Journal of Multivariate Analysis 104 (2012) 16–38

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Adaptive nonparametric regression on spin fiber bundles

Claudio Durastanti [a], Daryl Geller [b], Domenico Marinucci [c,*]

[a] *Dipartimento di Matematica, Università di Pavia, Italy*
[b] *Department of Mathematics, Stony Brook University, United States*
[c] *Dipartimento di Matematica, Università di Roma Tor Vergata, Italy*

ARTICLE INFO

Article history:
Received 7 December 2010
Available online 27 May 2011

AMS 2000 subject classifications:
62G08
62G20
42B35
42C10
42C40
46E35

Keywords:
Spin fiber bundles
Mixed spin needlets
Adaptive nonparametric regression
Thresholding
Spin Besov spaces

ABSTRACT

The construction of adaptive nonparametric procedures by means of wavelet thresholding techniques is now a classical topic in modern mathematical statistics. In this paper, we extend this framework to the analysis of nonparametric regression on sections of spin fiber bundles defined on the sphere. This can be viewed as a regression problem where the function to be estimated takes as its values algebraic curves (for instance, ellipses) rather than scalars, as usual. The problem is motivated by many important astrophysical applications, concerning, for instance, the analysis of the weak gravitational lensing effect, i.e. the distortion effect of gravity on the images of distant galaxies. We propose a thresholding procedure based upon the (mixed) spin needlets construction recently advocated by Geller and Marinucci (2008, 2010) and Geller et al. (2008, 2009), and we investigate their rates of convergence and their adaptive properties over spin Besov balls.

1. Introduction

Over the last two decades, wavelet techniques have become a well-established tool for the analysis of statistical nonparametric problems, especially in the framework of minimax estimation. The seminal contribution in this area was provided by Donoho et al. in [16], where it was proved that nonlinear wavelet estimators based on thresholding techniques achieve nearly optimal minimax rates (up to logarithmic terms) for a wide class of nonparametric estimation of unknown density and regression functions. The theory has been enormously developed ever since (we refer to [33] for a textbook reference).

The bulk of this literature has focussed on estimation in standard Euclidean frameworks, such as $\mathbb{R}$ or $\mathbb{R}^n$. More recently, applications from various scientific fields have drawn a lot of attention on more general settings, such as spherical data or more general manifolds (see [1]). This environment has recently experienced a remarkable amount of activity, both from the purely mathematical point of view and in terms of applications to empirical data.

In particular, a highly successful construction of a second-generation wavelet system on the sphere (the so-called needlets) has been introduced by [51,52]; this approach has been extended to more general manifolds and unbounded support in the harmonic domain by [23–25]. The investigation of the stochastic properties of needlets when implemented on spherical random fields is due to [2,3,43,44,49], where applications to several statistical procedures are also considered. These procedures have been mainly motivated by issues arising in Cosmology and Astrophysics, and indeed several applications to experimental data have already been implemented: for instance, those from the satellite WMAP mission

* Corresponding author.
E-mail address: marinucc@mat.uniroma2.it (D. Marinucci).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2011.05.012

from NASA, focussing on the so-called Cosmic Microwave Background radiation, see [56,48,55,18,57,58,14,59,12,30]. These applications, however, have not been focussed on thresholding estimates and minimax results, but rather to random fields issues, such as angular power spectrum estimation, higher-order spectra, testing for Gaussianity and isotropy, and several others (see also [47,11]).

More recently, a few papers have focussed on the use of needlets to develop estimators within the thresholding paradigm, in the framework of directional data. The pioneering contribution here is due to [4], see also [37,38,36,32]; applications to astrophysical data is still under way. Earlier results on minimax estimators for spherical data, outside the needlets approach, are due to Kim and coauthors (see [40,39,42]).

Another important generalization of the needlet approach has been recently advocated by Geller and Marinucci [21]; applications to statistics can be found in [20]. This development is again motivated by Cosmology and Astrophysics. In particular, we noted above as some extremely influential satellite missions from NASA and ESA (WMAP and Planck, respectively) are currently collecting data on the so-called Cosmic Microwave Background radiation, which can be viewed as the realization of a scalar, isotropic, mean-square continuous spherical random field (see for instance [15] for a review). These same experiments are also collecting data on a much more elusive cosmological feature, the so-called polarization of CMB. The latter can be loosely described as observations on random ellipses living on the tangent planes for each location on the celestial sphere. Mathematically, this can be expressed by defining random sections of spin fiber bundles, a generalization of the notion of scalar random fields (see [21,20,22,26,27] and Section 2 below for much more details and discussion). Quite interestingly, exactly the same mathematical framework describes the so-called weak gravitational lensing induced on the observed shape of distant galaxies by clusters of matter. The applications of spin needlets to CMB polarization data is discussed in [19]. The characterization of spin Besov spaces by means of needlets decompositions is discussed by [5,22]; the latter reference also introduces an alternative construction for needlets on spin fiber bundles (so-called mixed needlets), and provide its analytical and statistical properties.

Our purpose in this article is to exploit these results and classical techniques to introduce and develop spin nonparametric regression, with a view to applications to polarization and weak lensing data. The latter is again a major issue in the analysis of astrophysical data (see for instance [7,41] and the references therein). To motivate our analysis, we describe gravitational lensing in the subsection below.

1.1. Motivations: gravitational lensing

Gravitational lensing is an astrophysical phenomenon due to the presence of celestial bodies along the path of the light during its travel from a source (i.e. a galaxy) to an observer. More precisely, the mass of these bodies "lenses" the path of photons creating a distortion on the image perceived by the observer. The effect of (weak) gravitational lensing on the perceived image can be described (to the first order) as follows. For a given position in the sky $\left(\widehat{\vartheta},\widehat{\varphi}\right)$, let us label x_1, y_1 and x_2, y_2 the coordinates of the original and the observed light distribution, (respectively); the following relationship holds (see for instance [41]):

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - |g| \begin{pmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix}\right] \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}.$$

The first order effects of weak gravitational lensing are described by the second addendum on the right term in the last equation, which is called the *shear*. The scalar $|g|$ conveys the stretching of the observed object, whose direction is represented by the angle ϕ. The factor 2 in the trigonometric terms entails that the shear is a *spin 2 random field*, that is, it is symmetric under rotations of 180 degrees, see the following sections for a much more detailed discussion on spin quantities. It is then natural to describe the shear in a given location as a complex variable:

$$g = |g|\, e^{2i\phi} = |g| \left(\cos(2\phi) + i \sin(2\phi)\right) = g_1 + ig_2. \tag{1}$$

It is then possible to rewrite (1) as:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}.$$

$$x_1 + iy_1 = (1-g)(x_2 + iy_2).$$

We are actually interested in a shear field, i.e. to analyze how g varies over the celestial sphere S^2. Mathematically, this is described by viewing $g(\vartheta,\varphi)$ as a complex line bundle, see below. Loosely speaking, g will reflect the distribution of the gravitational potential across all observed directions. Huge amount of observational data are expected in the next decade, by means of satellite missions in preparations such as Euclid.

1.2. The statistical problem

As we wrote before, our aim here is to use spin and mixed needlets as a tool for the optimal recovery of gravitational shear from observed data. In particular, we investigate the properties of nonlinear hard thresholding estimates, and we establish rates of convergence over a wide class of L^p_s norms and spin Besov spaces (see again [5,22] and the sections to follow for more detailed definitions). More precisely, we shall assume to have observations on independent pairs of random variables, respectively scalar and spin, $\left(X_i, Y_{i;s}\right)$, $i = 1, \dots, n$, $(X_i) \in \mathbb{S}^2$; we view (X_i) as uniform random locations on the sphere, which correspond for instance to the positions of observed galaxies. We shall then be concerned with the regression model:

$$Y_{i;s} = F_s\left(X_i\right) + \varepsilon_{i;s}, \tag{2}$$

where $F_s(\cdot)$ is an unknown section of a spin fiber bundle; as explained above, for $s = 2$ $F_s = g$ can be taken to represent the geometric effect of the gravitational shear. We assume that this section belongs to $L^p_s\left(\mathbb{S}^2\right)$, the space of the spin s, p-integrable sections on the sphere. On the other hand, we assume the $\varepsilon_{i;s}$ are i.i.d. spin random variables, which can be viewed as an observational error (to be interpreted, for instance, as the intrinsic shape of the galaxy). We are then led to nonparametric estimation over an unknown functional class, and we aim at procedures which are robust (i.e. nearly optimal) for a wide class of L^p_s norms, $1 \le p \le \infty$. To address this issue, and given the properties of (mixed) spin needlets established in [21,22], we follow a classical approach, as discussed for the classical case on $\mathbb{R}$ by Donoho et al. [16] and Hardle et al. [33,37], and many other papers, see for instance [8,13,34] for some recent developments. In particular, as mentioned before we introduce thresholding estimates and establish convergence rates for the resulting nonlinear estimators. We stress that we consider at the same time estimators based upon both spin constructions we have mentioned before, i.e. pure and mixed spin needlets; the results with the two approaches are identical. Sharp adaptation results for nonparametric regression on vector bundles were recently established in an important paper by [40]. These authors focus on the $p = 2$, and therefore exploit Fourier methods rather than wavelets thresholding. For $s = 1$, our method can be viewed as a form of adaptive regression for vector fields, and in this sense it relates also to recent work on filament estimation by [28,29]. See also [60] for some recent work on statistical analysis for tensor-valued data.

The plan of the paper is as follows. In Section 2 we review some background material on spin fiber bundles, while in Section 3 we recall the construction of spin and mixed spin needlets; for both sections we follow closely earlier references, in particular [21,22]. In Section 4 we review some crucial material on spin Besov spaces, as discussed earlier by [5,22]. Sections 5 and 6 include the most important contributions of this paper, namely the presentation of the thresholding procedure and the investigation of its asymptotic properties. The proofs here follow classical arguments in this literature, as discussed for instance by [33,38,54].

2. Spin functions

2.1. Background and definitions

The purpose of this section is to review some background material on spin fiber bundles; our presentation follows closely [21,20,22], to which we refer for more discussion and details. The concept of a spin function was introduced in the sixties by Newman and Penrose in [53], while working on gravitational radiation, see also [31,17]. Writing in a physicists' jargon, they said that a function η has an integer-valued spin weight s (or, briefly, that η is a spin s quantity) if, whenever a tangent vector at point $x \in \mathbb{S}^2$ is rotated by an angle ψ under a coordinate change, η transforms as $\eta' = e^{is\psi}\eta$. This same idea is formalized as follows by [21]. Let $U_I := \mathbb{S}^2 \setminus \{N, S\}$ be the chart that covers the sphere with the North and the South poles subtracted: here we adopt the usual angular coordinates (ϑ, φ), $\vartheta \in (0, \pi)$ and $\varphi \in [-\pi, \pi]$. Define the rotated charts $U_R = RU_I$, where $R \in SO(3)$ (the special group of rotations) and label the corresponding coordinates (ϑ_R, φ_R). For any $x \in \mathbb{S}^2$, we can fix a "reference direction" in the tangent plane at x (labeled as usual $T_x(\mathbb{S}^2)$) by considering $\rho_I(x) = \partial/\partial\varphi$, the unitary tangent vector in the direction of the circle where ϑ is constant and φ is increasing.

For every x belonging to the intersection between the charts corresponding to U_R and U_I, we can uniquely measure the angle associated to a change of coordinate by considering the angle between the reference vector in the map U_I, and the reference vector in the rotated chart, namely $\rho_R(x) = \partial/\partial\varphi_R$ (Fig. 1). More generally, given $x \in \mathbb{S}^2$ and two charts U_{R_1} and U_{R_2} such that $x \in U_{R_1} \cap U_{R_2}$, the angle between U_{R_1} and U_{R_2}, ψ_{x,R_1,R_2} is defined as the angle between $\rho_{R_1}(x)$ and $\rho_{R_2}(x)$ (Fig. 2), see [21,20] for a discussion on the orientation of this angle.

Fix now an open subset $G \subset \mathbb{S}^2$. The collection of functions $\{F_R\}_{R\in SO(3)}$ is a spin s function F_s if and only if $\forall R_1, R_2 \in SO(3)$ and all $x \in U_{R_1} \cap U_{R_2} \cap G$ we have:

$$F_{R_2} = e^{is\psi_{x,R_1,R_2}} F_{R_1}.$$

We write $F_s \in C^\infty_s(G)$, if for every $R \in SO(3)$ the application $x \to F_s(x)$ is smooth. Note that for $s = 0$ we are back to the usual scalar functions.

From a differential geometry point of view, C^∞_s is the space of sections over G of the complex line bundle over the sphere $\mathbb{S}^2$ (see also [45,46] for more discussion on this point of view). The functional spaces $L^p_s\left(\mathbb{S}^2\right)$ are then defined as

$$F_s \in L^p_s\left(\mathbb{S}^2\right) \Leftrightarrow \|F_s\|_{L^p_s(\mathbb{S}^2)} = \left(\int_{\mathbb{S}^2} |F_s(x)|^p \, dx\right)^{1/p} < \infty.$$

Note that, while $F_s(x)$ is a section of the fiber bundle on $\mathbb{S}^2$, $|F_s(x)|$ is a real valued function on the sphere, because the modulus of F_s does not depend on the choice of the coordinate system: therefore the $L^p_s\left(\mathbb{S}^2\right)$ is well defined.

