sum_m \psi_{j\eta}^{lm} \hat{f}_{\varepsilon^{-1}mn}^l|^2 \leq C 2^{2j\nu}, \ \forall \, j \geq 0. \tag{36}$$

Let us take $\kappa^2 \geq 16p$ and $2^J = d[t_N]^{\frac{-1}{(\nu+1)}}$ with t_N as in (34) and d is a positive constant. Then if $\pi \geq 1$, $s > 2/\pi$, $r \geq 1$ (with the restriction $r \leq \pi$ if $s = (\nu+1)(\frac{p}{\pi} - 1)$), there exists a constant C such that:

$$\sup_{f \in B_{\pi,r}^s(M)} \mathbb{E}\|\hat{f} - f\|_p^p \leq C(\log(N))^{p-1}[N^{-1/2}\sqrt{\log(N)}]^{\mu p}, \tag{37}$$

where

$$\mu = \frac{s}{s+\nu+1}, \quad if \ s \geq (\nu+1)(\frac{p}{\pi}-1)$$

$$\mu = \frac{s-2/\pi+2/p}{s+\nu-2/\pi+1}, \quad if \ \frac{2}{\pi} < s < (\nu+1)(\frac{p}{\pi}-1).$$

The proof of this theorem is given in section 6.

Remarks

1. The rates of convergence found here are standard in inverse problems. They can be related to rates found in Kim and Koo (2002) in the same deconvolution problem, with a $\mathbb{L}_2$ loss and constraints on the spaces comparable to $B_{s2}^2(M)$. In the deconvolution problem on the interval, similar rates are found even for $\mathbb{L}_p$ losses (with standard modifications since the dimension here is 2 instead of 1): see for instance Johnstone et al. (2004) [4]. These results are proved to be minimax (see Kim and Koo (2002) [7]) up to logarithmic factors, for the case $p = 2$ with a $B_{s2}^2(M)$ constraint on the object to estimate.
2. It is worthwhile to notice that the procedure is adaptive, meaning that it does not require a priori knowledge on the regularity (or sparsity) of the function. It also adapt to non homogeneous smoothness of the function. The logarithmic factor is a standard price to pay for adaptation.
3. The parameter ν appearing here is often called degree of ill-posedness of the problem (DIP). It appears here through condition (36) which is essential in this problem. In [7] for instance, and very often in diverse inverse problems, this DIP parameter is introduced with the help of the eigenvalues of the operator (i.e. here the discrepancy of the coefficients of f_ε in its expansion along the spherical harmonics). In the following subsection, we prove that (36) is in fact a consequence of the standard 'ordinary smooth' condition.

4.2. *Condition (36) and the smoothness of f_ε.* Following Kim and Koo (2002) [7] (condition 3.6), we can define the smoothness of f_ε spectrally. We place ourselves in the 'ordinary smooth' case

$$(38) \qquad \|(\hat{f}_\varepsilon^l)^{-1}\|_{op} \leq d_0 l^\nu \quad \text{and} \quad \|\hat{f}_\varepsilon^l\|_{op} \leq d_1 l^{-\nu} \quad \text{as} \quad l \to \infty,$$

14

for some positive constants d_0, d_1 and nonnegative constant ν, and where the operator norm of the rotational Fourier transform $\hat{f}^l_\varepsilon$ is defined as

$$\|\hat{f}^l_\varepsilon\|_{op} = \sup_{h\neq 0,\, h\in\mathcal{E}_l} \frac{\|\hat{f}^l_\varepsilon h\|_2}{\|h\|_2},$$

$\mathcal{E}_l$ being the $(2l+1)$-dimensional vector space spanned by $\{Y^l_m : -l \leq m \leq l\}$. The following proposition states that condition (35) is satisfied in the ordinary smooth case by the needlets system.

PROPOSITION 1.   *If $\|\hat{f}^l_\varepsilon\|_{op} \leq d_0 l^{-\nu}$, then*

$$|\sigma_j|^2 := M^2 \sum_{ln} |\sum_m \psi^{lm}_{j\eta} \hat{f}^l_{\varepsilon^{-1}mn}|^2 \leq C2^{2j\nu},\ \forall\, j \geq 0.$$

Proof. Since $\psi^{lm}_{j\eta} \neq 0$ only for $2^{j-1} < l < 2^{j+1}$,

$$\sum_{l=2^{j-1}}^{2^{j+1}} \sum_{n=-l}^{l} |\sum_m \psi^{lm}_{j\eta} \hat{f}^l_{\varepsilon^{-1}mn}|^2 = \sum_{l=2^{j-1}}^{2^{j+1}} \|(\hat{f}^l_\varepsilon)^{-1}\psi_{j\eta}\|_2^2$$

$$\leq \sum_{l=2^{j-1}}^{2^{j+1}} \|(\hat{f}^l_\varepsilon)^{-1}\|_{op}^2 \|\psi_{j\eta}\|_2^2$$

which proves the result using inequality (25).

We now give a brief review of some examples of smooth distributions which are discussed in depth in Healy et *al* (1998) [3] and Kim and Koo (2002) [7].

4.2.1. *Rotational Laplace distribution.*   This distribution can be viewed as an exact analogy on $SO(3)$ of the Laplace distribution on $\mathbb{R}$. Spectrally, for some $\rho^2 > 0$, this distribution is characterized by

(39) $$\hat{f}^l_{\varepsilon,mn} = (1 + \rho^2 l(l+1))^{-1}\delta_{mn},$$

for $-l \leq m,n \leq l$ and $l = 0,1,...$, and where $\delta_{mn} = 1$ if $m = n$ and 0 otherwise.

4.2.2. *The Rosenthal distribution.*   This distribution has its origin in random walks in groups (for details see Rosenthal (1994) [11]).

