LOCALIZED SPHERICAL DECONVOLUTION

BY GÉRARD KERKYACHARIAN, THANH MAI PHAM NGOC
AND DOMINIQUE PICARD

CNRS and LPMA, Université Paris Sud and Université Paris VII

We provide a new algorithm for the treatment of the deconvolution problem 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 behavior 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 [19] and subsequently solved in Healy, Hendriks and Kim [5]. Kim and Koo [11] 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 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 is very rarely 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. [10] for the Wicksell case, Kerkyacharian et al. [7] for the Radon transform, which utilizes the construction borrowed from Narcowich, Petrushev and Wald [12, 13] of a tight frame (i.e., a redundant family) staying sufficiently close to the SVD decomposition but which enjoys at the same time enough localization properties to be successfully used for statistical

estimation (see, for instance, Baldi et al. [1, 2], Pietrobon et al. [14] for other types of applications). The construction [13] produces a family of functions which very much resemble 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 deconvolution, due to the closeness to the SVD basis) with a subsequent fine tuning 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 as well as $\mathbb{L}_\infty$. We find back these results in the simulation study where the effect of the localization are highlighted for instance by a comparison of the performances, on a bell-density example between the procedure provided here and the SVD methods (detailed below) proving that our quality of reconstruction of the peak is notably better.

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 Hölder spaces and include at the same time the Sobolev regularity spaces used in Kim and Koo [11].

Sphere deconvolution has a vast domain of application such as medical imaging (see Tournier et al. [18]) and astrophysics. Indeed, our results are especially motivated by many recent developments in the area of observational astrophysics.

It is a common problem in astrophysics to analyze 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) illustrates well the type of applications covered by 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, taking 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 perturbed by a secondary noise (ε_i) which leads to the deconvolution problem described below. Following Healy, Hendriks and Kim [5], Kim and Koo [11], the spherical deconvolution problem can be described as follows. Consider the situation where we observe $Z_1, \ldots, Z_N$, N i.i.d. observations with

$$(1) \qquad\qquad Z_i = \varepsilon_i X_i,$$

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 distributions of X and Z are absolutely continuous with respect to the uniform probability measure on $\mathbb{S}^2$, and that the distribution of ε is absolutely continuous with respect to the Haar measure of $SO(3)$. We will denote the densities of resp. Z, X and ε resp. f_Z, f_X, f_ε.

Then

$$(2) \qquad\qquad f_Z = f_\varepsilon * f_X,$$

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 recalls the necessary definitions. It is largely inspired by Kim and Koo [11] and Healy, Hendriks and Kim [5].

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 an expanded form in Vilenkin [20], Talman [16], Terras [17], Kim and Koo [11], and Healy, Hendriks and Kim [5]. Let

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

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 about the z-axis first by an angle ψ, followed by a rotation about the y-axis by an angle θ, and finally by another rotation again about the z-axis by an 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, Hendriks and Kim [5] 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,

$$(4) \qquad D_{mn}^l(\phi, \theta, \psi) = e^{-i(m\phi+n\psi)} P_{mn}^l(\cos\theta),$$

where the associated Legendre functions P_{mn}^l for $-l \leq m, n \leq l, l = 0, 1, \ldots,$ are fully described in Vilenkin [20]. The functions D_{mn}^l for $-l \leq m, n \leq l, l = 0, 1, \ldots,$ are the eigenfunctions of the Laplace Beltrami operator on $SO(3)$, hence, $\sqrt{2l+1}D_{mn}^l, -l \leq m, n \leq l, l = 0, 1, \ldots$ 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

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

where for $-l \leq m, n \leq l, l = 0, 1, \ldots,$ and $g \in SO(3)$, they constitute the collection of inequivalent irreducible representations of $SO(3)$ (for further details, see Vilenkin [20]).

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

$$(6) \qquad \hat{f}_{mn}^l = \int_{SO(3)} f(g) D_{mn}^l(g) \, dg,$$

where dg is the probability Haar measure on $SO(3)$ and we define the following matrix of dimension $(2l+1) \times (2l+1)$

$$\hat{f}^l = [\hat{f}_{mn}^l]_{-l \leq m, n \leq l}, \qquad l = 0, 1, \ldots.$$

The rotational inversion can be obtained by

$$
\begin{aligned}
f(g) &= \sum_l \sum_{-l \leq m, n \leq l} \hat{f}_{mn}^l \overline{D_{mn}^l(g)} \\
&= \sum_l \sum_{-l \leq m, n \leq l} \hat{f}_{mn}^l D_{mn}^l(g^{-1}),
\end{aligned}
$$

(7)

(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

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

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 Legendre functions.

The functions Y_m^l obey

$$(9) \qquad Y_{-m}^l(\theta, \phi) = (-1)^m \overline{Y_m^l(\theta, \phi)}.$$

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

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

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

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

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 $-l \leq m \leq l, l = 0, 1, \ldots,$

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

2.1. *The SVD method.* The singular value method (see Healy, Hendriks and Kim [5] and Kim and Koo [11]) 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:

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

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

3. Needlet construction. This construction is due to Narcowich et al. [12]. 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. [10], Baldi et al. [1, 2], Pietrobon et al. [14]). As described above, we have the following decomposition:

$$
\mathbb{L}_2(\mathbb{S}^2) = \bigoplus_{l=0}^{\infty} \mathbb{H}_l,
\tag{15}
$$

where $\mathbb{H}_l$ is the space spanned by $\{Y_m^l, -l \leq m \leq l\}$ 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 kernel operator:

$$
\forall f \in \mathbb{L}_2(\mathbb{S}^2) \qquad P_{\mathbb{H}_l} f(x) = \int_{\mathbb{S}^2} L_l(\langle x, y\rangle) f(y)\, dy,
\tag{16}
$$

where,

$$
L_l(x, y) = \sum_{m=-l}^{l} Y_m^l(x)\overline{Y_m^l(y)} = L_l(\langle x, y\rangle),
$$

and where $\langle x, y\rangle$ is the standard scalar product of $\mathbb{R}^3$, and L_l is the Legendre polynomial of degree l, defined on $[-1, +1]$ and verifying

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

where $\delta_{l,k}$ is the Kronecker symbol.