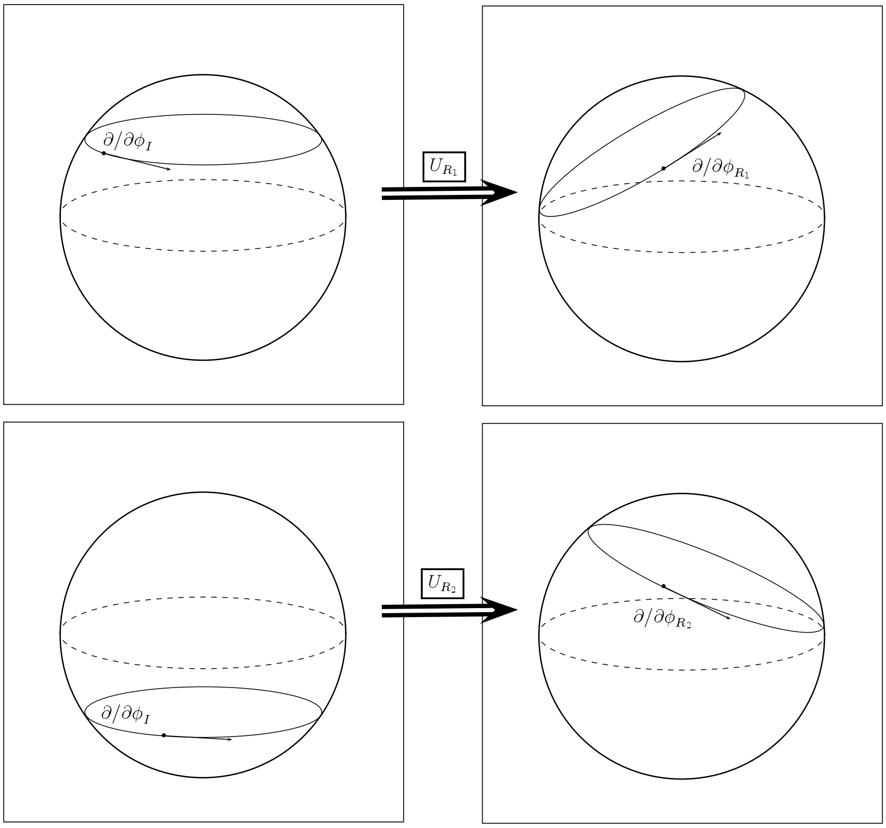

Fig. 1. The rotated charts U_{R_1} and U_{R_2}.

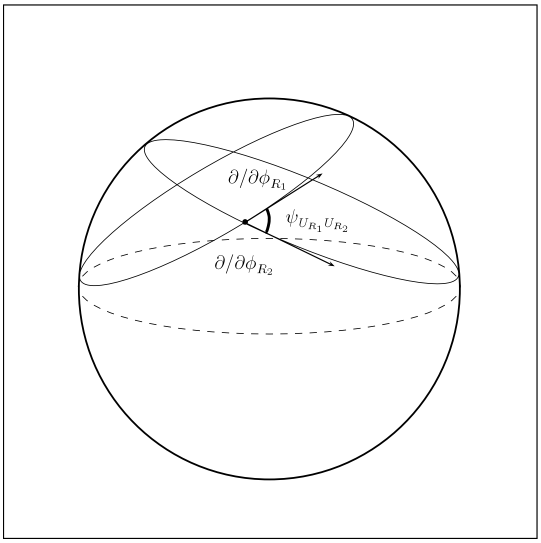

Fig. 2. The angle $\psi_{U_{R_1}U_{R_2}}$.

2.2. Spin spherical harmonics

We start by recalling the well-known expression for the spherical Laplacian $\Delta_{\mathbb{S}^2}$,

$$\Delta_{\mathbb{S}^2} := \frac{1}{\sin^2\vartheta}\frac{\partial}{\partial\varphi^2} + \frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left\{\sin\vartheta\frac{\partial}{\partial\vartheta}\right\}.$$

A complete orthonormal set of eigenfunctions for the spherical Laplacian is provided by the family of spherical harmonics $\{Y_{lm}\}$, $l = 0, 1, 2, \ldots, m = -l, \ldots, l$:

$$\Delta_{\mathbb{S}^2} Y_{lm} = -l(l+1)Y_{lm}, \qquad \int_{\mathbb{S}^2} Y_{lm}(x)\overline{Y}_{lm}(x)\mathrm{d}x = \delta_l^{l\prime}\delta_m^{m\prime},$$

where δ_a^b denotes the Kronecker delta function. In the spherical coordinates (ϑ, φ)

$$Y_{lm}(\vartheta, \varphi) = \mathrm{e}^{\mathrm{i}m\varphi}\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_{lm}(\cos\vartheta),$$

$$P_{lm}(x) = (1-x^2)^{m/2}\frac{\mathrm{d}}{\mathrm{d}x^m}P_l(x),$$

where $P_l(x)$ denotes the Legendre polynomials, see for instance [62] for more analytic expressions and discussion. Denoting by $\{\mathcal{H}_l\}$ the linear spaces spanned by the spherical harmonics, the following decomposition holds (see for instance [1]):

$$L^2\left(\mathbb{S}^2\right) = \bigoplus_{l\geq 0}\mathcal{H}_l,$$

that is, in the L^2 sense, for all $f \in L^2\left(\mathbb{S}^2\right)$

$$f(x) = \sum_{l,m} a_{lm}Y_{lm}(x), \quad a_{lm} = \int_{\mathbb{S}^2} f(x)\overline{Y}_{lm}(x)\mathrm{d}x.$$

It is possible to introduce spin spherical harmonics as the eigenfunctions of a second-order differential operator which generalizes the spherical Laplacian (refer again to [64,21,50] for more details). To this aim, consider the (*spin raising* and *spin lowering*) operators $\eth$ and $\overline{\eth}$, whose action on a spin function $F_s(\cdot)$ is provided by:

$$\eth F_s(\vartheta, \varphi) = -(\sin(\theta))^s\left[\frac{\partial}{\partial\vartheta} + \frac{\mathrm{i}}{\sin(\vartheta)}\frac{\partial}{\partial\varphi}\right](\sin(\theta))^{-s}F_s(\vartheta, \varphi),$$

$$\overline{\eth}F_s(\vartheta, \varphi) = -(\sin(\theta))^{-s}\left[\frac{\partial}{\partial\vartheta} - \frac{\mathrm{i}}{\sin(\vartheta)}\frac{\partial}{\partial\varphi}\right](\sin(\theta))^{s}F_s(\vartheta, \varphi).$$

It should be noted that $\eth$ transforms spin s functions into spin $s+1$ functions, $\eth C_s^\infty \to C_{s+1}^\infty$, while $\overline{\eth}$ transforms spin s functions into spin $s-1$ functions, $\overline{\eth}C_s^\infty \to C_{s-1}^\infty$, which justifies their names. The previous expressions should be written more rigorously in terms of $\eth_R$, $\overline{\eth}_R$, ϑ_R, φ_R, $F_{s;R}$, because both the operators and the spin functions depend on the choice of coordinates. More important, $\eth$, $\overline{\eth}$ can be used to define a differential operator $\eth\overline{\eth}$, which can be viewed as a generalization of the scalar spherical Laplacian; indeed

$$-\overline{\eth}\eth Y_{lm;s} = e_{ls}Y_{lm;s},$$

where $\{e_{ls}\}_{l=s,s+1} = \{(l-s)(l+s+1)\}_{l=s,s+1}$ is the associated sequence of eigenvalues and $\left\{Y_{lm;s}\right\}$, $l = s, s+1, \ldots; m = -l, \ldots, l$ is the sequence of orthonormal spherical harmonics, which we define by

$$Y_{lm;s} := \sqrt{\frac{(l-s)!}{(l+s)!}}\eth Y_{lm} \quad \text{for } s > 0,$$

$$Y_{lm;s} := \sqrt{\frac{(l+s)!}{(l-s)!}}\eth Y_{lm} \quad \text{for } s < 0.$$

Again, as before it should be noted that in the spin case the operators depend on the choice of the coordinates, differently from the scalar case. As discussed by [20,45,46] the spin construction could be alternatively provided in terms of the so-called spin-weighted representation of the special group of rotations $SO(3)$, indeed spin spherical harmonics can be related to the so-called Wigner's matrices, see again [62,63]. In particular, it is then possible to show that the spin spherical harmonics are themselves an orthonormal system, i.e. they satisfy

$$\int_{\mathbb{S}^2} Y_{lm;s}\overline{Y}_{lm;s}\mathrm{d}x = \int_0^{2\pi}\int_0^{\pi} Y_{lm;s}(\vartheta, \varphi)\overline{Y}_{lm;s}(\vartheta, \varphi)\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi = \delta_l^{l'}\delta_m^{m'}.$$

As for the scalar case,

$$L_s^2\left(\mathbb{S}^2\right) = \bigoplus_{l=0}^{\infty}\mathcal{H}_l \quad \mathcal{H}_l := span\left\{Y_{lm;s}; m = -l, \ldots, l\right\},$$

and the following representation holds

$$F_s(x) = \sum_l \sum_m a_{lm;s} Y_{lm;s}(x),$$

in the L_s^2 sense, i.e.

$$\lim_{L\to\infty} \int_{\mathbb{S}^2} \left| F_s(x) - \sum_{l=|s|}^{L} \sum_{m=-l}^{l} a_{lm;s} Y_{lm;s}(x) \right|^2 dx = 0.$$

Here, the spherical harmonics coefficients $a_{lm;s} := \int_{\mathbb{S}^2} F_s \overline{Y}_{lm} dx$ are such that

$$a_{lm;s} = a_{lm;E} + i a_{lm;M},$$

where $\{a_{lm;E}\}, \{a_{lm;M}\}$ are the coefficients of two standard (scalar-valued) spherical functions, which in the physical literature are labeled the electric and magnetic components of the spin function F_s, see again [21,22] for more discussion.

3. Spin and mixed needlets

3.1. Definition

We start by recalling the definition of scalar needlets, which were introduced by [51,52] as:

$$\psi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_l b\left(\frac{l}{B^j}\right) \sum_{m=-l}^{l} Y_{lm}(x) \overline{Y}_{lm}(\xi_{jk}), \quad \forall x \in \mathbb{S}^2;$$

here $\{\xi_{jk}, \lambda_{jk}\}$ are a set of cubature points and weights ensuring that:

$$\sum_{jk} \lambda_{jk} Y_{lm}(\xi_{jk}) \overline{Y}_{l'm'}(\xi_{jk}) = \int_{\mathbb{S}^2} Y_{lm}(x) \overline{Y}_{l'm'}(x)\, dx = \delta_l^{l'} \delta_m^{m'},$$

$b(\cdot)$ is a compactly supported C^∞ function satisfying the partition of unity property:

$$\sum_j b^2\left(\frac{l}{B^j}\right) \equiv 1$$

for all $l \geq 1$, and $B > 1$ is a bandwidth parameter. For a fixed value of B, we denote $\{\mathcal{X}_j\}_{j\geq 0}$ the nested sequence of cubature points corresponding to the space $\mathcal{K}_{[2B^j+1]}$, where $[\cdot]$ represents the integer part and $\mathcal{K}_L = \bigoplus_{l=0}^{L} \mathcal{H}_l$ is the space spanned by spherical harmonics up to order L. For each j, the cubature points are almost distributed as an α_j-net, with $\alpha_j := kB^{-j}$, the coefficients $\{\lambda_{jk}\}$ are such that $cB^{-2j} \leq \lambda_{jk} \leq CB^{-2j}$, with $c, C \in \mathbb{R}$, and $N_j = card\{\mathcal{X}_j\} \approx B^{2j}$, see for instance [3] for more details.

The construction of spin needlets (as provided by [21]) is formally similar to the scalar case, although as we discuss below it entails deep differences in terms of the spaces involved. Indeed, spin needlets are defined as follows:

$$\psi_{jk;s}(x) = \sqrt{\lambda_{jk}} \sum_l b\left(\frac{\sqrt{e_{l,s}}}{B^j}\right) \sum_{m=-l}^{l} \overline{Y}_{lm;s}(\xi_{jk}) Y_{lm;s}(x), \tag{3}$$

where $\{\lambda_{jk}, \xi_{jk}\}$ are, as before, cubature weights and cubature points, $b(\cdot) \in C^\infty$ is nonnegative, it is compactly supported in $[1/B, B]$ and satisfies the partition of unity property. Note, however, that the mathematical meaning of (3) is rather different from the scalar case; indeed $\psi_{jk;s}(x)$ is to be viewed as a spin s function with respect to rotations of the tangent plane $\mathbb{T}_x$, and a spin $-s$ function with respect to rotations of the tangent plane $\mathbb{T}_{\xi_{jk}}$. Moreover, as $Y_{lm;s}(\xi_{jk}), Y_{lm;s}(x)$ live on two different tangent planes $\mathbb{T}_{\xi_{jk}}, \mathbb{T}_x$, the product $\overline{Y}_{lm;s}(\xi_{jk}) Y_{lm;s}(x)$ is not defined and the notation $\overline{Y}_{lm;s}(\xi_{jk}) \otimes Y_{lm;s}(x)$ would be more appropriate. As a consequence, the spin needlet operators acts on spin s functions to produce spin s coefficients

$$\begin{aligned}\langle F_s, \psi_{jk;s}(x)\rangle &= \int_{\mathbb{S}^2} F_s(x) \overline{\psi}_{jk;s}(x)\, dx \\ &= \sqrt{\lambda_{jk}} \sum_{lm} b\left(\frac{\sqrt{e_{l,s}}}{B^j}\right) a_{lm;s} Y_{lm;s}(\xi_{jk}) \\ &=: \beta_{jk;s}.\end{aligned} \tag{4}$$

Therefore, $\psi_{jk;s}$ induces the linear map (4) from spin s quantities to spin s wavelet coefficients $\beta_{jk;s}$, while in the scalar case ($s = 0$) needlets generate a linear map from scalar quantities to scalar quantities. Indeed, if u is a spin s vector at ξ_{jk}, $\psi_{jk;s}(x)\, u$ becomes a spin s vector at ξ_{jk}, since the product of spin $-s$ and spin s vectors *at a point* x is a well-defined complex number, independently of the choice of coordinate system.