If one considers the situation where f_ε is a p-fold convolution product of conjugate invariant random for a fixed axis, then, Rosenthal (1994) [11] (p.407) showed that

$$\hat{f}^l_{\varepsilon,mn} = \left(\frac{\sin(l+1/2)\theta}{(2l+1)\sin\theta/2} \right)^p \delta_{mn},$$

for $-l \leq m, n \leq l$ and $l = 0, 1, \dots$ and where $0 < \theta \leq \pi$ and $p > 0$.

5. Practical performances. In this section we produce the results of numerical experiments on the sphere $\mathbb{S}^2$. The sets of cubature points $\xi_{j\eta}$ in the simulations that follow have been generated with the HEALPix pixelisation. For each resolution level j, the HEALPix pixelisation gives 12.2^{2j} cubature points.

In the two examples below we considered samples of cardinality $N = 1500$. The maximal resolution level J is taken such that $J = (1/2) \log_2 \left(\frac{N}{\log N} \right)$. In order not to have more cubature points than observations we set $J = 3$ for $N = 1500$. We recall the expression of the estimate of the needlets coefficients of the density of interest:

$$(40) \quad \hat{\beta}_{j\eta} = \frac{1}{N} \sqrt{\lambda_{j\eta}} \sum_{l=2^{j-1}}^{2^{j+1}} b(l/2^j) \sum_{m=-l}^{l} Y^l_m(\xi_{j\eta}) \sum_{n=-l}^{l} \hat{f}^l_{\varepsilon^{-1},mn} \sum_{u=1}^{N} \bar{Y}^l_n(Z_u),$$

where the cubature weight is equal to $\lambda_{j\eta} = 4\pi/(12.2^{2j})$.

We replace the rotational Fourier transform $(\hat{f}^l_\varepsilon)_{mn} := \hat{f}^l_{\varepsilon,mn}$ (defined in (6)) by its empirical version. Indeed, we suppose that the noise is unknown, this is a situation which is very likely to occur for instance in the context of astrophysics.

We precise again that $\hat{f}^l_{\varepsilon^{-1},mn}$ denotes the (m,n) element of the matrix $(\hat{f}^l_\varepsilon)^{-1} := \hat{f}^l_{\varepsilon^{-1}}$ which is the inverse of the $(2l+1) \times (2l+1)$ matrix $(\hat{f}^l_\varepsilon)$. In order to get the empirical version $\hat{f}^{l,N}_{\varepsilon^{-1},mn}$ of $\hat{f}^l_{\varepsilon^{-1},mn}$, we have first to compute the empirical matrix $(\hat{f}^{l,N}_\varepsilon)$ then to inverse it to get the matrix $(\hat{f}^{l,N}_\varepsilon)^{-1} := \hat{f}^{l,N}_{\varepsilon^{-1}}$. The (m,n) entry of the matrix $(\hat{f}^{l,N}_\varepsilon)$ is given by the formula:

$$\hat{f}^{l,N}_{\varepsilon,mn} = 1/N \sum_{j=1}^{N} D^l_{m,n}(\varepsilon_j),$$

where the rotational harmonics $D^l_{m,n}$ have been defined in (4). The ε_j's are i.i.d realizations of the variable $\varepsilon \in SO(3)$.

For the generation of the random variable $\varepsilon \in SO(3)$, we chose the Oz as the rotation axis and an angle ϕ following a uniform law. Various cases

are investigated for the angle ϕ, that is to say a uniform law with different supports such as $[0, \pi/8]$, $[0, \pi/4]$, $[0, \pi/2]$. The more the support of the uniform law is large, the more the level noise will be intense.

This particular choice of rotation matrix entails in the decomposition of an element of SO(3) (see formula (3)) the angles ψ and θ to be both equal to zero. For this specific setting of perturbation, we deduce the following form for the rotational harmonics:

$$D^l_{m,n}(\varepsilon_j) = P^l_{mn}(1)e^{-i\phi_j} = \delta_{mn}e^{-i\phi_j},$$

where $\phi_j \sim U[0, a]$ and a is a positive constant which will be specified later. For the reconstruction of the density f, we recall that

$$\hat{f} := \frac{1}{|\mathbb{S}^2|} + \sum_{j=0}^{J} \sum_{\eta=1}^{12.2^{2j}} \hat{\beta}_{j\eta}\psi_{j\eta}I_{\{|\hat{\beta}_{j\eta}|\geq\kappa t_n\sigma_j\}},$$

the expression of σ_j is given in (35), and $t_N = \sqrt{\log(N)/N}$. The constant M which appears in the expression of σ_j is such that $\|f\|_\infty \leq M$.

Example 1. In this first example, we consider the case of the uniform density $f = \frac{1}{4\pi}$. It is easy to verify that $\beta_{j\eta} = \langle f, \psi_{j\eta} \rangle_{L^2} = 0$ for every j and every η. Following Baldi et al. (2009) [1], a simple way of assessing the performance of the procedure is to count the number of coefficients surviving thresholding.

	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$\kappa_0 = 0.08$	2	1	75	337
$\kappa_0 = 0.29$	0	0	2	8
$\kappa_0 = 0.34$	0	0	0	2

TABLE 1

number of coefficients surviving thresholding for various values of κ_0, $\phi \sim U[0, \pi/8]$.

	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$\kappa_0 = 0.05$	4	35	135	556
$\kappa_0 = 0.17$	0	4	24	103
$\kappa_0 = 0.30$	0	0	1	5
$\kappa_0 = 0.34$	0	0	0	3

TABLE 2

number of coefficients surviving thresholding for various values of κ_0, $\phi \sim U[0, \pi]$.