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

$$
\int_{\mathbb{S}^2} L_l(\langle x, y\rangle) L_k(\langle y, z\rangle)\, dy = \delta_{l,k} L_l(\langle x, z\rangle).
\tag{18}
$$

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 $|\xi| \leq \frac{1}{2}$.

$$
b^2(\xi) = \phi\left(\frac{\xi}{2}\right) - \phi(\xi) \geq 0,
$$

so that

$$
\forall |\xi| \geq 1 \qquad \sum_{j \geq 0} b^2\left(\frac{\xi}{2^j}\right) = 1.
\tag{19}
$$

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

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

We obviously have

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

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

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

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

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

the space spanned by the spherical harmonics of 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

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

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

$$(23) \qquad \frac{1}{c} 2^{2j} \leq \#\mathscr{Z}_j \leq c 2^{2j}, \qquad \frac{1}{c} 2^{2j} \leq \lambda_\eta \leq c 2^{2j}$$

for some $c > 0$. It holds, using (20):

$$f = L_0(f) + \sum_j \sum_{\eta \in \mathscr{Z}_j} \langle f, \psi_{j,\eta} \rangle_{\mathbb{L}_2(\mathbb{S}^2)} \psi_{j,\eta}.$$

The main result of Narcowich et al. [12] 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 $\xi \in \mathbb{S}^2$,

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

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. Their proof can be found in [12, 13], also in [8].

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

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

$$(26) \qquad c 2^j \leq \|\psi_{j\eta}\|_\infty \leq C 2^j,$$

$$(27) \qquad \left\| \sum_{\eta \in \mathscr{Z}_j} \lambda_\eta \psi_{j,\eta} \right\|_\pi \leq c 2^{2j(1/2-1/\pi)} \left(\sum_{\eta \in \mathscr{Z}_j} |\lambda_\eta|^\pi \right)^{1/\pi},$$

$$(28) \qquad \left\| \sum_{\eta \in \mathscr{Z}_j} \lambda_\eta \psi_{j\eta} \right\|_p^p \leq D_p \sum_{\eta \in \mathscr{Z}_j} |\lambda_\eta|^p \|\psi_{j\eta}\|_p^p,$$

$$(29) \qquad \left\| \sum_{\eta \in \mathscr{Z}_j} u_\eta \psi_{j\eta} \right\|_\infty \leq C \sup_{\eta \in \mathscr{Z}_j} |u_\eta| 2^j.$$

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 $\eta = 250$.

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 [12].

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 $\mathscr{P}_k$ of spherical harmonics of degree less than k. The Besov space $B_{\pi,r}^s$ is defined as the space of functions such that

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

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

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

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 needlet expansion. We provide in Figure 1 a graphical representation of a needlet which shows its localization around the associated cubature point.

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

$$\langle f, \psi_{j,\eta} \rangle = \int_{\mathbb{S}^2} f(x) \psi_{j,\eta}(x) \, dx \overset{def.}{=} \beta_{j,\eta},$$

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

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

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

As has been shown above,

$$c 2^{2j(1/2 - 1/\pi)} \le \| \psi_{j,\eta} \|_\pi \le C 2^{2j(1/2 - 1/\pi)}$$
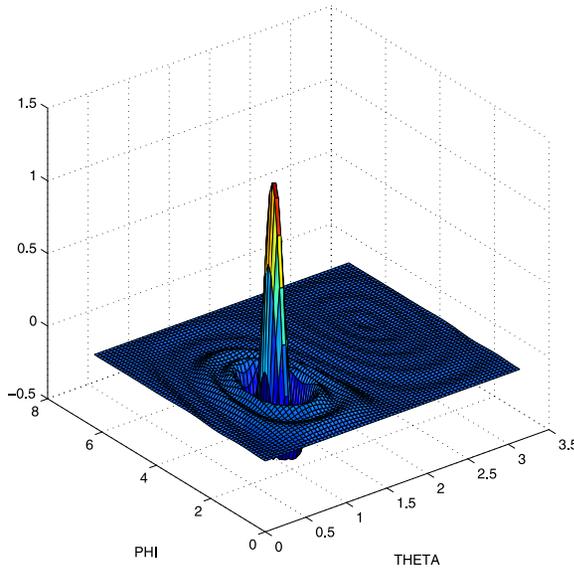
FIG. 1. *A spherical needlet.*

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

$$(30) \qquad \|f\|_{B_{\pi,r}^s} := \left\| \left(2^{j[s+2(1/2-1/\pi)]} \|(\beta_{j\eta})_{\eta \in \mathscr{Z}_j}\|_{\ell_\pi} \right)_{j \geq 0} \right\|_{\ell_r} < \infty,$$

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

$$(31) \qquad \begin{aligned} B_{\pi,r}^s &\subset B_{p,r}^s \qquad \text{for } p \leq \pi, \\ B_{\pi,r}^s &\subset B_{p,r}^{s-2(1/\pi-1/p)} \qquad \text{for } \pi \leq p \text{ and } s > 2\left(\frac{1}{\pi} - \frac{1}{p}\right). \end{aligned}$$

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

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

4. Needlet algorithm: Thresholding the 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). It is worthwhile pointing out that the SVD-estimate of the Fourier coefficient $\hat{f}_m^{l,N}$ appears in the expression of the estimate $\hat{\beta}_{j\eta}$. This underlines that the Needlet dictionary does not depart too much from the Fourier basis and hence benefits from the inversion property while being very well localized.