To provide a clearer interpretation to the previous expression, recall the decomposition of the functional space $L^2_s(\mathbb{S}^2) = \bigoplus_{l\geq 0} \mathcal{H}_l$. We can hence define the following operators on $\mathcal{H}_l$:

$$K_j(x,y) = \sum_l b^2\left(\frac{\sqrt{e_{l,s}}}{B^j}\right) Y_{lm;s}(x)\,\overline{Y}_{lm;s}(y)$$

$$\Lambda_j(x,y) = \sum_l b\left(\frac{\sqrt{e_{l,s}}}{B^j}\right) Y_{lm;s}(x)\,\overline{Y}_{lm;s}(y)$$

such that the reproducing kernel property holds:

$$\int_{\mathbb{S}^2} \Lambda_j(x,y)\,\overline{\Lambda}_j(y,z)\,dy = K_j(x,z).$$

Spin needlets can be derived by discretizing this operator by using the reproducing kernel property. In fact Λ_j is such that:

$$z \to \Lambda_j(x,z) \in \mathcal{K}_{[B^{2j+1}]},$$

and therefore:

$$z \to \Lambda_j(x,z)\,\overline{\Lambda}_j(z,y) \in \mathcal{K}_{[B^{4j+2}]}.$$

After discretization, we obtain:

$$K_j(x,y) = \sum_{\xi_{jk}\in\mathcal{K}_{[B^{4j+2}]}} \lambda_{jk}\Lambda_j(x,\xi_{jk})\,\overline{\Lambda}_j(\xi_{jk},y),$$

where we exploit the fact that the pairs $\{\lambda_{jk},\xi_{jk}\}$ can be chosen to form exact cubature points and weights [5]. Then

$$\begin{aligned} K_j f(x) &= \int_{\mathbb{S}^2} K_j(x,y) f(y)\,dy \\ &= \int_{\mathbb{S}^2} \sum_{\xi_{jk}\in\mathcal{K}_{[B^{4j+2}]}} \lambda_{jk}\Lambda_j(x,\xi_{jk})\,\overline{\Lambda}_j(\xi_{jk},y) f(y)\,dy \\ &= \sum_{\xi_{jk}\in\mathcal{K}_{[B^{4j+2}]}} \sqrt{\lambda_{jk}}\Lambda_j(x,\xi_{jk}) \int_{\mathbb{S}^2} \sqrt{\lambda_{jk}}\,\overline{\Lambda}_j(\xi_{jk},y) f(y)\,dy \\ &= \sum_{\xi_{jk}\in\mathcal{K}_{[B^{4j+2}]}} \beta_{jk;s}\psi_{jk;s}, \end{aligned}$$

where

$$\psi_{jk;s} = \sqrt{\lambda_{jk}}\Lambda_j(x,\xi_{jk}).$$

As a minor point, note that for the argument of the function $b(\cdot)$ we have used here the square root of $e_{l,s}$, the eigenvalue of the corresponding spin spherical harmonics, while in the scalar case [51,52] proposed to adopt l. However it is trivial to observe that, for fixed s:

$$\lim_{l\to\infty} \frac{\sqrt{e_{l,s}}}{l} = \lim_{l\to\infty} \frac{\sqrt{(l-s)(l+s+1)}}{l} = 1.$$

3.2. Some properties

We report some important properties for spin needlets, very similar to those in scalar case (see [51,52]). Indeed, from the previous discussion it follows easily that $\left|\psi_{jk;s}\right|^2$ is a well-defined scalar quantity. The following Localization property is hence well defined (see [21]): for any $M \in \mathbb{N}$, there exists a constant $c_M > 0$ such that for every $x \in \mathbb{S}^2$:

$$\left|\psi_{jk;s}(x)\right| \leq \frac{c_M B^j}{\left(1 + B^j \arccos\left(\langle \xi_{jk}, x\rangle\right)\right)^M}.$$

Let us recall from (4) that

$$\beta_{jk;s} = \int_{\mathbb{S}^2} F_s(x)\,\overline{Y}_{jk;s}(x)\,dx = \sqrt{\lambda_{jk}} \sum_l b\left(\frac{\sqrt{e_{ls}}}{B^j}\right) \sum_{m=-l}^{l} a_{lm;s} Y_{lm;s}(\xi_{jk}),$$

and the following *reconstruction formula* holds:

$$F_s(x) = \sum_j \sum_k \beta_{jk;s} \psi_{jk;s}(x).$$

It is simple to check that the squared coefficients $\left|\beta_{jk;s}\right|^2$ following quantities are scalar. In the following, we will need both the $L_s^2(\mathbb{S}^2)$ and the $L_s^p(\mathbb{S}^2)$ norm of $\psi_{jk;s}$. Let us start by observing that:

$$\begin{aligned}
\left\|\psi_{jk;s}\right\|^2_{L_s^2(\mathbb{S}^2)} &= \lambda_{jk} \sum_l b^2\left(\frac{\sqrt{e_{ls}}}{B^j}\right) \sum_{m=-l}^{l} Y_{lm;s}(\xi_{jk}) \overline{Y}_{lm;s}(\xi_{jk}) \int_{\mathbb{S}^2} Y_{lm;s}(x) \overline{Y}_{lm;s}(x)\, dx \\
&= \lambda_{jk} \sum_{l=B^{j-1}}^{B^{j+1}} b^2\left(\frac{\sqrt{e_{ls}}}{B^j}\right) \sum_{m=-l}^{l} Y_{lm;s}(\xi_{jk}) \overline{Y}_{lm;s}(\xi_{jk}) \\
&= \lambda_{jk} \sum_{l=B^{j-1}}^{B^{j+1}} \frac{2l+1}{4\pi} b^2\left(\frac{\sqrt{e_{ls}}}{B^j}\right) =: \tau^2_{jk;s}.
\end{aligned}$$

As discussed by [3,5,21], there exist positive constants c_1, c_2 such that $c_1 N_j^{-1} \leq \lambda_{jk} \leq c_2 N_j^{-1}$.

Throughout the rest of the paper, to simplify notations we shall assume to be dealing with sections of line bundles such that $F_s = (I - P_s)F_s$, P_s denoting the projection operator on the s spin spherical harmonics. In other words, the component at $l = s$ is assumed to be null; from the point of view of motivating applications, this is a very reasonable assumption, indeed for polarization or weak lensing experiments the so-called quadrupole term $l = s = 2$ has no physical meaning. The situation is indeed analogous to the standard scalar case, where the constant term $s = 0$ cannot even be measured by ongoing (so-called *differential*) experiments. Under these circumstances, as shown in [5], spin needlets make up a tight frame system, i.e. for all $F_s \in L_s^2(\mathbb{S}^2)$,

$$\|F_s\|^2_{L_s^2(\mathbb{S}^2)} = \sum_{jk} \left|\beta_{jk;s}\right|^2,$$

whence we have easily

$$\sum_{jk} \left|\langle \psi_{j_1,k_1;s}, \psi_{jk;s}\rangle\right|^2 = \left\|\psi_{j_1,k_1;s}\right\|^4_{L_s^2(\mathbb{S}^2)} + \sum_{j\neq j_1, k\neq k_1} \left|\langle \psi_{j_1,k_1;s} | \psi_{jk;s}\rangle\right|^2 \leq \left\|\psi_{j_1,k_1;s}\right\|^2_{L_s^2(\mathbb{S}^2)},$$

whence

$$\left\|\psi_{jk;s}\right\|_{L_s^2(\mathbb{S}^2)} \leq 1.$$

More generally, it is shown in [5,22] that for all $1 \leq p \leq \infty$, there exist positive constants c_p, C_p such that

$$c_p B^{2j\left(\frac{1}{2}-\frac{1}{p}\right)} \leq \left\|\psi_{jk;s}\right\|_{L_s^p(\mathbb{S}^2)} = \left(\int_{\mathbb{S}^2} \left|\psi_{jk;s}\right|^p dx\right)^{\frac{1}{p}} \leq C_p B^{2j\left(\frac{1}{2}-\frac{1}{p}\right)}. \tag{5}$$

3.3. Mixed needlets and their properties

Mixed Needlets were introduced in [22]; they are defined as

$$\psi_{jk;s\mathcal{M}}(x) = \sqrt{\lambda_{jk}} \sum_{l\geq|s|} b\left(\frac{\sqrt{e_{ls}}}{B^j}\right) \sum_m Y_{lm;s}(x) \overline{Y}_{lm}(\xi_{jk}),$$

with corresponding needlet coefficients

$$\beta_{jk;s\mathcal{M}} = \int_{\mathbb{S}^2} \overline{\psi}_{jk;s\mathcal{M}}(x) F_s(x)\, dx.$$

Mixed needlets form a tight frame system, with the same set of cubature points and weights as for the scalar case, $\left\{\xi_{jk}, \lambda_{jk}\right\}$. When $F_s \in L_s^2(\mathbb{S}^2)$, we have also

$$\beta_{jk;s\mathcal{M}} = \sqrt{\lambda_{jk}} \sum_{l\geq|s|} b\left(\frac{\sqrt{e_{ls}}}{B^j}\right) \sum_m a_{lm;s} Y_{lm}(\xi_{jk}),$$

and mixed needlets form a tight frame system. It should be noted that the coefficients $\left\{\beta_{jk;s\mathcal{M}}\right\}$ are scalar, complex-valued random variables, indeed for square integrable sections we have

$$\beta_{jk;s\mathcal{M}} = \sqrt{\lambda_{jk}} \sum_{l\geq|s|} b\left(\frac{\sqrt{e_{ls}}}{B^j}\right) \sum_m \left\{a_{lm;E} + ia_{lm;M}\right\} Y_{lm}\left(\xi_{jk}\right)$$
$$=: \beta_{jk;E} + i\beta_{jk;M},$$

where $\beta_{jk;E}$, $\beta_{jk;M}$ could be viewed as the scalar needlet coefficients of standard square integrable functions on the sphere. For general $F_s \in L^p_s\left(\mathbb{S}^2\right)$ the reconstruction formula holds, in the L^p_s sense:

$$F_s = \sum_j \sum_k \beta_{jk;s\mathcal{M}} \psi_{jk;s\mathcal{M}}.$$

Other properties of mixed needlets are analogous to those for the pure spin construction. In particular, note that scalar and pure spin needlets are both constructed by a convolution of a smooth function $b(.)$ with projection operators such as, for instance, $\sum_m Y_{lm}(x)\overline{Y}_{lm}(y)$, $\sum_m Y_{lm;s}(x)\overline{Y}_{lm;s}(y)$.

On the other mixed needlets are built by convolving $b(.)$ with $\sum_m Y_{lm}(x)\overline{Y}_{lm;s}(y)$, which is not a projection operator (indeed $\sum_m Y_{lm}(x)\overline{Y}_{lm;s}(x) \equiv 0$). It comes therefore to some extent as a surprise that mixed needlets do indeed enjoy localization properties, indeed we have (see again [22]): for each $M > 0$ there exists a constant C_M such that:

$$\left|\psi_{jk;s\mathcal{M}}\right| \leq \frac{C_M B^j}{\left(1 + B^j \arccos\left(\langle x, \xi_{jk}\rangle\right)\right)^M}.$$

Building upon this localization property, it is indeed possible to establish the following bounds (see for more details [22]):

$$c_1 B^{2j\left(\frac{1}{2}-\frac{1}{p}\right)} \leq \left\|\psi_{jk;s\mathcal{M}}\right\|_{L^p_s\left(\mathbb{S}^2\right)} \leq c_2 B^{2j\left(\frac{1}{2}-\frac{1}{p}\right)}, \quad c_1, c_2 > 0. \tag{6}$$

These constraints on the L^p_s norms will have the greatest importance for our results to follow. Also, for positive constants c_3, c_4 and arbitrary coefficients λ_k we have

$$c_3 \sum_k |\lambda_k|^p \left\|\psi_{jk;s\mathcal{M}}\right\|^p_{L^p_s\left(\mathbb{S}^2\right)} \leq \left\|\sum_k \lambda_k \psi_{jk;s\mathcal{M}}\right\|^p_{L^p_s\left(\mathbb{S}^2\right)} \leq c_4 \sum_k |\lambda_k|^p \left\|\psi_{jk;s\mathcal{M}}\right\|^p_{L^p_s\left(\mathbb{S}^2\right)}. \tag{7}$$

Remark 1. While the mathematical construction and the properties that can be developed on the mixed needlets are very similar to the spin case, there is a very relevant difference among these approaches that will be very important for our purposes. While, as we have already seen, $\psi_{jk;s}$ is formed by a tensorial product among two terms belonging to two different spaces of spin $-s$ and s such that β_{jk_s} belongs to the spin s space, $\psi_{jk;s\mathcal{M}}$ induces a linear map from a spin s vector at ξ_{jk} to a scalar (spin 0) quantity, such that for a spin s quantity u, the product $\overline{\psi}_{jk;s\mathcal{M}} \cdot u$ is always a scalar quantity.

4. Spin Besov spaces

Our aim in this section is to recall the definition of spin Besov spaces in terms of approximation properties. These definitions and their characterizations were provided by [5,22], to which we refer for further details and discussion. Define first,

$$G_k\left(F_s, \pi\right) = \inf_{H_s\in\mathcal{H}_k} \left\|F_s - H_s\right\|_{L^\pi_s\left(\mathbb{S}^2\right)},$$

i.e. the approximation error when replacing F_s by an element in $\mathcal{H}_{k;s}$. Then the Besov spin space $\mathcal{B}^r_{pq;s}$ is defined as the space of functions such that $F_s \in L^p_s\left(\mathbb{S}^2\right)$ and

$$\left(\sum_{k=0}^\infty \frac{1}{k}\left(k^r G_k\left(F_s, \pi\right)\right)^q\right) < \infty.$$

As usual, the last condition can be easily shown to be equivalent to

$$\left(\sum_{j=0}^\infty \left(B^{jr} G_{B^j}\left(F_s, \pi\right)\right)^q\right) < \infty.$$

Moreover, $F_s \in \mathcal{B}^r_{\pi q;s}$ if and only if, for every $j = 1, 2, \ldots$

$$\left(\sum_k \left(\left| \beta_{jk;s} \right| \left\| \psi_{jk;s} \right\|_{L^\pi_s(\mathbb{S}^2)} \right)^\pi \right)^{\frac{1}{\pi}} = \varepsilon_j B^{-jr}$$

where $\varepsilon_j \in \ell_q$ and $B > 1$. By defining the Besov norm as follows,

$$\|F_s\|_{\mathcal{B}^r_{\pi q;s}} = \begin{cases} \|F_s\|_{L^\pi_s(\mathbb{S}^2)} + \left[\sum_j B^{jq\left(r+\frac{1}{2}-\frac{1}{\pi}\right)} \left\{ \sum_k \left|\beta_{jk;s}\right|^\pi \right\}^{\frac{q}{\pi}} \right]^{\frac{1}{q}} & \text{if } q < \infty \\ \|F_s\|_{L^\pi_s(\mathbb{S}^2)} + \sup_j B^{j\left(r+\frac{1}{2}-\frac{1}{\pi}\right)} \left\| \left(\beta_{jk;s}\right)_k \right\|_{\ell_\pi} & \text{if } q = \infty, \end{cases}$$

we obtain that, if $\max(0, 1/\pi - 1/q) < r$ and $\pi, q > 1$, then

$$F_s \in \mathcal{B}^r_{\pi q;s} \Leftrightarrow \|F_s\|_{\mathcal{B}^r_{\pi q;s}} < \infty.$$