We precise that both in the cases of an angle following a law $U[0, \pi/8]$ or $U[0, \pi]$, all the coefficients are killed for $\kappa_0 = 0.38$. Accordingly, we can conclude that the thresholding procedure based on spherical needlets is very efficient.

Example 2. We will now deal with the example of a density of the form $f(\omega) = ce^{-4|\omega - \omega_1|^2}$, with $\omega_1 = (0, 1, 0)$ and $c = 1/0.7854$. With this choice of density, it turns out that $\|f\|_\infty = 1/0.7854 = 1.2732$. Hence we can set M such that $M = 1.2732$. The graph of f in the spherical coordinates (Φ, Θ) ($\Phi = \text{longitude}, 0 \leq \Phi \leq 2\pi$, $\Theta = \text{colatitude}, 0 \leq \Theta \leq \pi$) is given in Fig 2. We also plot the noisy observations for different cases of perturbations. We remark that for rather big rotation angles such that $\phi \sim U[0, \pi/4]$ or $\phi \sim U[0, \pi/2]$, the observations tend to be spread over a large region on the sphere and not being concentrated in a specific region any more. Consequently, denoising might prove to be difficult. In the context of the deconvolution on the sphere, a large amount of noise corresponds to a rotation around the Oz axis with a large angle.
A fundamental issue in the field of astrophysics is to detect the place of the peak of the bell which in the present density case is localized in ($\Theta = \pi/2$, $\Phi = \pi/2$). For each case of noise, we plot the observations both on the sphere and on the flattened sphere and give the reconstructed density in the spherical coordinates.
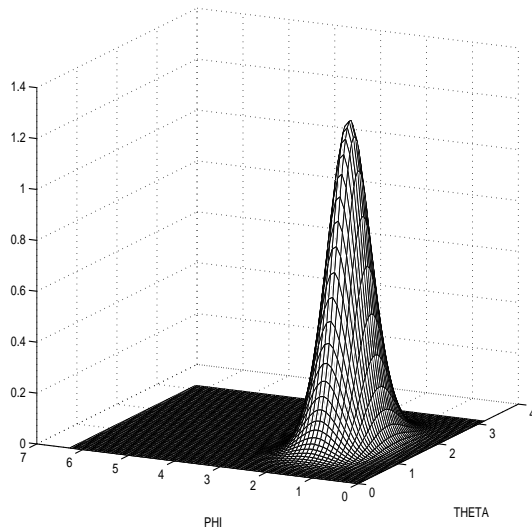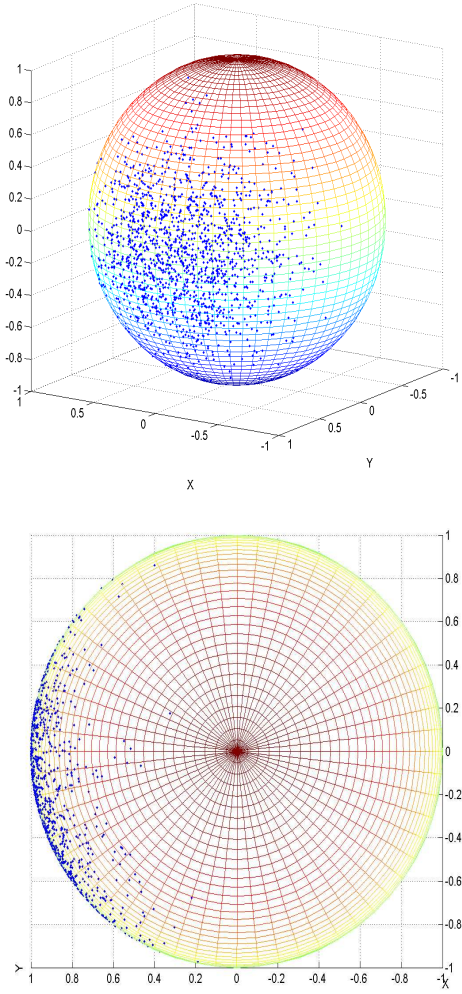

FIG 2. *The target density*
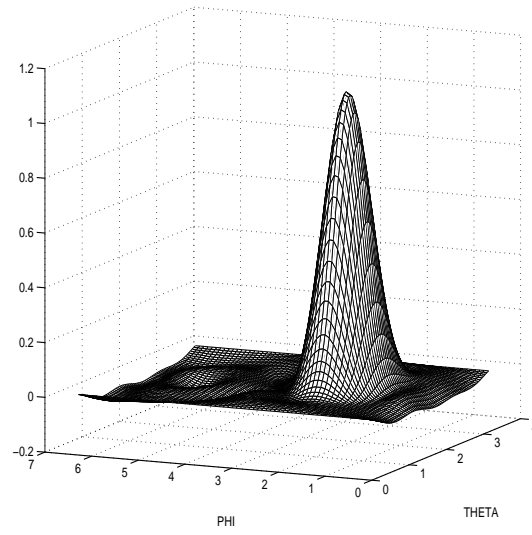
FIG 3. *Observations* $\phi \sim U[0, \pi/8]$.