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:

$$\hat{f} = \sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} t(\hat{\beta}_{j\eta}) \psi_{j\eta}, \tag{33}$$

where t is a thresholding operator defined by

$$t(\hat{\beta}_{j\eta}) = \hat{\beta}_{j\eta} I\{|\hat{\beta}_{j\eta}| \geq \kappa t_N |\sigma_j|\} \qquad \text{with} \tag{34}$$

$$t_N = \sqrt{\frac{\log N}{N}}, \tag{35}$$

$$\sigma_j^2 = A \sum_{ln} \left| \sum_m \psi_{j\eta}^{lm} \hat{f}_{\varepsilon^{-1}mn}^l \right|^2. \tag{36}$$

Here κ is a tuning parameter of the method which will be properly selected later on. A is chosen such that $\|f_Z\|_\infty \leq A$. The choice of these parameters will be discussed later. Notice that the thresholding depends on the resolution level j through the constant σ_j. As usual in inverse problems, the upper level of details J will be chosen depending of the degree of ill-posedness. It is precisely defined in Theorem 2.

4.1. *Performances of the procedure.*

The following theorem considers the case of a $\mathbb{L}_p$ loss with $1 \leq p < \infty$. The case $p = \infty$ is studied in Theorem 3.

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

$$\sigma_j^2 := A \sum_{ln} \left| \sum_m \psi_{j\eta}^{lm} \hat{f}_{\varepsilon^{-1}mn}^l \right|^2 \leq C 2^{2j\nu} \qquad \forall j \geq 0. \tag{37}$$

A is chosen such that $\|f_Z\|_\infty \leq A$. Let us choose κ such that $\kappa \geq \sqrt{3\pi} A$ and $\sqrt{3\pi} A\kappa > \max\{8p, 2p + 1\}$. Let us also take $2^J = \tau[t_N]^{-1/(\nu+1)}$ with t_N as in

(35) *and τ 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 T such that*

$$(38) \qquad \sup_{f \in B_{\pi,r}^s(M)} \mathbb{E}\|\hat{f} - f\|_p^p \leq T (\log(N))^{p-1} [N^{-1/2}\sqrt{\log(N)}]^{\mu p},$$

where

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

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

The following theorem considers the case of L_∞ norm loss.

THEOREM 3. *For $\nu > 0$, let us assume that there exist two constants C, C' such that*

$$(39) \qquad C'2^{2j\nu} \leq \sigma_j^2 := A \sum_{ln} \left| \sum_m \psi_{j\eta}^{lm} \hat{f}_{\varepsilon^{-1}mn}^l \right|^2 \leq C2^{2j\nu} \qquad \forall j \geq 0.$$

A is chosen such that $\|f_Z\|_\infty \leq A$. Let us choose κ such that $\kappa \geq \sqrt{3\pi}A$ and $\sqrt{3\pi}A\kappa > 16$. Let us also take $2^J = \tau[t_N]^{-1/(\nu+1)}$ with t_N as in (35) and τ is a positive constant. Then if $\pi \geq 1$, $s > 2/\pi$, $r \geq 1$, there exists a constant T such that

$$(40) \qquad \sup_{f \in B_{\pi,r}^s(M)} \mathbb{E}\|\hat{f} - f\|_\infty \leq T \log(N)[t_N]^{\mu'},$$

where

$$\mu' = \frac{s - 2/\pi}{s + \nu - 2/\pi + 1}.$$

The proof of these theorems is given in Section 6.

REMARKS.

1. In Theorem 2, the rates of convergence obtained for larger s are usually referred to as the dense case, whereas the other case is referred to as the sparse case.
2. The parameter ν appearing here is often called degree of ill-posedness of the problem (DIP). It appears here through conditions (37) and (39) which are essential in this problem. In [11], 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 rotational harmonics). In the following subsection, we prove that (37) and (39) are in fact a consequence of (and even equivalent to) the standard "ordinary smooth" condition.

3. The rates of convergence found here are standard in inverse problems. They can be related to rates found in Kim and Koo [11] in the same deconvolution problem, with a $\mathbb{L}_2$ loss and constraints on the spaces comparable to $B_{22}^s(M)$. 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. [6]. These results are proved to be minimax (see Kim and Koo [11]) up to logarithmic factors, for the case $p = 2$ with a $B_{22}^s(M)$ constraint on the object to estimate. With methods comparatively similar to those in Willer [21], it could be proved that our rates are also minimax in the general case (again up to logarithmic factors) if we assume condition (39) (or, in other terms, that the DIP is exactly of the order ν).

4. It is interesting to notice that Theorem 3 proves that the same algorithm is also working for the $\mathbb{L}_\infty$ norm, with a slightly more sophisticated proof. It is often the most useful loss function in practice. The proof requires (39) instead of (37). We do not know if this condition is necessary.

5. It is worthwhile noticing that the procedure is adaptive, meaning that it does not require a priori knowledge on the regularity (or sparsity) of the function. It also adapts to nonhomogeneous smoothness of the function. The logarithmic factor is a standard price to pay for adaptation.

6. The procedure requires the knowledge of the constant A which is corresponding to the $\mathbb{L}_\infty$ norm of the density f_Z. It is obvious (because of the convolution) that $A \leq M$. However, it should be better to obtain a procedure not depending on M either. For that, we advocate that $\| f_Z \|_\infty$ can be replaced by $\| \hat{f}_Z(j_N)) \|_\infty$ in practice where $\hat{f}_Z(j_N))$ is an undersmoothed needlet estimator of the density f_Z close to those introduced in Baldi et al. [1] but with no thresholding and with the level j_N chosen so that $2^{2j_N} \simeq N/(\log N)^2$. Standard arguments similar as in Giné and Nickl [4]) imply that this random quantity exponentially concentrates around $\| f_Z \|_\infty$. We can also adopt a more straightforward strategy as detailed in Section 5.

4.2. *Conditions* (37), (39) *and the smoothness of* f_ε. Following Kim and Koo ([11], condition (3.6)), we can define the smoothness of f_ε spectrally. We place ourselves in the "ordinary smooth" case

$$(41) \qquad \| (\hat{f}_\varepsilon^l)^{-1} \|_{\mathrm{op}} \leq d_0 l^\nu \quad \text{and} \quad \| \hat{f}_\varepsilon^l \|_{\mathrm{op}} \leq d_1 l^{-\nu} \qquad \text{as } l \to \infty$$

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

$$\| \hat{f}_\varepsilon^l \|_{\mathrm{op}} = \sup_{h \neq 0, h \in \mathbb{H}_l} \frac{\| \hat{f}_\varepsilon^l h \|_2}{\| h \|_2},$$

$\mathbb{H}_l$ being the $(2l + 1)$-dimensional vector space spanned by $\{ Y_m^l : -l \leq m \leq l \}$.