Besov spaces are characterized by come convenient embeddings, which (as always in this literature) will play a crucial role in our proofs to follow. More precisely, we have that, for $\pi_1 \le \pi_2$, $q_1 \le q_2$

$$\mathcal{B}^r_{\pi q_1;s} \subset \mathcal{B}^r_{\pi q_2;s}, \ \mathcal{B}^r_{\pi_2 q;s} \subset \mathcal{B}^r_{\pi_1 q;s}, \ \mathcal{B}^r_{\pi_1 q;s} \subset \mathcal{B}^{r-\frac{1}{\pi_1}+\frac{1}{\pi_2}}_{\pi_2 q;s}. \tag{8}$$

The proof of (8) is exactly the same as for the scalar case, see [4]. In particular

$$\begin{aligned} \mathcal{B}^r_{\pi_1 q;s} \subset \ \mathcal{B}^{r-\frac{1}{\pi_1}}_{\infty\infty;s} &\Longrightarrow \sup_k \left|\beta_{jk;s}\right| \left\|\psi_{jk}\right\|_{L^\infty_s(\mathbb{S}^2)} = \varepsilon_j B^{-j\left(r-\frac{1}{\pi_1}\right)} \\ &\Rightarrow B^{j(r+1-\frac{1}{\pi_1})} \sup_k \left|\beta_{jk;s}\right| < \infty \\ &\Rightarrow B^j \sup_k \left|\beta_{jk;s}\right| < \infty. \end{aligned}$$

5. Nonparametric regression on spin fiber bundles

5.1. The regression model

We start by recalling the regression formula (2):

$$Y_{i;s} = F_s(X_i) + \varepsilon_{i;s}.$$

Throughout this paper, we shall also assume that $\sup_x |F_s(x)| = M < \infty$. As discussed in the Introduction, we envisage a situation where it is possible to collect data which can be viewed as measurements on a spin fiber bundles, i.e. for instance the polarization of the Cosmic Microwave Background (see [35,61,10,20,19]), or the Weak Gravitational Lensing effect on the images of distant Galaxies (see [7]). To fix ideas, we focus on this second example. As discussed for instance in [15], the gravitational *shear*effect may be loosely described as gravity transforming into a more elliptical shape the image of galaxies. Of course the measurement of this shear is subject to an experimental error, for instance because of the unknown intrinsic ellipticity of the observed galaxy. Likewise, the weak gravitational lensing may produce an alignment in the inclination of nearby observations, but again this could be brought in by random fluctuations. We refer to [7] for much more detailed discussion on motivations and related challenges, which currently involve huge amount of physicists; major satellite experiments are at the planning stage, such as Euclid, see for instance http://hetdex.org/other_projects/euclid.php. To model the above discussed framework, we introduce random directions of observations $\left\{X_i \in \mathbb{S}^2\right\}$, which we take to be uniformly sampled over the sky, and observational errors $\left\{\varepsilon_{i;s}\right\}$, $i = 1, 2, \ldots, n$; the latter are independent and identically distributed spin s random variables, which we assume to be invariant in law with respect to rotations in the tangent plane:

$$\varepsilon'_{i;s} \overset{d}{=} \varepsilon_{i;s} e^{is\psi}, \quad \text{for all } \psi \in [0, 2\pi], \ i = 1, 2, \ldots, n, \tag{9}$$

$\overset{d}{=}$ denoting equality in law. As in [6], (9) implies that

$$\mathrm{Re}\, \varepsilon_{i;s} \overset{d}{=} \mathrm{Im}\, \varepsilon_{i;s} \overset{d}{=} \widetilde{\varepsilon}_i. \tag{10}$$

From (9) and (10) we have immediately

$$\begin{aligned} &E\left[\varepsilon_{i;s}\right] = E\left[\mathrm{Re}\, \varepsilon_{i;s} + i \mathrm{Im}\, \varepsilon_{i;s}\right] = 0, \\ &Var\left(\varepsilon_{i;s}\right) = E\left|\varepsilon_{i;s}\right|^2 = 2E\widetilde{\varepsilon}_i^2 =: \sigma^2_\varepsilon. \end{aligned}$$

Moreover, we shall assume that $\{\widetilde{\varepsilon}_i\}$ follows a sub-Gaussian distribution (Ref. [9]), i.e. there exists a number $a \geq 0$ such that for all $\lambda \in \mathbb{R}$ the following inequality holds:

$$E\left[e^{\lambda\widetilde{\varepsilon}_i}\right] \leq e^{\left(\frac{a^2\lambda^2}{2}\right)}. \tag{11}$$

We also define the *sub-Gaussian standard*of the random variable $\widetilde{\varepsilon}_i$ as:

$$\tau\left(\widetilde{\varepsilon}_i\right) = \inf\left\{a \geq 0 : E\left[e^{\lambda\widetilde{\varepsilon}_i}\right] \leq e^{\left(\frac{a^2\lambda^2}{2}\right)}, \lambda \in \mathbb{R}\right\} < \infty.$$

It is immediate to check (see [9]) that:

$$\tau\left(\widetilde{\varepsilon}_i\right) = \sup_{\lambda\neq 0}\left[\frac{2\log\left(E\left[e^{\lambda\widetilde{\varepsilon}_i}\right]\right)}{\lambda^2}\right]^{\frac{1}{2}}, \qquad E\left[e^{\lambda X}\right] \leq e^{\frac{\lambda^2\tau^2(\widetilde{\varepsilon}_i)}{2}}.$$

As is well known, a random variable is sub-Gaussian if and only if the moment generating function is majorized by the moment generating function of a zero-mean Gaussian random variable, whence the name *sub-Gaussian*. Indeed, the class of sub-Gaussian random variables contains, apart from the Gaussian themselves, all bounded zero-mean random variables and, more generally, all those random variables whose distribution tails decrease no slower than the tails of the Gaussian. We recall the following, simple results, whose proofs are available in [9]:

Lemma 2 (*Moment Characterization for Sub-Gaussian Random Variables*)**.** *Let $\tilde{\varepsilon}$ be a sub-Gaussian random variable such that $E\left(\tilde{\varepsilon}\right) = 0$. We have that $E\left(\left(\tilde{\varepsilon}\right)^2\right) \leq \tau\left(\tilde{\varepsilon}\right)$ and for all $p > 0$ $E\left(|\tilde{\varepsilon}|^p\right) < \infty$.*

In view of Lemma 2, sub-Gaussian random variables enjoy the same moment inequalities and concentration properties as Gaussian or bounded ones, and hence allow the implementation of the main technical tools in the proofs of our asymptotic results to follow. In this sense, they seem to provide a natural general framework for the analysis we must pursue.

5.2. *The estimation procedure*

The procedure we are going to investigate can be viewed as a form of needlet thresholding in the spin fiber bundles case (we refer to [4] for a similar approach, in the case of density estimation for standard scalar directional data). As discussed in the Introduction, we have now two alternative forms of needlets construction for the spin case, i.e. the pure spin needlets of [21] and the mixed spin needlets of [22]. Our approach could be implemented for both techniques, and indeed the proofs would be nearly identical. For definiteness, we shall focus on the mixed needlets constructions, which yields coefficients which are standard, complex-valued variables. For brevity's sake, however, we drop the subscript $\mathcal{M}$. We start by defining, as usual, an unbiased estimator for needlet coefficients. More precisely, we define

$$\widehat{\beta}_{jk;s} := \frac{1}{n}\sum_{i=1}^{n} Y_i\overline{\psi}_{jk;s}\left(X_i\right), \quad i = 1, 2, \ldots, n.$$

We have immediately:

$$\begin{aligned} E\left(\widehat{\beta}_{jk;s}\right) &= \frac{1}{n}\sum_{i=1}^{n} E\left[\overline{\psi}_{jk;s}\left(X_i\right)F_s\left(X_i\right) + \overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right] \\ &= \int_{\mathbb{S}^2}\overline{\psi}_{jk;s}\left(X_i\right)F_s\left(X_i\right) = \beta_{jk;s}. \end{aligned} \tag{12}$$

Moreover

$$\begin{aligned} Var\left(\widehat{\beta}_{jk;s}\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n}\overline{\psi}_{jk;s}\left(X_i\right)F_s(X_i) + \frac{1}{n}\sum_{i=1}^{n}\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} Var\left(\overline{\psi}_{jk;s}\left(X_i\right)F_s(X_i)\right) + \frac{1}{n^2}\sum_{i=1}^{n} Var\left(\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right). \end{aligned} \tag{13}$$

Now

$$\frac{1}{n^2}\sum_{i=1}^{n} Var\left(\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right) = \frac{1}{n}\sigma_\varepsilon^2\left\|\psi_{jk;s}\right\|^2_{L^2_s(\mathbb{S}^2)} = \frac{1}{n}\sigma_\varepsilon^2\tau_j^2 =: \frac{1}{n}\sigma^2_{1\varepsilon,j}$$

where in the last equality we used the independence of the $\varepsilon_{i;s}$. Note that obviously $\sigma_{1\varepsilon,j}^2 \leq \sigma_\varepsilon^2$. Also

$$0 \leq \frac{1}{n^2}\sum_{i=1}^{n} Var\left(\overline{\psi}_{jk;s}(X_i)\, F_s(X_i)\right) = \frac{1}{n}\int_{\mathbb{S}^2}\left|\overline{\psi}_{jk;s}(x)\, F_s(x)\right|^2 dx \leq \frac{M^2}{n},$$

and we define $\sigma_{\varepsilon,j}^2 := \sigma_{1\varepsilon,j}^2 + \frac{M^2}{n}$. We then proceed with the (now classical) hard thresholding procedure (see for instance [16,33,13]). In particular, we fix the threshold as

$$\kappa t_n = \kappa\sqrt{\frac{\log n}{n}}, \tag{14}$$

where κ is a real positive constant, whose value will be discussed later. Hence we define as usual

$$\beta_{jk;s}^{*} = w_{jk}\widehat{\beta}_{jk;s}, \quad w_{jk} = \mathbb{I}_{\{|\widehat{\beta}_{jk;s}|>\kappa t_n\}}, \tag{15}$$

where $\mathbb{I}_A$ denotes as usual the indicator function of the set A. The thresholding estimator is hence

$$F_s^{*}(x) = \sum_{j=1}^{J_n}\sum_{k=1}^{N_j}\beta_{jk;s}^{*}\psi_{jk;s}(x). \tag{16}$$

In (16), J_n represents a cut-off frequency, which we shall fix at $B^{J_n} = \sqrt{\frac{n}{\log n}}$, whereas N_j is the cardinality of the cubature point set at frequency j; it is known (see for instance [3]) that there exist positive constants c_1, c_2 such that $c_1 B^{2j} \leq N_j \leq c_2 B^{2j}$ (written $N^j \approx B^{2j}$). Our main result is to show that thresholding estimates achieve 'nearly optimal' (up to logarithmic factors) rates with respect to general $L_s^p\left(\mathbb{S}^2\right)$ loss functions.

Theorem 3. *Let $F_s \in \mathcal{B}_{\pi q;s}^{r}(G)$, the Besov ball such that $\|F_s\|_{\mathcal{B}_{\pi q;s}^{r}} \leq G < \infty$, $r - \frac{2}{\pi} > 0$, and consider F_s^* defined by (16), (14) and (15). For $1 \leq p < \infty$, there exist $\kappa > 0$ such that we have*

$$\sup_{F_s\in\mathcal{B}_{\pi q;s}^{r}} E\left\|F_s^{*} - F_s\right\|_{L_s^p}^{p} \leq C_p\{\log n\}^p\left[\frac{n}{\log n}\right]^{-\alpha(r,\pi,p)},$$

$$\alpha(r,\pi,p) = \begin{cases} \dfrac{rp}{2r+2} & \text{for } \pi \geq \dfrac{2p}{2r+2} \text{ (regular zone)} \\[2ex] \dfrac{p\left(r - 2\left(\frac{1}{\pi}-\frac{1}{p}\right)\right)}{2\left(r-2\left(\frac{1}{\pi}-\frac{1}{2}\right)\right)} & \text{for } \pi \leq \dfrac{2p}{2r+2} \text{ (sparse zone).}\end{cases}$$

Also, for $p = \infty$

$$\sup_{F_s\in\mathcal{B}_{\pi q;s}^{r}} E\left\|F_s^{*} - F_s\right\|_{L_s^\infty} \leq C_\infty\left[\frac{n}{\log n}\right]^{-\alpha(r,\pi,\infty)}, \quad \alpha(r,\pi,\infty) = \frac{\left(r-\frac{2}{\pi}\right)}{2\left(r-2\left(\frac{1}{\pi}-\frac{1}{2}\right)\right)}.$$

Remark 4. The definitions of regular and sparse zones are classical, and so are the rates we obtained, which indeed correspond (for instance) to those presented by [4]. For brevity's sake, we do not prove that these rates are indeed minimax (up to logarithmic terms), but it seems easy to achieve this goal by application of classical arguments, as for instance presented by [37]. It is trivial to note that for $\pi = \frac{2p}{2r+2} = \frac{p}{r+1}$ we have

$$\frac{\left(r-2\left(\frac{1}{\pi}-\frac{1}{p}\right)\right)}{2\left(r-2\left(\frac{1}{\pi}-\frac{1}{2}\right)\right)} = \frac{\left(r-2\left(\frac{r+1}{p}-\frac{1}{p}\right)\right)}{2\left(r-2\left(\frac{r+1}{p}-\frac{1}{2}\right)\right)} = \frac{r(p-2)}{2r(p-2)+2(p-2)}$$
$$= \frac{r}{2r+2},$$

and also

$$\frac{rp}{2r+2} \geq \frac{p\left(r-2\left(\frac{1}{\pi}-\frac{1}{p}\right)\right)}{2\left(r-2\left(\frac{1}{\pi}-\frac{1}{2}\right)\right)} \quad \text{in the regular zone,}$$
$$\frac{rp}{2r+2} \leq \frac{p\left(r-2\left(\frac{1}{\pi}-\frac{1}{p}\right)\right)}{2\left(r-2\left(\frac{1}{\pi}-\frac{1}{2}\right)\right)} \quad \text{in the sparse zone.}$$

Of course $\alpha(r,\pi,p) < \frac{1}{2}$, $\lim_{r\to\infty}\alpha(r,\pi,p) = \frac{1}{2}$.

Remark 5. For $s = 0$, our results cover adaptive nonparametric regression for complex-valued, scalar functions. Again, the rates correspond to the usual nearly minimax bounds.