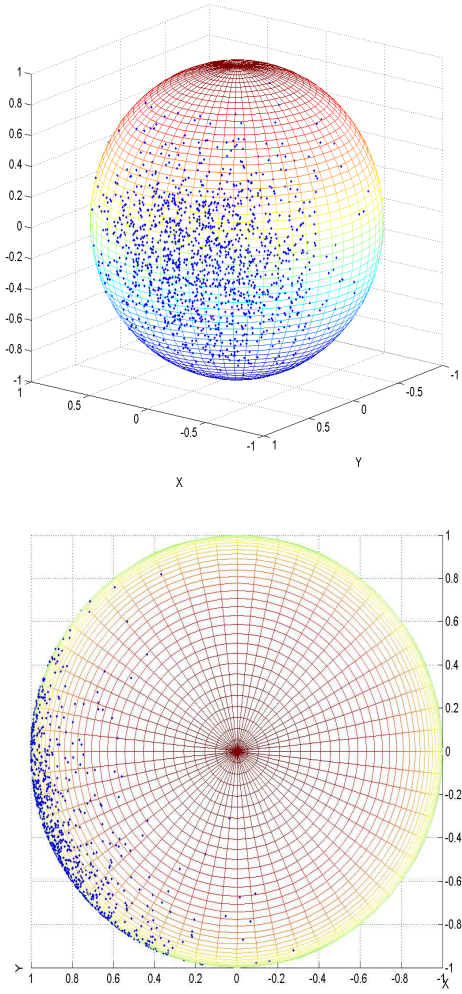FIG 4. *The estimated density, bandwith* $\kappa_0 = 0.43$.

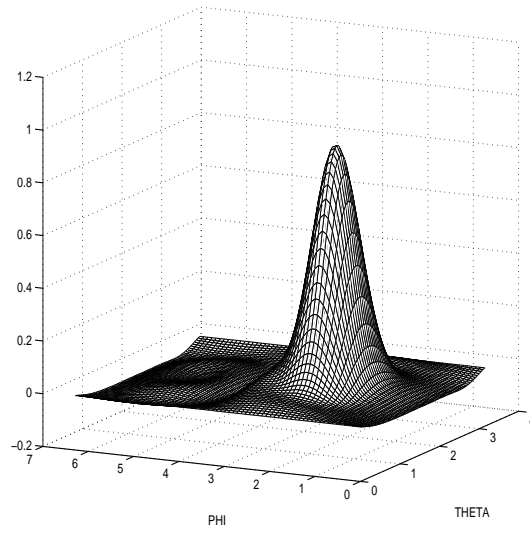FIG 5. *Observations $\phi \sim U[0, \pi/4]$.*

FIG 6. *The estimated density, bandwith $\kappa_0 = 0.46$.*
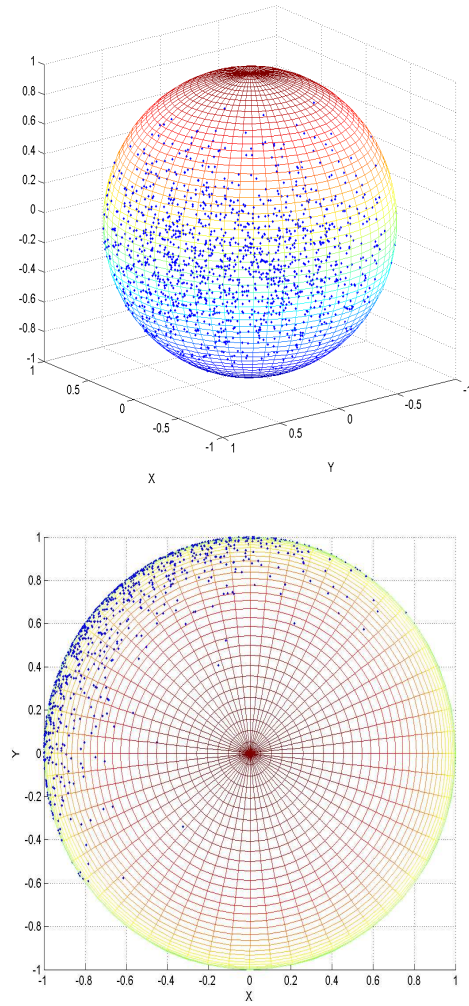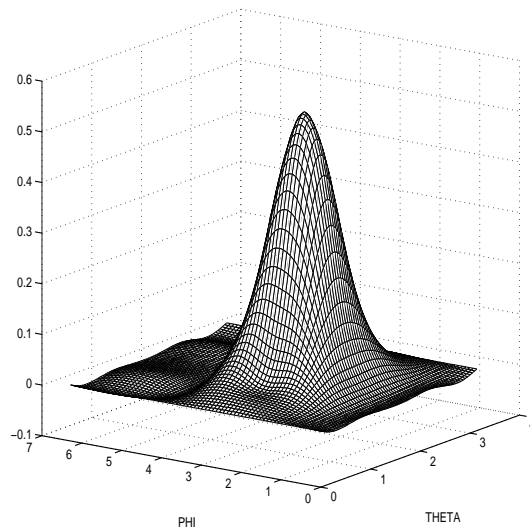
FIG 7. *Observations* $\phi \sim U[0, \pi/2]$.

FIG 8. *The estimated density, bandwith $\kappa_0 = 0.56$.*

At a closer inspection, we notice that the position of the peak of the estimated bell is pretty well localized whatever the amount of noise. Nevertheless in the case of the law $U[0, \pi/2]$, the longitude coordinate of the peak tends to slightly move, its colatitude coordinate remaining well localized. Therefore even if in the case of rather big rotations such that $\phi \sim U[0, \pi/4]$ and $\phi \sim U[0, \pi/2]$, our estimation procedure allows us to detect the position of the peak. Of course, one remarks that the base of the bell tends to become a bit larger when the noise increases, this is due to the fact that the observations are not concentrated in a specific region any longer, but the genuine form of the density is well preserved.