The following proposition states that condition (37) [resp. (39)] is verified in the ordinary smooth case by the needlets system.

PROPOSITION 1. *If* $\|(\hat{f}_\varepsilon^l)^{-1}\|_{op} \le d_0 l^\nu$, *then there exists a constant C such that*

$$|\sigma_j|^2 := A \sum_{ln} \left| \sum_m \psi_{j\eta}^{lm} \hat{f}_{\varepsilon^{-1}mn}^l \right|^2 \le C 2^{2j\nu} \qquad \forall j \ge 0.$$

If $\|\hat{f}_\varepsilon^l\|_{op} \le d_0 l^{-\nu}$, *then there exists a constant C' such that*

$$|\sigma_j|^2 := A \sum_{ln} \left| \sum_m \psi_{j\eta}^{lm} \hat{f}_{\varepsilon^{-1}mn}^l \right|^2 \ge C' 2^{2j\nu} \qquad \forall j \ge 0.$$

The proof of this proposition is given in the supplement article (Kerkyacharian et al. [9]).

Notice also that the super smooth case (corresponding to exponential spectral decreasing) is also considered in Kim and Koo [11]. We will not consider this case here, although this could be done, basically because this case corresponds to very poor rates of convergence (logarithmic in N). As well, this case does not require a thresholding since the adaptation is obtained almost for free.

We now give a brief review of some examples of smooth distributions which are discussed in depth in Healy, Hendriks and Kim [5] and Kim and Koo [11].

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

(42) $$\hat{f}_{\varepsilon,mn}^l = \left(1 + \rho^2 l(l+1)\right)^{-1} \delta_{mn}$$

for $-l \le m, n \le l$ and $l = 0, 1, \ldots,$ 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 [15]).

If one considers the situation where f_ε is a p-fold convolution product of conjugate invariant random for a fixed axis, then Rosenthal ([15], page 407) showed that

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

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

5. Practical performances. In this section, we produce the results of numerical experiments on the sphere $\mathbb{S}^2$. Numerical work has been conducted using the spherical pixelization HEALPix software package. HEALPix provides an approximate quadrature of the sphere with a number of data points of order $C 2^{2J}$ and a number of quadrature weights of order $\frac{1}{C 2^{2J}}$, for some positive constant C. This approximation is considered as reliable enough and commonly used in astrophysics.

In the two examples below, we considered samples of cardinality $N = 1500$. The maximal resolution level J is taken such that $J = (1/2)\log_2(\frac{N}{\log N})$. 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:

$$(43) \quad \hat{\beta}_{j\eta} = \frac{1}{N}\sqrt{\lambda_{j\eta}} \sum_{l=2^{j-1}}^{2^{j+1}} b(l/2^j) \sum_{m=-l}^{l} \overline{Y_m^l(\xi_{j\eta})} \sum_{n=-l}^{l} \hat{f}_{\varepsilon^{-1},mn}^l \sum_{u=1}^{N} Y_n^l(Z_u),$$

where the quadrature weight are approximately uniform, $\lambda_{j\eta} \simeq 4\pi/(12.2^{2j})$. We replace the rotational Fourier transform $(\hat{f}_\varepsilon^l)_{mn} := \hat{f}_{\varepsilon,mn}^l$ [defined in (6)] by its empirical version, more tractable in our simulation study. Note also that this situation is very likely to occur for instance in the context of astrophysics.

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

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

where the rotational harmonics $D_{m,n}^l$ 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 Oz as the rotation axis and an angle ϕ following a uniform law with different supports such as $[0, \pi/8]$, $[0, \pi/4]$, $[0, \pi/2]$. The larger the support of distribution is, the more intense the effect of the noise will be.

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

$$D_{m,n}^l(\varepsilon_j) = P_{mn}^l(1)e^{-ni\phi_j} = \delta_{mn}e^{-in\phi_j},$$

where $\phi_j \sim U[0, a]$ and a is a positive constant which will be specified later.

Choosing σ_j. As one may notice, the estimator $\hat{f}$ [see (33)] relies on the knowledge of A which controls the sup norm of the density f_Z and appears in the formula of σ_j^2, see (36). Different ways to circumvent this difficulty can be used, for instance, estimating it as explained above. However, we have adopted in this section a different approach. The quantity σ_j^2 constitutes a control of the variance of the estimated coefficients $\hat{\beta}_{j\eta}$ as it is shown by inequality (5) in the supplementary material (see Kerkyacharian et al. [9]). Using this remark, instead of using an

upper bound of the variance of $\hat{\beta}_{j\eta}$, we will directly plug in an estimation of this variance. Hence, σ_j^2 in (34) is replaced by

$$v_{j\eta}^2 = \frac{1}{(N-1)} \sum_{i=1}^{N} |G_{j\eta}(Z_i) - \widetilde{G_{j\eta}}(Z)|^2,$$

where

$$G_{j\eta}(x) = \sum_{lm} \psi_{j\eta}^{lm} \sum_n \hat{f}_{\varepsilon^{-1},mn}^l Y_n^l(x)$$

and

$$\widetilde{G_{j\eta}}(Z) = \frac{1}{N} \sum_{i=1}^{N} G_{j\eta}(Z_i).$$

Remark that $\hat{\beta}_{j\eta} = \frac{1}{N} \sum_{i=1}^{N} G_{j\eta}(Z_i)$.