The Proof of Theorem 3 is provided in the section to follow.

6. Proofs

Our arguments will follow closely classical approaches in this area, as presented for instance by [37], see also [4].

6.1. An auxiliary result

We shall need, in what follows, some sharp bounds which are provided in the following result. The arguments are close, for instance, to those for the inequality (65) on page 1088 of [37] where the case of a scalar Gaussian noise is considered: see also Proposition 15 in [4].

Proposition 6. *Let $\{\varepsilon_{i;s}\}$ be such that (9) and (11) are fulfilled. Assume also that $M := \|F_s\|_\infty < \infty$. For all $\gamma > 0$ and for all j such that $B^j \leq \sqrt{n/\log n}$, there exists $\kappa_\gamma > 0$ such that for $\kappa > \kappa_\gamma$ the following inequality holds:*

$$\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\overline{\psi}_{jk;s}\varepsilon_{i;s}\right| > \kappa\sqrt{\frac{\log n}{n}}\right) \leq Cn^{-\gamma}. \tag{17}$$

where $\gamma \approx \kappa^{4/3}$. Moreover, for all $p > 0$ we have

$$E\left[\left|\widehat{\beta}_{jk;s} - \beta_{jk;s}\right|^p\right] \leq C_p n^{-\frac{p}{2}} \tag{18}$$

$$E\left[\sup_k \left|\widehat{\beta}_{jk;s} - \beta_{jk;s}\right|^p\right] \leq C_\infty (j+1)^{p-1} n^{-p/2}. \tag{19}$$

Remark 7. It is possible to obtain sharp analytic expressions for κ, C_p, C_∞, for instance by arguing as in Lemma 16 of [4].

Proof. Note first that

$$\widehat{\beta}_{jk;s} = \frac{1}{n}\sum_{i=1}^{n}\overline{\psi}_{jk;s}\left(F_s\left(X_i\right) + \varepsilon_{i;s}\right) \tag{20}$$

$$\beta_{jk;s} = E\left(\widehat{\beta}_{jk;s}\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(\overline{\psi}_{jk;s}\left(X_i\right) F_s\left(X_i\right)\right), \tag{21}$$

and

$$\begin{aligned}\widehat{\beta}_{jk;s} - \beta_{jk;s} &= \frac{1}{n}\sum_{i=1}^{n}\left\{\left(\overline{\psi}_{jk;s}F_s\left(X_i\right)\right) - E\left(\overline{\psi}_{jk;s}F_s\left(X_i\right)\right)\right\} + \frac{1}{n}\sum_{i=1}^{n}\overline{\psi}_{jk;s}\varepsilon_{i;s}\\ &= \frac{1}{n}\sum_{i=1}^{n}\Psi_{jk;s}(X_i) + \frac{1}{n}\sum_{i=1}^{n}\overline{\psi}_{jk;s}\varepsilon_{i;s}\end{aligned}$$

where

$$\Psi_{jk;s}(X_i) := \overline{\psi}_{jk;s}(X_i)F_s\left(X_i\right) - E\left(\overline{\psi}_{jk;s}(X_i)F_s\left(X_i\right)\right).$$

Consider $\mathbb{P}_\beta\left(x\right) := \mathbb{P}\left(\left|\widehat{\beta}_{jk;s} - \beta_{jk;s}\right| > x\right)$:

$$\mathbb{P}_\beta\left(x\right) \leq \mathbb{P}_F\left(x\right) + \mathbb{P}_\varepsilon\left(x\right) \tag{22}$$

where:

$$\mathbb{P}_F\left(x\right) = \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\Psi_{jk;s}(X_i)\right| > \frac{1}{2}x\right),$$

$$\mathbb{P}_\varepsilon\left(x\right) = \mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^{n}\overline{\psi}_{jk;s}\varepsilon_{i;s}\right| > \frac{1}{2}x\right).$$

As before, we can split these sums into a real and imaginary part, to which we can apply separately the following procedures for both real and imaginary part in $\mathbb{P}_F\left(x\right)$ and $\mathbb{P}_\varepsilon\left(x\right)$, that give the same results.

As far as $\mathbb{P}_F(x)$ is concerned, we use the fact that $\Psi_{jk;s}(X_i)$ are i.i.d random variables such that for each of them:

$$\sup\left|\Psi_{jk;s}(X_i)\right| \leq 2cMB^j$$
$$E\left(\left|\Psi_{jk;s}(X_i)\right|^2\right) \leq E\left(\left|\overline{\psi}_{jk;s}(X_i)\,F_s(X_i)\right|^2\right) \leq M^2\left\|\psi_{jk;s}\right\|^2_{L^2_s(\mathbb{S}^2)} \leq M^2.$$

We therefore apply Bernstein inequality, to obtain:

$$\mathbb{P}_F(x) \leq 4\exp\left(-\frac{n\frac{x^2}{4}}{2\left(M^2+\frac{1}{3}cMB^jx\right)}\right), \tag{23}$$

where the value 4 takes on count both real and imaginary parts.

Fixing $x=\kappa t_n$, the following result is obtained:

$$\mathbb{P}_F(\kappa t_n) \leq 4\exp\left(-\frac{n\left((k/2)\sqrt{\log n/n}\right)^2}{\frac{2}{3}\left(3M^2+cMB^jk\sqrt{\frac{\log n}{n}}\right)}\right),$$

and by choosing j such that $B^j \leq \sqrt{\frac{n}{\log n}}$

$$\mathbb{P}_F(\kappa t_n) \leq 4\exp\left(-\frac{3k^2\log n}{8M\left(3M+ck\right)}\right) = 2n^{-\frac{3k^2}{8M(3M+ck)}}. \tag{24}$$

As far as $\mathbb{P}_\varepsilon(x)$ is concerned, consider that conditionally on $\left(X'_1,\ldots,X'_n\right)$, $\frac{1}{n}\sum_i\overline{\psi}_{jk;s}\left(X'_i\right)\varepsilon_{i;s}$ is a complex-valued sub-Gaussian variable with mean 0 and variance $\frac{1}{n^2}\sum_{i=1}^n\left|\psi_{jk;s}(X_i)\right|^2\sigma_\varepsilon^2$. Therefore, by using the Markov's inequality, we obtain:

$$\mathbb{P}_\varepsilon(x) \leq E\left(\exp\left(-\frac{-nx^2}{\sigma_\varepsilon^2\frac{8}{n}\sum\limits_{i=1}^n\left|\psi_{jk;s}\right|^2}\right)\middle| X'_1,\ldots,X'_n\right).$$

Observe that $\left|\psi_{jk;s}\left(X'_i\right)\right|^2$ are i.i.d. variables bounded by CB^{2j}, such that $E\left(\left|\psi_{jk;s}\left(X'_i\right)\right|^2\right)=\int_{\mathbb{S}^2}\left|\psi_{jk;s}\right|^2\mathrm{d}x=\left\|\psi_{jk;s}\right\|^2_{L^2_s(\mathbb{S}^2)}\leq 1$. Therefore we split the denominator into 2 terms, using

$$\mathbb{I}_{\left\{\left|\frac{1}{n}\sum\limits_{i=1}^n\left|\psi_{jk;s}(X'_i)\right|^2-\left\|\psi_{jk;s}\right\|^2_{L^2_s(\mathbb{S}^2)}\right|<\alpha\right\}} \quad\text{and}\quad \mathbb{I}_{\left\{\left|\frac{1}{n}\sum\limits_{i=1}^n\left|\psi_{jk;s}(X'_i)\right|^2-\left\|\psi_{jk;s}\right\|^2_{L^2_s(\mathbb{S}^2)}\right|\geq\alpha\right\}},\quad \alpha>0,$$

$$\mathbb{P}_\varepsilon(x) \leq \exp\left(-\frac{nx^2}{8\sigma_\varepsilon^2(1+\alpha)}\right)+\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^n\left|\psi_{jk;s}\right|^2-\left\|\psi_{jk;s}\right\|_2^2\right|>\alpha\right).$$

Now, by fixing $x=\kappa t_n$, we obtain the following result:

$$\mathbb{P}_\varepsilon(\kappa t_n) \leq \exp\left(-\frac{k^2\log n}{8\sigma_\varepsilon^2(\alpha+1)}\right)+\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n\left|\psi_{jk;s}\left(X'_i\right)\right|^2-\left\|\psi_{jk;s}\right\|^2_{L^2_s(\mathbb{S}^2)}\right|\geq\alpha\right).$$

Now, we use on the second term the Hoeffding's inequality:

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^n\left|\psi_{jk;s}\left(X'_i\right)\right|^2-\left\|\psi_{jk;s}\right\|^2_{L^2_s(\mathbb{S}^2)}\right|\geq\alpha\right) \leq 2\exp\left\{-\frac{2n^2\alpha^2}{ncB^{2j}}\right\}.$$

Again, because $B^{2j}\leq\frac{n}{\log n}$, we obtain:

$$\begin{aligned}\mathbb{P}_\varepsilon(\kappa t_n) &\leq 2\left\{\exp\left(-\frac{2\alpha^2\log n}{c}\right)+\exp\left(-\frac{k^2\log n}{8\sigma_\varepsilon^2(\alpha+1)}\right)\right\}\\ &= 2\left\{n^{-\frac{2\alpha^2}{c}}+n^{-\frac{k^2}{8\sigma_\varepsilon^2(\alpha+1)}}\right\}.\end{aligned} \tag{25}$$

We fix $\alpha \sim k^{\frac{2}{3}}$ in order to obtain the same order of magnitude between the two terms, and by using (24) and (25) finally we obtain:

$$\mathbb{P}_{\varepsilon}, \mathbb{P}_F \leq C \cdot n^{-ck^{4/3}}.$$

In order to prove (18), we use again (20), to obtain:

$$\begin{aligned} E\left[\left|\widehat{\beta}_{jk;s} - \beta_{jk;s}\right|^p\right] &\leq 2^{p-1}\left(E\left[\left|\frac{1}{n}\sum_{i=1}^{n}\left(\overline{\psi}_{jk;s}F_s\left(X_i\right)\right) - E\left(\overline{\psi}_{jk;s}F_s\left(X_i\right)\right)\right|^p\right] + E\left[\left|\frac{1}{n}\sum_{i=1}^{n}\overline{\psi}_{jk;s}\varepsilon_{i;s}\right|^p\right]\right) \\ &= 2^{p-1}\left(E_F + E_{\varepsilon}\right). \end{aligned}$$

We need to split again both E_F and E_ε into real and imaginary parts. Note that

$$\begin{aligned} E_F &= E\left[\left|\frac{1}{n}\sum_{i=1}^{n}\mathrm{Re}\,\Psi_{jk;s}(X_i) + \mathrm{Im}\,\Psi_{jk;s}(X_i)\right|^p\right] \\ &\leq 2^{p-1}\left(E\left(\left|\frac{1}{n}\sum_{i=1}^{n}\mathrm{Re}\,\Psi_{jk;s}(X_i)\right|^p\right) + E\left(\left|\frac{1}{n}\sum_{i=1}^{n}\mathrm{Im}\,\Psi_{jk;s}(X_i)\right|^p\right)\right) \\ &\leq 2^{p-1}\left(E_F^1 + E_F^2\right) \end{aligned}$$

and

$$\begin{aligned} E_\varepsilon &= E\left[\left|\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{Re}\left\{\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right\} + \mathrm{Im}\left\{\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right\}\right)\right|^p\right] \\ &\leq 2^{p-1}\left(E\left(\left|\frac{1}{n}\sum_{i=1}^{n}\mathrm{Re}\left\{\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right\}\right|^p\right) + E\left(\left|\frac{1}{n}\sum_{i=1}^{n}\mathrm{Im}\left\{\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right\}\right|^p\right)\right) \\ &\leq 2^{p-1}\left(E_\varepsilon^1 + E_\varepsilon^2\right). \end{aligned} \tag{26}$$

If for $0 < p \leq 2$, we apply the classical convexity inequality, in the case $2 < p < \infty$, we obtain a very similar result by applying the Rosenthal inequality to each term in (26) to obtain:

$$E_F^1 \leq C_p\left(\frac{E\left(\left|\mathrm{Re}\,\Psi_{jk;s}(X_i)\right|^p\right)}{n^{p-1}} + \frac{\left(E\left(\left|\mathrm{Re}\,\Psi_{jk;s}(X_i)\right|^2\right)\right)^{\frac{p}{2}}}{n^{\frac{p}{2}}}\right), \tag{27}$$

and similar results for E_F^2, E_ε^1 and E_ε^2. Recalling that $B^j \leq \sqrt{\frac{n}{\log n}} \leq \sqrt{n}$, we obtain:

$$\begin{aligned} E\left(\mathrm{Re}\left|\Psi_{jk;s}(X_i)\right|^p\right) &= E\left(\mathrm{Im}\left|\Psi_{jk;s}(X_i)\right|^p\right) \\ &\leq E\left(\left|\overline{\psi}_{jk;s}\left(X_i\right)F_s\left(X_i\right)\right|^p\right) \leq \int_{\mathbb{S}^2}\left|\overline{\psi}_{jk;s}\left(X_i\right)F_s\left(X_i\right)\right|^p \mathrm{d}x \\ &\leq cM^pB^{j(p-2)} \leq cM^p n^{-\frac{p-2}{2}}. \end{aligned}$$

As far as the noise-related terms, we obtain:

$$\begin{aligned} E\left(\left|\mathrm{Re}\left(\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right)\right|^p\right) &= E\left(\left|\mathrm{Im}\left(\overline{\psi}_{jk;s}\left(X_i\right)\varepsilon_{i;s}\right)\right|^p\right) \\ &\leq E\left(\left|\varepsilon_{i;s}\right|^p\right)cB^{j(p-2)} \leq cn^{-\frac{p-2}{2}}. \end{aligned}$$

Then, by substituting the last inequalities in (27) and the correspondent inequalities for E_F^2, E_ε^1 and E_ε^2, we obtain:

$$\frac{n^{\frac{p}{2}-1}}{n^{p-1}} = n^{-\frac{p}{2}}.$$

Now we study the case $p = \infty$: in order to prove (19), we majorize:

$$E\left[\sup_k\left|\widehat{\beta}_{jk;s} - \beta_{jk;s}\right|\right] \leq \int_{\mathbb{R}^+} x^{p-1}\mathbb{P}\left(\sup_k\left|\widehat{\beta}_{jk;s} - \beta_{jk;s}\right|^p > x\right)\mathrm{d}x. \tag{28}$$