6. Proof of Theorem 2 . In this proof, C will denote an absolute constant which may change from one line to the other.

We begin with the following proposition

PROPOSITION 2. *For any $p \geq 1$,*

$$\tag{41} \mathbb{E}|\hat{\beta}_{j\eta} - \beta_{j\eta}|^p \leq c[\frac{\sigma_j^2}{N}]^{\frac{p}{2}}$$

$$\tag{42} \mathbb{P}[|\hat{\beta}_{j\eta} - \beta_{j\eta}| \geq \sigma_j \kappa t_N] \leq cN^{-\gamma(\kappa)}$$

$$\tag{43} \gamma(\kappa) = -\kappa^2/4$$

To prove (41), we will use Rosenthal inequality:

$$
\begin{aligned}
\hat{\beta}_{j\eta} - \beta_{j\eta} &= \frac{1}{N}\sum_{i=1}^{N} G(Z_i) - \mathbb{E}G(Z_i)\\
G(x) &= \sum_{lm}\psi_{j\eta}^{lm}\sum_n \hat{f}_{\varepsilon^{-1}mn}^l Y_n^l(x).
\end{aligned}
$$

Hence,

$$
VarG(Z) \leq \int_{\mathbb{S}^2}|G(\omega)|^2 f_Z(\omega)d\omega \leq M\int_{\mathbb{S}^2}|G(\omega)|^2 d\omega,
$$

since $\|f_Z\|_\infty \leq M$ (this is a consequence of the fact that $s > 2/\pi$ which in turns implies that f is continuous, and so is f_Z on the sphere). Then using Parseval:

$$
\int_{\mathbb{S}^2}|G(\omega)|^2 d\omega = \sum_{ln}|\sum_m \psi_{j\eta}^{lm}\hat{f}_{\varepsilon^{-1}mn}^l|^2 = \sigma_j^2.
$$

Moreover, using Cauchy Schwarz inequality, and noticing again that $\psi_{j\eta}^{lq} \neq 0$ only for $2^{j-1} \leq l \leq 2^{j+1}$:

$$
|G(\omega)| \leq [\sum_{ln}|\sum_m \psi_{j\eta}^{lm}f_{\varepsilon^{-1}mn}^l|^2]^{1/2}[\sum_{lm}I\{2^{j-1}\leq l \leq 2^{j+1}\}|Y_m^l(\omega)|^2]^{1/2} \leq C\sigma_j 2^{j/2},
$$

since $\sum_m |Y_m^l(\omega)|^2 = \frac{2l+1}{4\pi}$ because of the addition formula (see Kim and Koo (2002) [7]) inequality (6.1)). Then using Rosenthal inequality, we get (we recall that $2^j \leq N^{1/2}$ for $j \leq J$):

$$
\mathbb{E}|\frac{1}{N}\sum_{i=1}^{N} G(Z_i) - \mathbb{E}G(Z_i)|^p \leq C[\frac{1}{N^p}\sigma_j^2(\sigma_j 2^{j/2})^{p-2} + \frac{1}{N^p}(N\sigma_j^2)^{\frac{p}{2}}] \leq C\sigma_j^p N^{-p/2}.
$$

To prove (42), we use Bernstein inequality:

$$
\mathbb{P}[|\hat{\beta}_{j\eta} - \beta_{j\eta}| \geq \sigma_j \kappa t_N] \leq 2\exp-\frac{\kappa^2 \sigma_j^2 \log N}{2(\sigma_j^2 + \frac{1}{3}\kappa\sigma_j t_N\sigma_j 2^{j/2}\pi^{-1/2})} \leq 2N^{-\kappa^2/4}.
$$

This ends up the proof of the proposition.

Now, to get the result of Theorem 2, we begin by the following decomposition :

$$
\begin{aligned}
\mathbb{E}\|\hat{f} - f\|_p^p &\leq 2^{p-1}\{\mathbb{E}\|\sum_{j=-1}^{J}\sum_{\eta\in\mathscr{Z}_j}(t(\hat{\beta_{j\eta}}) - \beta_{j\eta})\psi_{j\eta}\|_p^p + \|\sum_{j>J}\sum_{\eta\in\mathscr{Z}_j}\beta_{j\eta}\psi_{j\eta}\|_p^p\}\\
&=: \ I + II
\end{aligned}
$$

The term II is easy to analyse, as follows. We observe first that since

$B_{\pi,r}^s(M) \subset B_{p,r}^s(M')$ for $\pi \geq p$, this case will be assimilated to the case $\pi = p$ and from now on, we will only consider $\pi \leq p$. Since f belongs to $B_{\pi,r}^s(M)$, using the embedding results recalled above in (31), we have that f also belongs to $B_{p,r}^{s-(\frac{2}{\pi} - \frac{2}{p})}(M')$, for some constant M' and for $\pi \leq p$. Hence

$$\| \sum_{j>J} \sum_{\eta \in \mathscr{Z}_j} \beta_{j\eta} \psi_{j\eta} \|_p \leq C 2^{-J[s-2(\frac{1}{\pi} - \frac{1}{p})]}.$$

Then we only need to verify that $\frac{s-2(\frac{1}{\pi} - \frac{1}{p})}{\nu+1}$ is always larger that μ, which is not difficult.

Indeed, on the first zone $s \geq (\nu+1)(p/\pi - 1)$. So, $s + \nu + 1 \geq (\nu+1)\frac{p}{\pi}$ which entails that $\frac{s}{(s+\nu+1)} \leq \frac{s}{(\nu+1)\frac{p}{\pi}}$. We need to check that $s - 2(\frac{1}{\pi} - \frac{1}{p}) \geq \frac{s\pi}{p}$. We have that $s - \frac{2}{\pi} + \frac{2}{p} - \frac{s\pi}{p} = 2(\frac{s\pi}{2} - 1)(\frac{1}{\pi} - \frac{1}{p}) \geq 0$ since $s \geq \frac{2}{\pi}$ and $p \geq \pi$.

On the second zone, we obviously have that $\frac{s-2(\frac{1}{\pi} - \frac{1}{p})}{\nu+1}$ is always larger that $\mu = \frac{s-2/\pi+2/p}{s+\nu-2/\pi+1}$.