Hence, for the reconstruction of the density f, we have the following Needlet estimator:

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

Choosing the tuning parameter κ. Before entering into the details of the numerical results, we will dwell on the methodology used in this part to calibrate the tuning paramater κ in practice. In a first time, we focus on the uniform density. Indeed, we make an upstream study with the uniform density $f = \frac{1}{4\pi} \mathbf{1}_{\mathbb{S}^2}$ which will allow us to determine a reasonable value for κ which will be kept in the sequel for other types of densities. On the one hand, the choice of the uniform density is not fortuitous because we know that in theory the needlet coefficient $\langle f, \psi_{j\eta} \rangle = 0$ for all j and η. On the other hand, for an upstream study, a prerequisite is to deal with a simple density, which is the case. Hence, we test various values of κ and see how many coefficients survive thresholding. Then we look at the smallest value of κ for which all the estimated coefficients are killed. It turns out that $\kappa = 0.5$. Consequently for $\kappa = 0.5$, we have noise-free reconstructions of the uniform density. Accordingly, this particular value of κ plays a benchmark value for the other types of density that we highlight. In other words, in the Example 2, which concerns the unimodal density we set $\kappa = 0.5$.

We can form a parallel to what happens on the real line. Indeed, in a certain way, this strategy for choosing κ in practise meets up with the universal threshold $\sqrt{2 \log n}$ put forward by Donoho and Johnstone [3] in the context of fixed design regression on $\mathbb{R}$ in order to "kill" asymptotically all the coefficients when estimating the zero function. Then, we compute the $\mathbb{L}_2$ and $\mathbb{L}_\infty$ losses and give the graphic reconstructions. This choice of κ turns out to be very reasonable and fruitful as the results prove to be good enough.

TABLE 1
Number of coefficients surviving thresholding for various values of κ,
$\phi \sim U[0, \pi/8]$

	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$\kappa = 0.2$	0	7	30	110
$\kappa = 0.3$	0	0	2	6
$\kappa = 0.4$	0	0	0	3

EXAMPLE 1.   In this first example, we consider the case of the uniform density $f = \frac{1}{4\pi}\mathbf{1}_{\mathbb{S}^2}$. It is easy to verify that $\beta_{j\eta} = \langle f, \psi_{j\eta} \rangle_{\mathbb{L}_2} = 0$ for every j and every η. Following Baldi et al. [1], a simple way of assessing the performance of the Needlet procedure is to count the number of coefficients surviving thresholding.

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 $\kappa = 0.5$ (see Tables 1 and 2). 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 $c = 1/0.7854$. The graph of f in the spherical coordinates (Φ, Θ) (Φ = longitude, $0 \le \Phi \le 2\pi$, Θ = colatitude, $0 \le \Theta \le \pi$) is given in Figure 2. We also plot the noisy observations for different cases of perturbations. For 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 to be 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 about the Oz axis by a large angle.

As motivated above, in the sequel, the tuning parameter κ for the Needlet procedure is set to 0.5 both for computing the quadratic loss, the L_∞ loss and the graphic reconstructions.

First of all, we have computed an estimate of the $\mathbb{L}_2$ and the $\mathbb{L}_\infty$ norms of the difference between $\hat{f}$ our Needlet estimator and f (see Table 3). For the quadratic

TABLE 2
Number of coefficients surviving thresholding for various values of κ,
$\phi \sim U[0, \pi]$

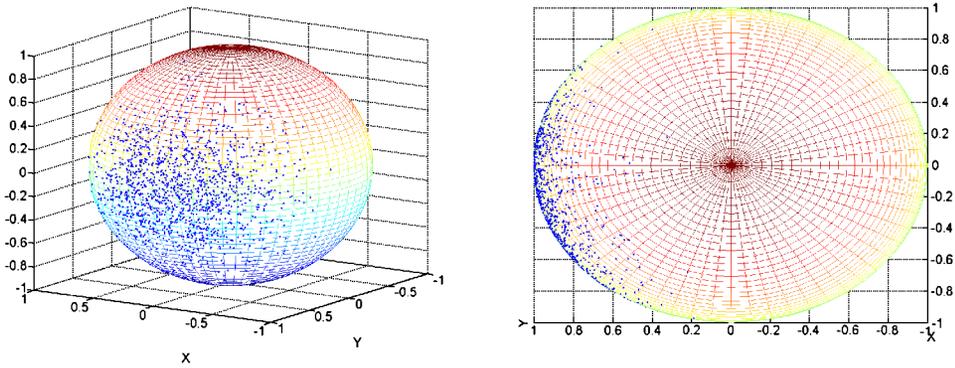|	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$\kappa = 0.2$	2	3	77	350
$\kappa = 0.3$	0	0	4	10
$\kappa = 0.4$	0	0	0	6

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

loss, we took the square root of the sum of the squares of the coefficients. As for the $\mathbb{L}_\infty$ distance, we chose an almost uniform grid of the sphere $\mathbb{S}^2$ given by the HEALPix pixelization program. We recall that

$$\|\hat{f} - f\|_\infty = \sup_{i=1,\ldots,L} |\hat{f}(\alpha_i) - f(\alpha_i)|,$$

where the α_i are the points of a uniform grid of the sphere. Here, we chose $L = 192 = 12.2^4$. All the estimated losses were computed over 50 runs and for $N = 1500$ observations. We considered the three cases of noise level described above. The results are summarized in the following table of errors which shows that the estimator performs quite well. In particular, its performances are deteriorating when the noise becomes more important (which was expected) and give very good results in $\mathbb{L}_\infty$ norm. We concentrated on particular phenomena instead of performing a large scanning of the errors in very diverse situations because the computations—although feasible—are rather costly in time when used repeatedly. For instance, to our knowledge, and probably for the same reason, no such study has been performed for the SVD method.