Recalling the procedure used in the proof of (17), for $B^j \leq \sqrt{n}$, (23) becomes:

$$\mathbb{P}_F(x) \leq 4\left(\exp\left(-\frac{nx^2}{16M^2}\right) + \exp\left(-\frac{3\sqrt{n}x}{16cM}\right)\right), \tag{29}$$

while, in a similar way, we split the first term on (25) as:

$$\exp\left(-\frac{nx^2}{8\sigma_\varepsilon^2(1+\alpha)}\right) \leq \exp\left(-\frac{nx^2}{16\sigma_\varepsilon^2}\right) + \exp\left(-\frac{nx^2}{16\sigma_\varepsilon^2\alpha}\right) = \mathbb{P}_\varepsilon^*(x) + \mathbb{P}_{\varepsilon,\alpha}^1. \tag{30}$$

By applying on the last term of (25) the Hoeffding inequality and for $B^j \leq \sqrt{n}$, we obtain:

$$\begin{aligned}\mathbb{P}\left(\frac{1}{n}\left|\sum_{i=1}^n \left|\psi_{jk;s}\right|^2 - \left\|\psi_{jk;s}\right\|_2^2\right| > \alpha\right) &\leq \exp\left(-\frac{2n^2\alpha^2}{ncB^{2j}}\right)\\ &\leq \exp\left(-\frac{2\alpha^2}{c}\right) = \mathbb{P}_{\varepsilon,\alpha}^2.\end{aligned}$$

We choose $\alpha = \left\{\frac{c^{1/3}}{32^{1/3}\sigma_\varepsilon^{2/3}} \cdot n^{1/3}x^{2/3}\right\}$, to obtain

$$\mathbb{P}_{\varepsilon,\alpha}^1 + \mathbb{P}_{\varepsilon,\alpha}^2 \leq C\exp\left(-\frac{n^{2/3}x^{4/3}}{2^{7/3}\sigma_\varepsilon^{4/3}c^{1/3}}\right), \tag{31}$$

and in view of (22), (29), (31) and (30)

$$\mathbb{P}_\beta(x) \leq C\left(\exp\left(-\frac{nx^2}{16\sigma_\varepsilon^2}\right) + \exp\left(-\frac{nx^2}{16M^2}\right) + \exp\left(-\frac{2\sqrt{n}x}{16cM}\right) + \exp\left(-\frac{n^{2/3}x^{4/3}}{2^{7/3}\sigma_\varepsilon^{4/3}c^{1/3}}\right)\right).$$

Now we fix a parameter $a = \max\left(4\sqrt{2}\sigma_\varepsilon, 4\sqrt{2}M, \frac{32}{3}cM, 2^{11/4}\sigma_\varepsilon c^{1/4}\right)$. Write (28) as:

$$\begin{aligned}E\left[\sup_k\left|\widehat{\beta}_{jk;s} - \beta_{jk;s}\right|^p\right] &\leq \int_{0\leq x\leq \frac{aj}{\sqrt{n}}} x^{p-1}\mathrm{d}x + 2c\int_{x>\frac{aj}{\sqrt{n}}} Cx^{p-1}B^j\\ &\quad\times\left[\exp\left(\frac{nx^2}{16\sigma_\varepsilon^2}\right) + \exp\left(-\frac{nx^2}{16M^2}\right) + \exp\left(-\frac{2\sqrt{n}x}{16cM}\right) + \exp\left(-\left(\frac{nx^2}{2^7\sigma_\varepsilon^2c^{\frac{1}{2}}}\right)^{\frac{2}{3}}\right)\right]\\ &= E_\infty^1 + E_\infty^2 + E_\infty^3.\end{aligned} \tag{32}$$

We observe that for each term depending on $\exp\left(-nx^2/C\right)$, where $C = 4\sqrt{2}\sigma_\varepsilon, 4\sqrt{2}M$ and for $x > aj/\sqrt{n}$, we have:

$$B^j\exp\left(-\frac{nx^2}{C}\right) \leq \exp\left(-\frac{nx^2}{2C} - \frac{nx^2}{2C} + j\right) \leq \exp\left(-\frac{nx^2}{2C}\right).$$

Similarly, we have for $x > aj/\sqrt{n}$:

$$B^j\exp\left(-\frac{2\sqrt{n}x}{16cM}\right) \leq \exp\left(-\frac{2\sqrt{n}x}{32cM}\right),$$

and finally, again for $x > aj/\sqrt{n}$

$$B^j\exp\left(-\frac{n^{2/3}x^{4/3}}{2^7/3\sigma_\varepsilon^{4/3}}\right) \leq \exp\left(-\frac{n^{2/3}x^{4/3}}{2^{10^3}\sigma_\varepsilon^{4/3}}\right).$$

Likewise, the integral E_∞^1 is simply majorized by:

$$E_\infty^1 \leq C\frac{1}{p}\left(\frac{j}{\sqrt{n}}\right)^p \leq C_pj^pn^{-p/2}. \tag{33}$$

As far as E_∞^2 is concerned, by using a change of variable $u = \sqrt{n}x$ we obtain:

$$E_\infty^2 \leq 2C\frac{1}{n^{-p/2}}\int_{u\geq aj} u^{p-1}\exp\left(-\frac{u^{4/3}}{2^{10/3}\sigma_\varepsilon^{4/3}c^{1/3}}\right)\mathrm{d}u \leq C_pn^{-p/2}. \tag{34}$$

A similar procedure is applied to E^3_∞ by using the same change of variable $u = \sqrt{n}x$ to obtain:

$$E^3_\infty \leq C'_p n^{-p/2}. \tag{35}$$

Finally by substituting (33)–(35) in (32) we obtain the thesis. □

6.2. Proof of Theorem 3

Again we follow closely standard arguments in the thresholding literature, as discussed for instance in [33]. More precisely, the proof can be customarily divided into different cases, as follows.

- Regular zone, $p < \infty$.

  We start as usual from

$$\begin{aligned} E\left\|F^*_s - F_s\right\|^p_{L^p_s} &= E\left\|\sum_{j\leq J_n}\sum_k w_{jk}\widehat{\beta}_{jk;s}\psi_{jk;s} - \sum_j\sum_k \beta_{jk;s}\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \\ &= E\left\|\sum_{j\leq J_n}\sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s} + \sum_{j>J_n}\sum_k \beta_{jk;s}\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \\ &\leq E\left\|\sum_{j\leq J_n}\sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} + \left\|\sum_{j>J_n}\sum_k \beta_{jk;s}\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \\ &=: I + II. \end{aligned}$$

For $p \leq \pi$, we have $\mathcal{B}^r_{\pi q;s} \subset \mathcal{B}^r_{pq;s}$, whence we can always take $\pi = p$ in this case; hence we focus on $p \geq \pi$. Here we have the embedding $\mathcal{B}^r_{\pi q;s} \subset \mathcal{B}^{r-\frac{2}{\pi}+\frac{2}{p}}_{pq;s}$, whence

$$\left\|\sum_{j>J_n}\sum_k \beta_{jk;s}\psi_{jk;s}\right\|_{L^p_s(\mathbb{S}^2)} = O\left(B^{-2j(\frac{r}{2}-\frac{1}{\pi}+\frac{1}{p})}\right) = O\left(\left\{\frac{n}{\log n}\right\}^{-(\frac{r}{2}-\frac{1}{\pi}+\frac{1}{p})}\right),$$

and because in the regular zone

$$r \geq \frac{2}{\pi}, \frac{r}{2r+2} = \frac{rp}{2(r+1)p} \leq \frac{r\pi}{2p},$$

we obtain

$$\left(\frac{r}{2} - \frac{1}{\pi} + \frac{1}{p}\right) - \frac{r}{2r+2} \geq \left(\frac{r}{2} - \frac{1}{\pi} + \frac{1}{p}\right) - \frac{r\pi}{2p} = \left(\frac{1}{\pi} - \frac{1}{p}\right)\left(\frac{r\pi}{2} - 1\right) > 0.$$

Hence the bias term is fixed. For the variance term we have

$$I \leq J_n^{p-1}\sum_{j\leq J_n} E\left\|\sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)}.$$

Now we split I in four zones; more precisely, we shall label A (respectively U) where the estimated coefficients is above (resp. under) the threshold κt_n, and a (respectively u) the regions where the deterministic coefficients are above or under a new threshold, which is $\frac{\kappa}{2}t_n$ in A and $2\kappa t_n$ in U. We hence obtain

$$\begin{aligned} &\sum_{j\leq J_n} E\left\|\sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \\ &= \sum_{j\leq J_n} E\left\|\sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \mathbb{I}_{\{|\widehat{\beta}_{jk;s}|\geq\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\geq\kappa t_n/2\}} \\ &\quad + \sum_{j\leq J_n} E\left\|\sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \mathbb{I}_{\{|\widehat{\beta}_{jk;s}|\geq\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\leq\kappa t_n/2\}} \\ &\quad + \sum_{j\leq J_n} E\left\|\sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \mathbb{I}_{\{|\widehat{\beta}_{jk;s}|<\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\geq 2\kappa t_n\}} \end{aligned}$$

$$
\begin{aligned}
&+ \sum_{j\le J_n} E\left\| \sum_k (w_{jk}\widehat{\beta}_{jk;s} - \beta_{jk;s})\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \mathbb{I}_{\{|\widehat{\beta}_{jk;s}|<\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\le 2\kappa t_n\}}.\\
&\le C\left\{ \sum_{j\le J_n}\sum_k \left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} E\left[\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p \mathbb{I}_{\{|\widehat{\beta}_{jk;s}|\ge\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}}\right]\right.\\
&+ \sum_{j\le J_n}\sum_k \left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} E\left[\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p \mathbb{I}_{\{|\widehat{\beta}_{jk;s}|\ge\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\le\kappa t_n/2\}}\right]\\
&+ \sum_{j\le J_n}\sum_k \left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \left|\beta_{jk;s}\right|^p E\left[\mathbb{I}_{\{|\widehat{\beta}_{jk;s}|<\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\ge 2\kappa t_n\}}\right]\\
&\left.+ \sum_{j\le J_n}\sum_k \left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} \left|\beta_{jk;s}\right|^p E\left[\mathbb{I}_{\{|\widehat{\beta}_{jk;s}|<\kappa t_n\}}\mathbb{I}_{\{|\beta_{jk;s}|\le 2\kappa t_n\}}\right]\right\}\\
&= Aa + Au + Ua + Uu.
\end{aligned}
$$

This argument is the same as in [4], where the regions are labeled instead Bb, Bs, Sb, Ss; we preferred to avoid B and b which have a different use in the present work. Heuristically, the cross/terms Au, Ua are easier to bound, as we can exploit quick decay of $\Pr\left\{\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right| > \frac{1}{2}t_n\right\}$; for Aa, Uu the crucial bounds will be derived by the tail behavior in the Besov balls $\mathcal{B}^r_{pq;s}(G)$.

Note firstly that

$$
\begin{aligned}
Aa &\le C\sum_{j\le J_n}\sum_k \left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)} E\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p \mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}}\\
&\le C\sum_{j\le J_n}\sum_k B^{j(p-2)}\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}} E\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p;
\end{aligned}
$$

now from (18) and (6) we know that

$$
E\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p \le C_p n^{-p/2}, \qquad \sum_k B^{j(p-2)} = O(B^{jp}).
$$

Write

$$
\begin{aligned}
&\sum_{j\le J_n}\sum_k B^{j(p-2)}\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}} E\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p\\
&\quad\le C\left\{ n^{-p/2}\sum_{j\le J_{1n}}\sum_k B^{j(p-2)}\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}} + n^{-p/2}\sum_{j>J_{1n}}\sum_k B^{j(p-2)}\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}}\right\}\\
&\quad\le C\left\{ n^{-p/2}B^{pJ_{1n}} + n^{-p/2}\sum_{j>J_{1n}}\sum_k B^{j(p-2)}\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}}\right\}.
\end{aligned}
$$

Fix

$$
B^{J_{1n}} = \kappa'\left\{\frac{n}{\log n}\right\}^{\frac{1}{2(r+1)}},
$$

and note that we have

$$
\begin{aligned}
\sum_{j>J_{1n}}\sum_k B^{j(p-2)}\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}} &\le \sum_{j>J_{1n}}\sum_k \left|\beta_{jk;s}\right|^p B^{j(p-2)}\left\{\kappa t_n/2\right\}^{-p}\\
&\le \left\{\frac{n}{\log n}\right\}^{p/2}\sum_{j>J_{1n}}\left\{\sum_k\left|\beta_{jk;s}\right|^p\left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)}\right\},
\end{aligned}
$$

where

$$
\left\{\sum_k\left|\beta_{jk;s}\right|^p\left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)}\right\} \le CB^{-prj},
$$

because by assumption $F_s \in \mathcal{B}^r_{pq;s}$. Hence

$$
\left\{\frac{n}{\log n}\right\}^{p/2}\sum_{j>J_{1n}}\left\{\sum_k\left|\beta_{jk;s}\right|^p\left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)}\right\} \le \left\{\frac{n}{\log n}\right\}^{p/2} B^{-prJ_{1n}} \le C\left\{\frac{n}{\log n}\right\}^{p/2}\left\{\frac{n}{\log n}\right\}^{-\frac{pr}{2(r+1)}}
$$

$$\leq C\left\{\frac{n}{\log n}\right\}^{\frac{p(r+1)-pr}{2(r+1)}} \leq C\left\{\frac{n}{\log n}\right\}^{\frac{p}{2(r+1)}} \leq CB^{pJ_{1n}},$$

and

$$C\sum_{j\leq J_n}\sum_k B^{j(p-2)}\mathbb{I}_{\{|\beta_{jk;s}|\geq\kappa t_n/2\}}E\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p \leq Cn^{-p/2}B^{pJ_{1n}} \leq C\left\{\frac{n}{\log n}\right\}^{\frac{p}{2(r+1)}} n^{-p/2} \leq C\left\{\frac{n}{\log n}\right\}^{\frac{-pr}{2(r+1)}}.$$