Bounding the term I is more involved. Using the triangular inequality together with Hölder inequality, and property (28) for the second line, we get

$$I \leq 2^{p-1} J^{p-1} \sum_{j=-1}^{J} \mathbb{E} \| \sum_{\eta \in \mathscr{Z}_j} (t(\hat{\beta}_{j\eta}) - \beta_{j\eta}) \psi_{j\eta} \|_p^p$$

$$\leq 2^{p-1} J^{p-1} C \sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} \mathbb{E} |t(\hat{\beta}_{j\eta}) - \beta_{j\eta}|^p \|\psi_{j\eta}\|_p^p.$$

Now, we separate four cases :

$$\sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} \mathbb{E} |t(\hat{\beta}_{j\eta}) - \beta_{j\eta}|^p \|\psi_{j\eta}\|_p^p = \sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} \mathbb{E} |t(\hat{\beta}_{j\eta}) - \beta_{j\eta}|^p \|\psi_{j\eta}\|_p^p \Big\{ I\{|\hat{\beta}_{j\eta}| \geq \kappa t_N \sigma_j\}$$

$$+ I\{|\hat{\beta}_{j\eta}| < \kappa t_N \sigma_j\} \Big\}$$

$$\leq \sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} \Big[\mathbb{E} |\hat{\beta}_{j\eta} - \beta_{j\eta}|^p \|\psi_{j\eta}\|_p^p I\{|\hat{\beta}_{j\eta}| \geq \kappa t_N \sigma_j\}$$

$$\Big\{ I\{|\beta_{j\eta}| \geq \frac{\kappa}{2} t_N \sigma_j\} + I\{|\beta_{j\eta}| < \frac{\kappa}{2} t_N \sigma_j\} \Big\}$$

$$+ |\beta_{j\eta}|^p \|\psi_{j\eta}\|_p^p I\{|\hat{\beta}_{j\eta}| < \kappa t_N \sigma_j\} \Big\{ I\{|\beta_{j\eta}| \geq 2\kappa t_N \sigma_j\}$$

$$+ I\{|\beta_{j\eta}| < 2\kappa t_N \sigma_j\} \Big\} \Big]$$

$$= : Bb + Bs + Sb + Ss.$$

Using Proposition 2, we have

$$
\begin{aligned}
Bb &\le C\sum_{j=-1}^{J}\sum_{\eta\in\mathscr{Z}_j}\sigma_j^p N^{-p/2}\|\psi_{j\eta}\|_p^p I\{|\beta_{j\eta}|\ge\tfrac{\kappa}{2}t_N\sigma_j\}\\
Ss &\le \sum_{j=-1}^{J}\sum_{\eta\in\mathscr{Z}_j}|\beta_{j\eta}|^p\|\psi_{j\eta}\|_p^p I\{|\beta_{j\eta}|<2\kappa t_N\sigma_j\}.
\end{aligned}
$$

And, using in addition (25) and (36)

$$
\begin{aligned}
Bs &\le \sum_{j=-1}^{J}\sum_{\eta\in\mathscr{Z}_j}[\mathbb{E}|\hat{\beta_{j\eta}}-\beta_{j\eta}|^{2p}]^{1/2}[\mathbb{P}\{|\hat{\beta_{j\eta}}-\beta_{j\eta}|\ge\tfrac{\kappa}{2}t_N\sigma_j\}]^{1/2}\|\psi_{j\eta}\|_p^p I\{|\beta_{j\eta}|<\tfrac{\kappa}{2}t_N\sigma_j\}\\
&\le C\sum_{j=-1}^{J}\sum_{\eta\in\mathscr{Z}_j}\sigma_j^p N^{-p/2}2^{1/2}N^{-\kappa^2/32}\|\psi_{j\eta}\|_p^p I\{|\beta_{j\eta}|<\tfrac{\kappa}{2}t_N\sigma_j\}\\
&\le C\sum_{j=-1}^{J}2^{jp(\nu+1)}N^{-p/2}N^{-\kappa^2/32}\le CN^{-\kappa^2/32}.
\end{aligned}
$$

Now, since f belongs to $B_{\pi,r}^s(M)$, using again Proposition 2 and (31), we have

$$
\begin{aligned}
Sb &\le \sum_{j=-1}^{J}\sum_{\eta\in\mathscr{Z}_j}|\beta_{j\eta}|^p\|\psi_{j\eta}\|_p^p\mathbb{P}\{|\hat{\beta_{j\eta}}-\beta_{j\eta}|\ge 2\kappa t_N\sigma_j\}I\{|\beta_{j\eta}|\ge 2\kappa t_N\sigma_j\}\\
&\le \sum_{j=-1}^{J}\sum_{\eta\in\mathscr{Z}_j}|\beta_{j\eta}|^p\|\psi_{j\eta}\|_p^p 2N^{-\kappa^2}I\{|\beta_{j\eta}|\ge 2\kappa t_N\sigma_j\}\\
&\le C\sum_{j=-1}^{J}2^{-jp(s-2(\frac{1}{\pi}-\frac{1}{p})_+)}N^{-\kappa^2}\le CN^{-\kappa^2}2^{2J}\le CN^{-\kappa^2+\frac{1}{(\nu+1)}}.
\end{aligned}
$$

It is easy to check that in any cases if $\kappa\ge 1-p$ the terms Bs and Sb are smaller than the rates announced in the theorem.