SVD versus needlet on a particular problem. A central issue in *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

TABLE 3
$\mathbb{L}_2$ and $\mathbb{L}_\infty$ estimated losses

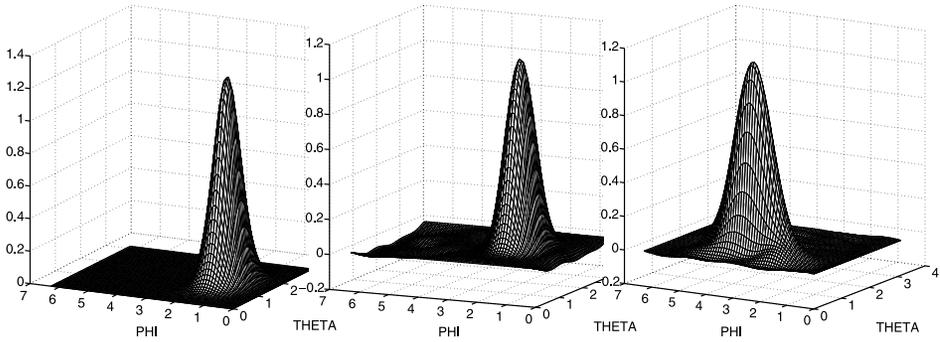|	$\phi \sim U[0, \pi/8]$	$\phi \sim U[0, \pi/4]$	$\phi \sim U[0, \pi/2]$
$\mathbb{L}_2$	0.3335	0.5523	0.7830
$\mathbb{L}_\infty$	0.1019	0.1677	0.1928

FIG. 3.   *The exact density, the Needlet procedure, the SVD method, $\phi \sim U[0, \pi/8]$.*

both on the sphere and on the flattened sphere and give the reconstructed density in the spherical coordinates for the Needlet procedure and the SVD estimator (see Figures 2, 4 and 6).

For each of the three groups of graphic reconstructions of the target density presented below corresponding to three cases of noise, you will find in order, the exact target density, then the one estimated with the Needlet procedure and finally the density estimated with the SVD method (see Figures 3, 5 and 7).

At a closer inspection, we notice that the position of the peak of the estimated bell is well localized by the Needlet procedure whatever the amount of noise, it is especially true for $\phi \sim U[0, \pi/8]$. Only in the case of the law $U[0, \pi/2]$, the longitude coordinate of the peak tends to slightly move away from the true value. As for the SVD method, for the three reconstructions, it fails to detect properly the exact position of the peak. Therefore, even if in the case of big rotations such that $\phi \sim U[0, \pi/4]$ and especially $\phi \sim U[0, \pi/2]$, the Needlet procedure allows us to have a rather good detection of the position of the peak. This is of course due to the remarkable concentration of the needlet. Of course, one remarks that the base
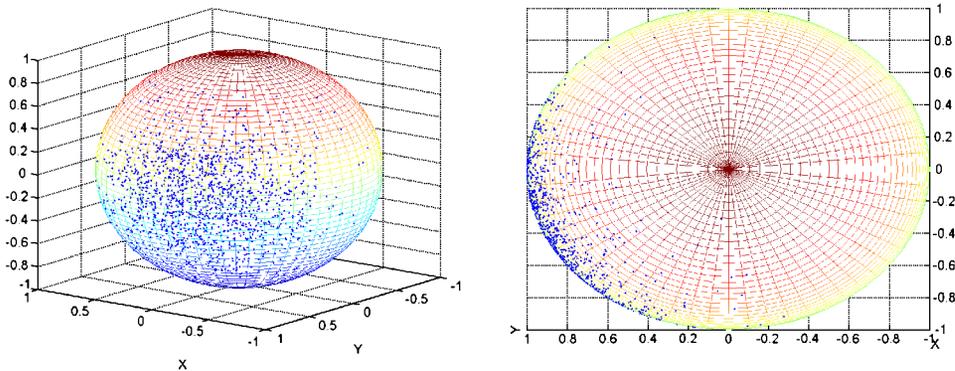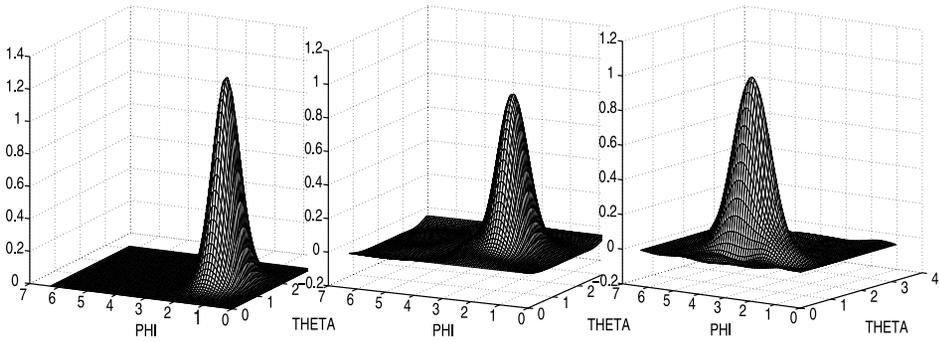

FIG. 4.   *Observations $\phi \sim U[0, \pi/4]$.*

FIG. 5. *The exact density, the Needlet procedure, the SVD method, $\phi \sim U[0, \pi/4]$.*

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 Theorems 2 and 3. 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 $q \geq 1$ there exist constants C, s_q, s_q', such that, as soon as $2^j \leq [\frac{N}{\log N}]^{1/2}$,*

$$(44) \qquad \mathbb{P}\{|\hat{\beta}_{j\eta} - \beta_{j\eta}| \geq \sigma_j v\} \leq 2 \exp\left\{-\frac{N v^2}{2(1 + 2/(\sqrt{3\pi} A) v 2^j)}\right\} \qquad \forall v > 0,$$

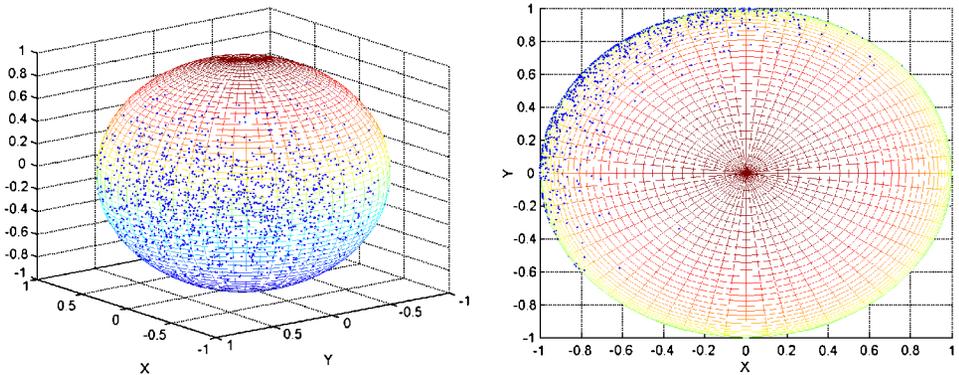$$(45) \qquad \mathbb{E}|\hat{\beta}_{j\eta} - \beta_{j\eta}|^q \leq s_q \left[\frac{\sigma_j^2}{N}\right]^{q/2},$$