Hence the term Aa is fixed. For the term Uu, it suffices to observe that

$$\begin{aligned} Uu &\leq C\sum_{j\leq J_n}\sum_k \left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)}\left|\beta_{jk;s}\right|^p\mathbb{I}_{\{|\beta_{jk;s}|\leq 2\kappa t_n\}} \\ &\leq C\left\{\sum_{j\leq J_{1n}}\sum_k B^{j(p-2)}\left|2\kappa t_n\right|^p + \sum_{j>J_{1n}}\sum_k B^{j(p-2)}\left|\beta_{jk;s}\right|^p\right\} \\ &\leq C\left\{B^{pJ_{1n}}\left\{\frac{n}{\log n}\right\}^{-p/2} + B^{-prJ_{1n}}\right\} \\ &\leq C\left\{\left[\frac{n}{\log n}\right]^{\frac{p}{2(r+1)}}\left[\frac{n}{\log n}\right]^{-\frac{p}{2}} + \left[\frac{n}{\log n}\right]^{-\frac{pr}{2(r+1)}}\right\} = O\left(\left[\frac{n}{\log n}\right]^{-\frac{pr}{2(r+1)}}\right). \end{aligned}$$

Now note that

$$\begin{aligned} Au &\leq C\sum_{j\leq J_n}\sum_k B^{j(p-2)}E\left[\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^p\mathbb{I}_{\{|\widehat{\beta}_{jk;s}-\beta_{jk;s}|\geq\kappa t_n/2\}}\right] \\ &\leq \sum_{j\leq J_n}\sum_k B^{j(p-2)}\left\{E\left[\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|^{2p}\right]\right\}^{1/2}\left\{\mathbb{P}\left[\left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|\geq\kappa t_n/2\right]\right\}^{1/2} \end{aligned}$$

and using (18)

$$Au \leq Cn^{-p/2}B^{pJ_n}n^{-\gamma/2} \leq Cn^{-p/2}\left[\frac{n}{\log n}\right]^{p/2} n^{-\gamma/2} = C\left(\log n\right)^{-\frac{p}{2}} n^{-\gamma/2}.$$

Finally

$$Ua \leq \sum_{j\leq J_n}\sum_k \left\|\psi_{jk;s}\right\|^p_{L^p_s(\mathbb{S}^2)}\left|\beta_{jk;s}\right|^p E\left[\mathbb{I}_{\{|\widehat{\beta}_{jk;s}-\beta_{jk;s}|>\kappa t_n\}}\right] \leq Cn^{-\gamma}\left\|F_s\right\|^p_{L^p_s(\mathbb{S}^2)}.$$

Because obviously $n^{-\gamma} \leq n^{-\gamma/2}$ we have to choose γ such that:

$$n^{-\gamma/2} \leq n^{-\frac{pr}{2r+2}} \longrightarrow \gamma \geq \frac{pr}{r+1}.$$

We can hence take $\kappa \sim \gamma^{3/4}$, which yields

$$\kappa \geq C\left(\frac{pr}{r+1}\right)^{\frac{3}{4}}.$$

- The case $p=\infty$.

  Assume first that $F_s \in \mathcal{B}^r_{\infty,\infty;s}$. Then

$$\begin{aligned} E\left\|F^*_s - F_s\right\|_{L^\infty_s(\mathbb{S}^2)} &\leq E\left\|\sum_{j\leq J_n}\sum_k (w_{jk}\widehat{\beta}_{jk;s}-\beta_{jk;s})\psi_{jk;s}\right\|_{L^\infty_s(\mathbb{S}^2)} + \left\|\sum_{j>J_n}\sum_k \beta_{jk;s}\psi_{jk;s}\right\|_{L^\infty_s(\mathbb{S}^2)} \\ &= I + II. \end{aligned}$$

For II, it is sufficient to note that

$$\begin{aligned} \left\|\sum_{j>J_n}\sum_k \beta_{jk;s}\psi_{jk;s}\right\|_{L^\infty_s(\mathbb{S}^2)} &\leq \sum_{j>J_n}\left\|\sum_k \beta_{jk;s}\psi_{jk;s}\right\|_{L^\infty_s(\mathbb{S}^2)} = O\left(B^{-rJ_n}\right) \\ &= O\left(\left[\frac{n}{\log n}\right]^{-r/2}\right) = O\left(\left[\frac{n}{\log n}\right]^{-r/2(r+1)}\right). \end{aligned}$$

On the other hand,

$$E\left\|\sum_{j\le J_n}\sum_k (w_{jk}\widehat{\beta}_{jk;s}-\beta_{jk;s})\psi_{jk;s}\right\|_{L_s^\infty(\mathbb{S}^2)} \le \sum_{j\le J_n} E\left\|\sum_k (w_{jk}\widehat{\beta}_{jk;s}-\beta_{jk;s})\psi_{jk;s}\right\|_{L_s^\infty(\mathbb{S}^2)}$$
$$\le C\sum_{j\le J_n} B^j E\left[\sup_k \left|w_{jk}\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|\right]$$
$$\le C\sum_{j\le J_n} B^j E\left[\sup_k \left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|\right]\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}} + C\sum_{j\le J_n} B^j E\left[\sup_k \left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|\right]\mathbb{I}_{\{|\widehat{\beta}_{jk;s}-\beta_{jk;s}|\ge\kappa t_n/2\}}$$
$$+C\sum_{j\le J_n} B^j \sup_k \left|\beta_{jk;s}\right| E\left[\mathbb{I}_{\{|\widehat{\beta}_{jk;s}-\beta_{jk;s}|>\kappa t_n\}}\right] + C\sum_{j\le J_n} B^j \sup_k \left|\beta_{jk;s}\right| \mathbb{I}_{\{|\beta_{jk;s}|\le 2\kappa t_n\}}$$
$$= Aa + Au + Ua + Uu.$$

Now as before, we note that it is possible to choose

$$J_{1n} : B^{J_{1n}} \sim \kappa'\left\{\frac{n}{\log n}\right\}^{\frac{1}{2(r+1)}} \quad \text{and for} \quad j > J_{1n}, \quad \mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}} \equiv 0.$$

Hence, by (19)

$$Aa \le C\sum_{j\le J_n} B^j E\left[\sup_k \left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|\right]\mathbb{I}_{\{|\beta_{jk;s}|\ge\kappa t_n/2\}}$$
$$\le C\sum_{j\le J_{1n}} B^j E\left[\sup_k \left|\widehat{\beta}_{jk;s}-\beta_{jk;s}\right|\right] \le CJ_{1n}n^{-\frac{1}{2}}B^{J_{1n}}$$
$$\le CJ_{1n}(\log n)^{-1/2}\left\{\frac{n}{\log n}\right\}^{-\frac{r}{2(r+1)}}.$$

Also

$$\sum_{j\le J_n} B^j \sup_k \left|\beta_{jk}\right| \mathbb{I}_{\{|\beta_{jk;s}|\le 2\kappa t_n\}} \le C\left\{t_n B^{J_{1n}} + \sum_{J_{1n}\le j<\infty} B^j \sup_k \left|\beta_{jk}\right|\right\}$$
$$\le C\left\{t_n B^{J_{1n}} + \sum_{J_{1n}\le j<\infty} \|F_s\|_{L_s^\infty}\right\}$$
$$\le C\left\{t_n B^{J_{1n}} + B^{-J_{1n}}\right\} \le C\left\{\frac{n}{\log n}\right\}^{-\frac{r}{2(r+1)}}.$$

For the remaining two terms the arguments is the same, actually easier. For general π and q, it is sufficient to note that $\mathcal{B}^r_{\pi q;s} \subset \mathcal{B}^{r'}_{\infty,\infty;s}$, $r' = r - 2/\pi$. By the previous argument

$$E\left\|F_s^* - F_s\right\|_{L_s^\infty} \le CJ_n\left\{\frac{n}{\log n}\right\}^{-\frac{r'}{2(r'+1)}} = CJ_n\left\{\frac{n}{\log n}\right\}^{-\frac{r-2/\pi}{2(r-2(1/\pi-1/2))}}.$$

Note that for $\pi = p = \infty$ the sparse and regular zone coincide; otherwise for $p = \infty$ we are always in the sparse zone

- The sparse case.

The argument is again similar to the previous cases and to [37,4]. Indeed we have $\mathcal{B}^r_{\pi q;s} \subset \mathcal{B}^{r-2(\frac{1}{\pi}-\frac{1}{p})}_{p,q;s}$,

$$E\left\|F_s^* - F_s\right\|^p_{L_s^p} \le E\left\|\sum_{j\le J_n}\sum_k (w_{jk}\widehat{\beta}_{jk;s}-\beta_{jk;s})\psi_{jk;s}\right\|^p_{L_s^p(\mathbb{S}^2)} + \left\|\sum_{j>J_n}\sum_k \beta_{jk;s}\psi_{jk;s}\right\|^p_{L_s^p(\mathbb{S}^2)},$$

$$\left\|\sum_{j>J_n}\sum_k \beta_{jk;s}\psi_{jk;s}\right\|^p_{L_s^p(\mathbb{S}^2)} \le CB^{-J_n(r-2(\frac{1}{\pi}-\frac{1}{p}))} \le CB^{-2J_n[(r-2(\frac{1}{\pi}-\frac{1}{p}))/2(r-2(\frac{1}{\pi}-\frac{1}{p}))]}$$
$$\le \left\{\frac{n}{\log n}\right\}^{-[(r-2(\frac{1}{\pi}-\frac{1}{p}))/2(r-2(\frac{1}{\pi}-\frac{1}{p}))]},$$

because $r - \frac{2}{\pi} + 1 \geq 1$, given that $r - \frac{2}{\pi} \geq 0$ by assumption. Hence the bias term has the correct order. For the variance term, the trick is very much as above, and we omit some details. It is possible to split the term to be bounded into four terms, after which the two "cross terms" Au and Ua are easy because they involve quantities like $\mathbb{P}\left\{|\widehat{\beta}_{jk;s} - \beta_{jk;s}| > \kappa t_n\right\}$, which can be made smaller than $n^{-p/2}$ for all $p > 0$, given a suitable choice of κ. Fix J_{2n} such that

$$B^{J_{2n}} \approx \left[\frac{n}{\log n}\right]^{\frac{1}{2((r-\frac{2}{\pi})+1)}},$$

so that

$$\left[\frac{n}{\log n}\right]^{\frac{\pi-p}{2}} B^{J_{2n}(p-\pi(r+1))} \approx \left[\frac{n}{\log n}\right]^{\frac{\pi-p}{2}} \left[\frac{n}{\log n}\right]^{\frac{(p-\pi(r+1))}{2((r-\frac{2}{\pi})+1)}}$$
$$\approx \left[\frac{n}{\log n}\right]^{\frac{(\pi-p)((r-\frac{2}{\pi})+1)|+(p-\pi(r+1))}{2((r-\frac{2}{\pi})+1)}}.$$

For the terms of the form Aa and Uu we have

$$J_n^{p-1} n^{-p/2} \sum_{j \leq J_{1n}} B^{j(p-2)} \sum_k \mathbb{I}_{\{|\beta_{jk;s}| \geq \kappa t_n/2\}} + J_n^{p-1} \sum_j B^{j(p-2)} \sum_k |\beta_{jk;s}|^p \mathbb{I}_{\{|\beta_{jk;s}| \leq 2\kappa t_n\}},$$

where to obtain the first summand we have exploited the embedding $\mathcal{B}^r_{\pi q;s} \subset \mathcal{B}^{r-\frac{2}{\pi}}_{\infty,\infty;s}$, whence for $j \geq J_{2n}$ one has $\mathbb{I}_{\{|\beta_{jk;s}| \geq \kappa t_n/2\}} \equiv 0$. Now

$$n^{-p/2} \sum_{j \leq J_{2n}} B^{j(p-2)} \sum_k \mathbb{I}_{\{|\beta_{jk;s}| \geq \kappa t_n/2\}} \leq C n^{-p/2} \sum_{j \leq J_{2n}} B^{j(p-2)} \sum_k |\beta_{jk;s}|^\pi t_n^{-\pi}$$
$$\leq C n^{-p/2} t_n^{-\pi} \sum_{j \leq J_{2n}} B^{j(p-\pi)} \sum_k B^{j(\pi-2)} |\beta_{jk;s}|^\pi$$
$$\leq C n^{-p/2} t_n^{-\pi} \sum_{j \leq J_{2n}} B^{j(p-\pi)} B^{-r\pi j} \leq C \left[\frac{n}{\log n}\right]^{\frac{\pi-p}{2}} B^{J_{2n}(p-\pi(r+1))}.$$

Likewise

$$\sum_{j \leq J_{2n}} B^{j(p-2)} \sum_k |\beta_{jk;s}|^p \mathbb{I}_{\{|\beta_{jk;s}| \leq 2\kappa t_n\}} \leq C \sum_{j \leq J_{2n}} B^{j(p-2)} \sum_k |\beta_{jk;s}|^\pi t_n^{p-\pi}$$
$$\leq C \left[\frac{n}{\log n}\right]^{\frac{\pi-p}{2}} \sum_{j \leq J_{2n}} B^{j(p-\pi)} \sum_k B^{j(\pi-2)} |\beta_{jk;s}|^\pi$$
$$\leq C \left[\frac{n}{\log n}\right]^{\frac{\pi-p}{2}} B^{J_{2n}(p-\pi(r+1))}.$$

Now

$$\frac{(\pi-p)\left(\left(r-\frac{2}{\pi}\right)+1\right)+(p-\pi(r+1))}{2\left(\left(r-\frac{2}{\pi}\right)+1\right)} = \frac{\left(\pi(r+1)-2-pr+\frac{2p}{\pi}-p\right)+(p-\pi(r+1))}{2\left(\left(r-\frac{2}{\pi}\right)+1\right)}$$
$$= -\frac{2+pr-\frac{2p}{\pi}}{2\left(\left(r-\frac{2}{\pi}\right)+1\right)} = -\frac{p\left(r-2\left(\frac{1}{\pi}-\frac{1}{p}\right)\right)}{2\left(r-2\left(\frac{1}{\pi}-\frac{1}{2}\right)\right)},$$

that is, these terms have the right order. So we are only left with

$$\sum_{j \geq J_{2n}} B^{j(p-2)} \sum_k |\beta_{jk;s}|^p \mathbb{I}_{\{|\beta_{jk;s}| \leq 2\kappa t_n\}}.$$