We have using Proposition 2, (25) and condition (36) for any $p\ge z\ge 0$:

$$
\begin{aligned}
Bb &\le CN^{-p/2}\sum_{j=-1}^{J}\sigma_j^p 2^{j(p-2)}\sum_{\eta\in\mathscr{Z}_j}I\{|\beta_{j\eta}|\ge\tfrac{\kappa}{2}t_N\sigma_j\}\\
&\le CN^{-p/2}\sum_{j=-1}^{J}\sigma_j^p 2^{j(p-2)}\sum_{\eta\in\mathscr{Z}_j}|\beta_{j\eta}|^z[t_N\sigma_j]^{-z}\\
&\le Ct_N^{p-z}\sum_{j=-1}^{J}2^{j[\nu(p-z)+p-2]}\sum_{\eta\in\mathscr{Z}_j}|\beta_{j\eta}|^z.
\end{aligned}
$$

Also, for any $p \geq z \geq 0$

$$Ss \leq C \sum_{j=-1}^{J} 2^{j(p-2)} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^z \sigma_j^{p-z} [t_N]^{p-z}$$

$$\leq C[t_N]^{p-z} \sum_{j=-1}^{J} 2^{j(\nu(p-z)+p-2)} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^z.$$

So in both cases we have the same bound to investigate. We will write this bound on the following form (forgetting the constant) :

$$A+B := t_N{}^{p-z_1} [\sum_{j=-1}^{j_0} 2^{j[\nu(p-z_1)+p-2]} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^{z_1}] + t_N{}^{p-z_2} [\sum_{j=j_0+1}^{J} 2^{j[\nu(p-z_2)+p-2]} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^{z_2}].$$

The constants z_i and j_0 will be chosen depending on the cases, with the only constraint $p \geq z_i \geq 0$.

We recall that we only need to investigate the case $p \geq \pi$, since when $p \leq \pi$, $B_{\pi r}^s(M) \subset B_{pr}^s(M')$.

Let us first consider the case where $s \geq (\nu + 1)(\frac{p}{\pi} - 1)$, put

$$q = \frac{p(\nu + 1)}{s + \nu + 1},$$

and observe that on the considered domain, $q \leq \pi$ and $p > q$. In the sequel it will be useful to observe that we have $s = (\nu + 1)(\frac{p}{q} - 1)$. Now, taking $z_2 = \pi$, we get:

$$B \leq t_N{}^{p-\pi} [\sum_{j=j_0+1}^{J} 2^{j[\nu(p-\pi)+p-2]} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^{\pi}].$$

Now, as

$$\frac{p}{q} - \frac{2}{\pi} + \nu(\frac{p}{q} - 1) = s + 1 - \frac{2}{\pi}$$

and

$$\sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^{\pi} = 2^{-j\pi(s+1-\frac{2}{\pi})} \tau_j^{\pi},$$

with $(\tau_j)_j \in l_r$ (this last thing is a consequence of the fact that $f \in B_{\pi,r}^s(M)$), we can write :

$$B \leq t_N{}^{p-\pi} \sum_{j=j_0+1}^{J} 2^{jp(1-\frac{\pi}{q})(\nu+1)} \tau_j^{\pi}$$

$$\leq C t_N{}^{p-\pi} 2^{j_0 p(1-\frac{\pi}{q})(\nu+1)}.$$

The last inequality is true for any $r \geq 1$ if $\pi > q$ and for $r \leq \pi$ if $\pi = q$. Notice that $\pi = q$ is equivalent to $s = (\nu + 1)(\frac{p}{\pi} - 1)$. Now if we choose j_0 such that $2^{j_0 \frac{p}{q}(\nu+1)} \sim t_N{}^{-1}$ we get the bound

$$t_N{}^{p-q},$$

which exactly gives the rate announced in the theorem for this case. As for the first part of the sum (before j_0), we have, taking now $z_1 = \tilde{q}$, with $\tilde{q} \leq \pi$ (and also $\tilde{q} \leq p$ since we investigate the case $p \geq \pi$), so that $[\frac{1}{2^{2j}} \sum_{\eta \in \mathscr{L}_j} |\beta_{j\eta}|^{\tilde{q}}]^{\frac{1}{\tilde{q}}} \leq [\frac{1}{2^{2j}} \sum_{\eta \in \mathscr{L}_j} |\beta_{j\eta}|^{\pi}]^{\frac{1}{\pi}}$, we get

$$
\begin{aligned}
A \;\leq\; & t_N{}^{p-\tilde{q}} [\sum_{-1}^{j_0} 2^{j[\nu(p-\tilde{q})+p-2]} \sum_{\eta \in \mathscr{L}_j} |\beta_{j\eta}|^{\tilde{q}}] \\
\leq\; & t_N{}^{p-\tilde{q}} [\sum_{-1}^{j_0} 2^{j[\nu(p-\tilde{q})+p-\frac{2\tilde{q}}{\pi}]} [\sum_{\eta \in \mathscr{L}_j} |\beta_{j\eta}|^{\pi}]^{\frac{\tilde{q}}{\pi}}] \\
\leq\; & t_N{}^{p-\tilde{q}} \sum_{-1}^{j_0} 2^{j[(\nu+1)p(1-\frac{\tilde{q}}{q})]} \tau_j^{\tilde{q}} \\
\leq\; & C t_N{}^{p-\tilde{q}} 2^{j_0 p[(\nu+1)(1-\frac{\tilde{q}}{q})]} \\
\leq\; & C t_N{}^{p-q}.
\end{aligned}
$$

The last two lines are valid if $\tilde{q}$ is chosen strictly smaller than q (this is possible since $\pi \geq q$).