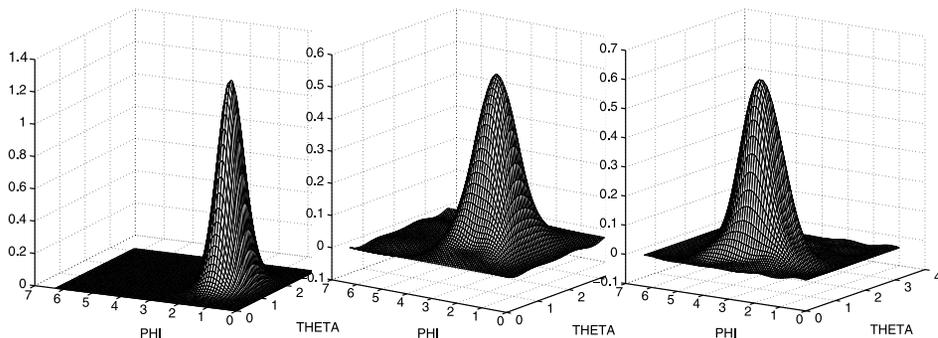

FIG. 6. *Observations $\phi \sim U[0, \pi/2]$.*

FIG. 7.   *The exact density, the Needlet procedure, the SVD method, $\phi \sim U[0, \pi/2]$.*

$$\text{(46)} \qquad \mathbb{E} \sup_{\eta} |\hat{\beta}_{j\eta} - \beta_{j\eta}|^q \leq s'_q (j+1)^q \left[\frac{\sigma_j^2}{N} \right]^{q/2},$$

$$\text{(47)} \qquad \mathbb{P}(|\hat{\beta}_{j\eta} - \beta_{j\eta}| \geq \sigma_j \kappa t_N) \leq 2N^{-\sqrt{3\pi} A\kappa/4} \qquad \forall \kappa \geq \sqrt{3\pi} A.$$

The proof of this proposition is given in the supplementary material (see Kerkyacharian et al. [9]).

6.1. *Proof of Theorem* 2.   Now, to get the result of Theorem 2, we begin by the following decomposition:

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

$$=: I + II.$$

The term *II* is easy to analyze, 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 $B_{p,r}^{s-(2/\pi-2/p)}(M')$, for some constant M' and for $\pi \leq p$. Hence,

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

Then we only need to verify that $\frac{s-2(1/\pi-1/p)}{\nu+1}$ 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 $\frac{s}{(s+\nu+1)} \leq \frac{s}{(\nu+1)p/\pi}$. 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 $\frac{s - 2(1/\pi - 1/p)}{\nu + 1}$ is always larger than $\mu = \frac{s - 2/\pi + 2/p}{s + \nu - 2/\pi + 1}$.

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

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

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

Now, we separate four cases:

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

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

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

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

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

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

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

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

We have the following upper bounds for the terms Bs and Sb:

$$Bs \leq C N^{-\sqrt{3\pi} A\kappa/16},$$

$$Sb \leq C N^{-\sqrt{3\pi} A\kappa/2 + p/(\nu+1)}.$$

It is easy to check that in any cases if $\sqrt{3\pi} A\kappa > \max\{8p, 2p + 1\}$ the terms Bs and Sb are smaller than the rates announced in the theorem. For the details of the above upper bounds of the terms Bs and Sb, see the supplementary material (Kerkyacharian et al. [9]).

Using Proposition 2, we have

$$Bb \leq C \sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} \sigma_j^p N^{-p/2} \|\psi_{j\eta}\|_p^p I\left\{|\beta_{j\eta}| \geq \frac{\kappa}{2} t_N \sigma_j\right\},$$

$$Ss \leq \sum_{j=-1}^{J} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^p \|\psi_{j\eta}\|_p^p I\{|\beta_{j\eta}| < 2\kappa t_N \sigma_j\}.$$

Using again Proposition 2, (25) and condition (37) for any $p \geq z \geq 0$:

$$Bb \leq C N^{-p/2} \sum_{j=-1}^{J} \sigma_j^p 2^{j(p-2)} \sum_{\eta \in \mathscr{Z}_j} I\left\{|\beta_{j\eta}| \geq \frac{\kappa}{2} t_N \sigma_j\right\}$$

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

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

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):

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

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

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

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

Now, as

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

and

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

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

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

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

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 $2^{j_0(p/q)(\nu+1)} \sim t_N{}^{-1}$, 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 $[\frac{1}{2^{2j}} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^{\tilde{q}}]^{1/\tilde{q}} \leq [\frac{1}{2^{2j}} \sum_{\eta \in \mathscr{Z}_j} |\beta_{j\eta}|^\pi]^{1/\pi}$, we get

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

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

$$\leq t_N{}^{p-\tilde{q}} \sum_{j=-1}^{j_0} 2^{j[(\nu+1)p(1-\tilde{q}/q)]} \tau_j^{\tilde{q}}$$

$$\leq C t_N{}^{p-\tilde{q}} 2^{j_0 p[(\nu+1)(1-\tilde{q}/q)]}$$

$$\leq C t_N{}^{p-q}.$$

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

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

In such a way that we easily verify that $p - q = p\frac{s - 2/\pi + 2/p}{1 + \nu + s - 2/\pi}, q - \pi = \frac{(p-\pi)(1+\nu) - \pi s}{s + \nu - 2/\pi + 1} > 0$. 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 $B^s_{\pi,r}(M)$,

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

$$\leq t_N{}^{p-\pi} \sum_{-1}^{j_0} 2^{j[(\nu+1-2/p)(p/q)(q-\pi)]} \tau_j^\pi$$

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

This is true since $\nu + 1 - \frac{2}{p}$ is also strictly positive since $\nu + 1 > \frac{s}{p/\pi - 1} \geq \frac{2}{p-\pi} \geq \frac{2}{p}$. If we now take $2^{j_0(p/q)(\nu+1-2/p)} \sim t_N{}^{-1}$, 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 $z_2 = \tilde{q} > q (> \pi)$

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

But

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

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

$$B \leq C t_N{}^{p-\tilde{q}} \sum_{j=j_0+1} 2^{j[(\nu+1-2/p)(p/q)(q-\tilde{q})]} \tau_j^{\tilde{q}}$$

$$\leq C t_N{}^{p-\tilde{q}} 2^{j_0[(\nu+1-2/p)(p/q)(q-\tilde{q})]}$$

$$\leq C t_N{}^{p-q},$$

which completes the proof of Theorem 2.