Let us now consider the case where $s < (\nu + 1)(\frac{p}{\pi} - 1)$, and choose now

$$q = p \frac{\nu + 1 - \frac{2}{p}}{s + \nu - \frac{2}{\pi} + 1}.$$

In such a way that we easily verify that $p - q = p \frac{s - 2/\pi + 2/p}{1 + \nu + s - 2/\pi}$, $q - \pi = \frac{(p-\pi)(1+\nu) - \pi s}{s + \nu - \frac{2}{\pi} + 1} > 0$. Furthermore we also have $s + 1 - \frac{2}{\pi} = \frac{p}{q} - \frac{2}{q} + \nu(\frac{p}{q} - 1)$. Hence taking $z_1 = \pi$ and using again the fact that f belongs to $B_{\pi,r}^s(M)$,

$$
\begin{aligned}
A \leq\; & t_N{}^{p-\pi} [\sum_{-1}^{j_0} 2^{j[\nu(p-\pi)+p-2]} \sum_{\eta \in \mathscr{L}_j} |\beta_{j\eta}|^{\pi}] \\
\leq\; & t_N{}^{p-\pi} \sum_{-1}^{j_0} 2^{j[(\nu+1-\frac{2}{p})\frac{p}{q}(q-\pi)]} \tau_j^{\pi} \\
\leq\; & C t_N{}^{p-\pi} 2^{j_0[(\nu+1-\frac{2}{p})\frac{p}{q}(q-\pi)]}.
\end{aligned}
$$

This is true since $\nu + 1 - \frac{2}{p}$ is also strictly positive since $\nu + 1 > \frac{s}{\frac{p}{\pi}-1} \geq \frac{2}{p-\pi} \geq \frac{2}{p}$. If we now take $2^{j_0 \frac{p}{q}(\nu+1-\frac{2}{p})} \sim t_N^{-1}$ we get the bound

$$t_N{}^{p-q},$$

which is the rate announced in the theorem for this case.

Again, for B, we have, taking now $z_2 = \tilde{q} > q(> \pi)$

$$B \leq t_N{}^{p-\tilde{q}}[\sum_{j=j_0+1}^{J} 2^{j[\nu(p-\tilde{q})+p-2]} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^{\tilde{q}}].$$

But

$$\sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^{\tilde{q}} \leq C 2^{-j\tilde{q}(s+1-\frac{2}{\pi})} \tau_j^{\tilde{q}},$$

and $s + 1 - \frac{2}{\pi} = \frac{p}{q} - \frac{2}{q} + \nu(\frac{p}{q} - 1)$, hence

$$\begin{aligned} B &\leq C t_N{}^{p-\tilde{q}} \sum_{j=j_0+1} 2^{j[(\nu+1-\frac{2}{p})\frac{p}{q}(q-\tilde{q})]} \tau_j^{\tilde{q}} \\ &\leq C t_N{}^{p-\tilde{q}} 2^{j_0[(\nu+1-\frac{2}{p})\frac{p}{q}(q-\tilde{q})]} \\ &\leq C t_N{}^{p-q}, \end{aligned}$$

which completes the proof of Theorem 2.

REFERENCES

[1] Paolo Baldi, Gérard Kerkyacharian, Domenico Marinucci, and Dominique Picard. Adaptive density estimation for directional data using needlets. *Ann. Statist. to appear*, 2009.

[2] Paolo Baldi, Gérard Kerkyacharian, Domenico Marinucci, and Dominique Picard. Subsampling needlet coefficients on the sphere. *Bernoulli*, 15(2):438–463, 2009.

[3] Dennis M. Healy, Jr., Harrie Hendriks, and Peter T. Kim. Spherical deconvolution. *J. Multivariate Anal.*, 67(1):1–22, 1998.

[4] I. Johnstone, G. Kerkyacharian, D. Picard, and M. Raimondo. Wavelet deconvolution in a periodic setting. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 66(3):547–573, 2004.

[5] G. Kerkyacharian, G. Kyriazis, E. Le Pennec, P. Petrushev, and D. Picard. Inversion of noisy radon transform by svd based needlets. *ACHA*, 2009.

[6] Gérard Kerkyacharian, Pencho Petrushev, Dominique Picard, and Thomas Willer. Needlet algorithms for estimation in inverse problems. *Electron. J. Stat.*, 1:30–76 (electronic), 2007.

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

[8] F. Narcowich, P. Petrushev, and J. Ward. Decomposition of Besov and Triebel-Lizorkin spaces on the sphere. *J. Funct. Anal.*, 238(2):530–564, 2006.

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

[10] D. Pietrobon, Paolo Baldi, Gérard Kerkyacharian, Domenico Marinucci, and Dominique Picard. Spherical needlets for cmb data analysis. 383:539–545, 2008.

[11] Jeffrey S. Rosenthal. Random rotations: characters and random walks on $SO(N)$. *Ann. Probab.*, 22(1):398–423, 1994.

[12] James D. Talman. *Special functions: A group theoretic approach.* Based on lectures by Eugene P. Wigner. With an introduction by Eugene P. Wigner. W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[13] Audrey Terras. *Harmonic analysis on symmetric spaces and applications. I.* Springer-Verlag, New York, 1985.

[14] Arnoud C. M. van Rooij and Frits H. Ruymgaart. Regularized deconvolution on the circle and the sphere. In *Nonparametric functional estimation and related topics (Spetses, 1990)*, volume 335 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 679–690. Kluwer Acad. Publ., Dordrecht, 1991.

[15] N. Ja. Vilenkin. *Fonctions spéciales et théorie de la représentation des groupes.* Traduit du Russe par D. Hérault. Monographies Universitaires de Mathématiques, No. 33. Dunod, Paris, 1969.

Gérard Kerkyacharian
CNRS, LPMA
175 rue du Chevaleret,
75013 Paris, France
E-mail: kerk@math.jussieu.fr

Thanh Mai Pham Ngoc
Laboratoire de Probabilités et Modèles aléatoires
Université Paris 6, case 188, 4, Pl. Jussieu, 75252 Paris France.
E-mail: thanh.pham_ngoc@upmc.fr

Dominique Picard
Université Paris VII,
175 rue du Chevaleret,
75013 Paris, France
E-mail: picard@math.jussieu.fr