6.2. *Proof of Theorem* 3. The proof of this theorem is entirely given in the supplementary material (see Kerkyacharian et al. [9]).

SUPPLEMENTARY MATERIAL

Supplement to "Localized spherical deconvolution" (DOI: 10.1214/10-AOS858SUPP; .pdf). We give in the supplement some technical details for the understanding of the proofs of Propositions 1 and 2 and of Theorems 2 and 3.

REFERENCES

[1] BALDI, P., KERKYACHARIAN, G., MARINUCCI, D. and PICARD, D. (2009). Adaptive density estimation for directional data using needlets. *Ann. Statist.* **37** 3362–3395. MR2549563

[2] BALDI, P., KERKYACHARIAN, G., MARINUCCI, D. and PICARD, D. (2009). Subsampling needlet coefficients on the sphere. *Bernoulli* **15** 438–463. MR2543869

[3] DONOHO, D. L and JOHNSTONE, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81** 425–455. MR1311089

[4] GINÉ, E. and NICKL, R. (2010). Adaptive estimation of a distribution function and its density in sup-norm loss by wavelet and spline projections. *Bernoulli*. To appear.

[5] HEALY JR., D. M., HENDRIKS, H. and KIM, P. T. (1998). Spherical deconvolution. *J. Multivariate Anal.* **67** 1–22. MR1659108

[6] JOHNSTONE, I., KERKYACHARIAN, G., PICARD, D. and RAIMONDO, M. (2004). Wavelet deconvolution in a periodic setting. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66** 547–573. MR2088290

[7] KERKYACHARIAN, G., KYRIAZIS, G., LE PENNEC, E., PETRUSHEV, P. and PICARD, D. (2010). Inversion of noisy Radon transform by SVD based needlets. *Appl. Comput. Harmonic Anal.* **28** 24–45. MR2563258

[8] KERKYACHARIAN, G. and PICARD, D. (2009). New generation wavelets associated with statistical problems. In *The 8th Workshop on Stochastic Numerics* 119–146. Research Institute for Mathematical Sciences, Kyoto Univ.

[9] KERKYACHARIAN, G., PHAM NGOC, T. M and PICARD, D. (2010). Supplement to "Localized spherical deconvolution." DOI:10.1214/10-AOS858SUPP.

[10] KERKYACHARIAN, G., PETRUSHEV, P., PICARD, D. and WILLER, T. (2007). Needlet algorithms for estimation in inverse problems. *Electron. J. Statist.* **1** 30–76 (electronic). MR2312145

[11] KIM, P. T. and KOO, J. Y. (2002). Optimal spherical deconvolution. *J. Multivariate Anal.* **80** 21–42. MR1889831

[12] NARCOWICH, F., PETRUSHEV, P. and WARD, J. (2006). Decomposition of Besov and Triebel–Lizorkin spaces on the sphere. *J. Funct. Anal.* **238** 530–564. MR2253732

[13] NARCOWICH, F., PETRUSHEV, P. and WARD, J. (2006). Local tight frames on spheres. *SIAM J. Math. Anal.* **38** 574–594. MR2237162

[14] PIETROBON, D., BALDI, P., KERKYACHARIAN, G., MARINUCCI, D. and PICARD, D. (2008). Spherical needlets for CMB data analysis. *Monthly Notices of the Roy. Astron. Soc.* **383** 539–545.

[15] ROSENTHAL, J. S. (1994). Random rotations: Characters and random walks on $SO(N)$. *Ann. Probab.* **22** 398–423. MR1258882

[16] TALMAN, J. D. (1968). *Special Functions: A Group Theoretic Approach*. W. A. Benjamin, New York. MR0239154

[17] TERRAS, A. (1985). *Harmonic Analysis on Symmetric Spaces and Applications. I.* Springer, New York. MR0791406

[18] TOURNIER J. D., CALAMANTE, F., GADIAN, D. G. and CONNELLY, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted mri data using spherical deconvolution. *NeuroImage* **23** 1176–1185.

[19] VAN ROOIJ, A. C. M. and RUYMGAART, F. H. (1991). Regularized deconvolution on the circle and the sphere. In *Nonparametric Functional Estimation and Related Topics (Spetses, 1990). NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.* **335** 679–690. Kluwer Academic, Dordrecht. MR1154359

[20] VILENKIN, N. J. (1969). *Fonctions spéciales et théorie de la représentation des groupes. Traduit du Russe par D. Hérault. Monographies Universitaires de Mathématiques* **33**. Dunod, Paris. MR0243143

[21] WILLER, T. (2009). Optimal bounds for inverse problems with Jacobi-type eigenfunctions. *Statist. Sinica* **19** 785–800. MR2514188

G. KERKYACHARIAN
CNRS, LPMA
175 RUE DU CHEVALERET
75013 PARIS
FRANCE
E-MAIL: kerk@math.jussieu.fr

T. M. PHAM NGOC
LABORATOIRE DE MATHÉMATIQUES,
   UMR CNRS 8628
UNIVERSITÉ PARIS SUD
91405 ORSAY CEDEX
FRANCE
E-MAIL: thanh.pham_ngoc@math.u-psud.fr

D. PICARD
UNIVERSITÉ PARIS VII
175 RUE DU CHEVALERET
75013 PARIS
FRANCE
E-MAIL: picard@math.jussieu.fr