Consider

$$m = \frac{p-2}{r-\frac{2}{\pi}+1};$$

note that

$$p-m = \frac{pr-\frac{2p}{\pi}+p-p+2}{r-\frac{2}{\pi}+1} = \frac{pr-\frac{2p}{\pi}+2}{r-\frac{2}{\pi}+1} > 0,$$

$$m-\pi = \frac{p-2}{r-\frac{2}{\pi}+1}-\pi = \frac{p-\pi(r+1)}{r-\frac{2}{\pi}+1} > 0,$$

because $p-\pi(r+1) > 0$ in the sparse zone. We have

$$\sum_{j\geq J_{2n}} B^{j(p-2)} \sum_k |\beta_{jk;s}|^p \mathbb{I}_{\{|\beta_{jk;s}|\leq 2\kappa t_n\}} \leq C \sum_{j\geq J_{2n}} B^{j(p-2)} \sum_k |\beta_{jk;s}|^m t_n^{p-m}$$
$$\leq C t_n^{p-m} \sum_{j\geq J_{2n}} B^{j(p-m)} \sum_k B^{j(m-2)} |\beta_{jk;s}|^m$$
$$\leq C t_n^{p-m} \sum_{j\geq J_{2n}} B^{j(p-m)} \sum_k \left\| \psi_{jk;s} \right\|^m_{L^m_s(\mathbb{S}^2)} |\beta_{jk;s}|^m. \quad (36)$$

Now, because $\mathcal{B}^r_{\pi q;s} \subset \mathcal{B}^{r-\frac{2}{\pi}+\frac{2}{m}}_{m,q;s}$,

$$\sum_k \left\| \psi_{jk;s} \right\|^m_{L^m_s(\mathbb{S}^2)} |\beta_{jk;s}|^m \leq C B^{-mj(r-\frac{2}{\pi}+\frac{2}{m})},$$

hence (36) is bounded by

$$C t_n^{p-m} \sum_{J_{2n}\leq j\leq J} B^{j(p-m-2)} B^{-j(r-\frac{2}{\pi}+\frac{2}{m})m} \leq C t_n^{p-m} \sum_{J_{2n}\leq j\leq J} B^{j\left[(p-m-2)-(r-\frac{2}{\pi}+\frac{2}{m})m\right]}.$$

Observe that

$$(p-m)-\left(r-\frac{2}{\pi}+\frac{2}{m}\right)m = \frac{pr-\frac{2p}{\pi}+2}{r-\frac{2}{\pi}+1}-\left(r-\frac{2}{\pi}\right)m-2$$
$$= \frac{pr-\frac{2p}{\pi}+2}{r-\frac{2}{\pi}+1}-\left(r-\frac{2}{\pi}\right)\frac{p-2}{r-\frac{2}{\pi}+1}-2$$
$$= \frac{2r+2\left(1-\frac{2}{\pi}\right)}{r-\frac{2}{\pi}+1}-2=0,$$

hence

$$\sum_{J_{2n}\leq j\leq J} B^{j(p-2)} \sum_k |\beta_{jk;s}|^p \mathbb{I}_{\{|\beta_{jk;s}|\leq 2\kappa t_n\}} \leq C J_n t_n^{p-m} \leq C \log n \left[\frac{n}{\log n}\right]^{-\frac{p(r-2(\frac{1}{\pi}-\frac{1}{p}))}{2(r-2(\frac{1}{\pi}-\frac{1}{2}))}}.$$

Thus the proof is completed. □

References

[1] R.J. Adler, J.E. Taylor, Random Fields and Geometry, Springer, 2007.

[2] P. Baldi, G. Kerkyacharian, D. Marinucci, D. Picard, Asymptotics for spherical needlets, Annals of Statistics 37 (3) (2009) 1150–1171. arXiv:math.st/0606599.

[3] P. Baldi, G. Kerkyacharian, D. Marinucci, D. Picard, Subsampling needlet coefficients on the sphere, Bernoulli 15 (2009) 438–463. arXiv:0706.4169.

[4] P. Baldi, G. Kerkyacharian, D. Marinucci, D. Picard, Adaptive density estimation for directional data using needlets, Annals of Statistics 37 (6A) (2009) 3362–3395. arXiv:0807.5059.

[5] P. Baldi, G. Kerkyacharian, D. Marinucci, D. Picard, Besov spaces for sections of spin fiber bundles on the sphere, preprint, 2009.

[6] P. Baldi, D. Marinucci, Some characterizations of the spherical harmonics coefficients for isotropic random fields, Statistics and Probability Letters 77 (2007) 490–496. arXiv:math/0606709.

[7] S. Bridles, et al., Handbook for the GREAT08 challenge: an image analysis competition for gravitational lensing, Annals of Applied Statistics 2 (2009) 6–37.

[8] L. Brown, T. Cai, H. Zhou, Nonparametric regression in exponential families, Annals of Statistics 38 (4) (2010) 2005–2046.

[9] V.V. Buldygin, Y.V. Kozachenko, Metric Characterization of Random Variables and Random Processes, American Mathematical Society, 2000.

[10] P. Cabella, M. Kamionkowski, Theory of cosmic microwave background polarization, 18 Mar 2005, arXiv:astro-ph/0403392v2.

[11] P. Cabella, D. Marinucci, Statistical challenges in the analysis of cosmic microwave background radiation, Annals of Applied Statistics 2 (2009) 61–95.
[12] P. Cabella, D. Pietrobon, M. Veneziani, A. Balbi, R. Crittenden, G. de Gasperis, C. Quercellini, N. Vittorio, Foreground influence on primordial non-Gaussianity estimates: needlet analysis of WMAP 5-year data, Monthly Notices of the Royal Astronomical Society 405 (2) (2010) 961–968. arXiv:0910.4362.
[13] T.T. Cai, M.G. Low, H.L. Zhao, Trade-offs between global and local risks in nonparametric function estimation, Bernoulli 13 (2007) 1–19.
[14] J. Delabrouille, J.-F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, F. Guilloux, A full sky, low foreground, high resolution cmb map from WMAP, Astronomy and Astrophysics 493 (3) (2009) 835–857. arXiv:0807.0773.
[15] S. Dodelson, Modern Cosmology, Academic Press, 2003.
[16] D. Donoho, I. Johnstone, G. Kerkyacharian, D. Picard, Density estimation by wavelet thresholding, Annals of Statistics 24 (1996) 508–539.
[17] M. Eastwood, P. Tod, Edth—a differential operator on the sphere, Mathematical Proceedings of the Cambridge Philosophical Society 92 (1982) 317–330.
[18] G. Faÿ, F. Guilloux, M. Betoule, J.-F. Cardoso, J. Delabrouille, M. Le Jeune, CMB power spectrum estimation using wavelets, Physical Review D (2008) D78:083013, arxiv:0807.1113.
[19] D. Geller, F.K. Hansen, D. Marinucci, G. Kerkyacharian, D. Picard, Spin needlets for cosmic microwave background polarization data analysis, Physical Review D (2008) D78:123533, arXiv:0811.2881.
[20] D. Geller, X. Lan, D. Marinucci, Spin needlets spectral estimation, Electronic Journal of Statistics 3 (2009) 1497–1530. arXiv:0907.3369.
[21] D. Geller, D. Marinucci, Spin wavelets on the sphere, Journal of Fourier Analysis and its Applications 16 (6) (2008) 840–884. arXiv:0811.2835.
[22] D. Geller, D. Marinucci, Mixed needlets, Journal of Mathematical Analysis and Applications 375 (2) (2010) 610–630. arXiv:1006.3835.
[23] D. Geller, A. Mayeli, Continuous wavelets on manifolds, Mathematische Zeitschrift 262 (2009) 895–927. arXiv:math/0602201.
[24] D. Geller, A. Mayeli, Nearly tight frames and space-frequency analysis on compact manifolds, Mathematische Zeitschrift 263 (2009) (2009) 235–264. arXiv:0706.3642.
[25] D. Geller, A. Mayeli, Besov spaces and frames on compact manifolds, Indiana University Mathematics Journal 58 (2009) 2003–2042. arXiv:0709.2452.
[26] D. Geller, A. Mayeli, Nearly tight frames of spin wavelets on the sphere, Sampling Theory in Signal and Image Processing 9 (2010) (2010) 25–57. arXiv:0907.3164.
[27] D. Geller, I. Pesenson, Band-limited localized parseval frames and besov spaces on compact homogeneous manifolds, Journal of Geometric Analysis 21 (2) (2011) 334–371.
[28] C.R. Genovese, M. Perone-Pacifico, I. Verdinelli, L. Wasserman, On the path density of a gradient field, Annals of Statistics 37 (6A) (2009) 3236–3271.
[29] C.R. Genovese, M. Perone-Pacifico, I. Verdinelli, L. Wasserman, Nonparametric filament estimation, 2010. arXiv:1003.5536.
[30] T. Ghosh, J. Delabrouille, M. Remazeilles, J.-F. Cardoso, T. Souradeep, Foreground maps in WMAP frequency bands, 2010. arXiv:1006.0916.
[31] J.N. Goldberg, E.T. Newman, Spin-s spherical harmonics and ð, Journal of Mathematical Physics 8 (11) (1967) 2155–2166.
[32] F. Guilloux, G. Fay, J.-F. Cardoso, Practical wavelet design on the sphere, Applied and Computational Harmonic Analysis 26 (2009) 143–160. arXiv:0706.2598.
[33] W. Hardle, G. Kerkyacharian, D. Picard, A. Tsybakov, Wavelets, Approximations and Statistical Application, Springer, Berlin, 1997.
[34] A.B. Juditsky, O.V. Lepski, A.B. Tsybakov, Nonparametric estimation of composite functions, Annals of Statistics 37 (2009) 1360–1404.
[35] M. Kamionkowski, A. Kosowski, A. Stebbins, Statistics of cosmic microwave background polarization, Physical Review D 55 (12) (1996) 7468–7388.
[36] G. Kerkyacharian, R. Nickl, D. Picard, Concentration inequalities and confidence bands for needlet density estimators on the unit sphere, Probability Theory and Related Fields 150 (2011).
[37] G. Kerkyacharian, D. Picard, Thresholding algorithms, maxisets and well concentrated bases, Test 9 (2000) 283–344.
[38] G. Kerkyacharian, D. Picard, Regression in random design and warped wavelets, Bernoulli 10 (6) (2004) 1053–1105.
[39] P.T. Kim, J.-Y. Koo, Optimal spherical deconvolution, Journal of Multivariate Analysis 80 (2002) 21–42.
[40] P.T. Kim, J.-Y. Koo, Z.-M. Luo, Weyl eigenvalue asymptotics and sharp adaptation on vector bundles, Journal of Multivariate Analysis 100 (2009) 1962–1978.
[41] T. Kitching, et al. gravitational lensing accuracy testing 2010 (great10) challenge handbook, preprint, 2010. arXiv:1009.0779.
[42] J.-Y. Koo, P.T. Kim, Sharp adaptation for spherical inverse problems with applications to medical imaging, Journal of Multivariate Analysis 99 (2008) 165–190.
[43] X. Lan, D. Marinucci, The needlets bispectrum, Electronic Journal of Statistics 2 (2008) 332–367. arXiv:0802.4020.
[44] X. Lan, D. Marinucci, On the dependence structure of wavelet coefficients for spherical random fields, Stochastic Processes and their Applications 119 (2008) 3749–3766. arXiv:0805.4154.
[45] N. Leonenko, L. Sakhno, On spectral representations of tensor random fields on the sphere, 2009. arXiv:0912.3389.
[46] A. Malyarenko, Invariant random fields in vector bundles and applications to cosmology, 2009. arXiv:0907.4620.
[47] D. Marinucci, High-resolution asymptotics for the angular bispectrum of spherical random fields, Annals of Statistics 34 (2006) 1–41. arXiv:math/0502434.
[48] D. Marinucci, D. Pietrobon, A. Balbi, P. Baldi, P. Cabella, G. Kerkyacharian, P. Natoli, D. Picard, N. Vittorio, Spherical needlets for CMB data analysis, Monthly Notices of the Royal Astronomical Society 383 (2) (2008) 539–545. arXiv:0707.0844.
[49] A Mayeli, Asymptotic uncorrelation for mexican needlets, Journal of Mathematical Analysis and Applications 363 (1) (2010) 336–344. arXiv:0806.3009.
[50] J.D. McEwen, Fast, exact (but unstable) spin spherical harmonic transforms, 2008. arXiv:astro-ph/0807.4494.
[51] F.J. Narcowich, P. Petrushev, J.D. Ward, Localized tight frames on spheres, SIAM Journal of Mathematical Analysis 38 (2006) 574–594.
[52] F.J. Narcowich, P. Petrushev, J.D. Ward, Decomposition of besov and triebel-lizorkin spaces on the sphere, Journal of Functional Analysis 238 (2) (2006) 530–564.
[53] E.T. Newman, R. Penrose, Note on the Bondi–Metzner–Sachs group, Journal of Mathematical Physics 7 (5) (1966) 863–870.
[54] T.M. Pham Ngoc, Regression in random design and Bayesian warped wavelets estimators, 2009. arXiv:0908.2965.
[55] D. Pietrobon, A. Amblard, A. Balbi, P. Cabella, A. Cooray, D. Marinucci, Needlet detection of features in WMAP CMB sky and the impact on anisotropies and hemispherical asymmetries, Physical Review D (2008) D78 103504. arXiv:0809.0010.
[56] D. Pietrobon, A. Balbi, D. Marinucci, Integrated sachs-wolfe effect from the cross correlation of WMAP3 year and the NRAO VLA sky survey data: new results and constraints on dark energy, Physical Review D (2006) id. D:74, 043524.
[57] D. Pietrobon, P. Cabella, A. Balbi, G. de Gasperis, N. Vittorio, Constraints on primordial non-gaussianity from a needlet analysis of the WMAP-5 data, Monthly Notices of the Royal Astronomical Society 396 (3) (2009) 1682–1688. arXiv:0812.2478.
[58] O. Rudjord, F.K. Hansen, X. Lan, M. Liguori, D. Marinucci, S. Matarrese, An estimate of the primordial non-gaussianity parameter f_{NL} using the needlet bispectrum from WMAP, Astrophysical Journal 701 (2009) 369–376. arXiv:0901.3154.
[59] O. Rudjord, F.K. Hansen, X. Lan, M. Liguori, D. Marinucci, S. Matarrese, Directional variations of the non-gaussianity parameter f_{NL}, Astrophysical Journal 708 (2) (2010) 1321–1325. arXiv:0906.3232.
[60] A. Schwartzman, W.F. Mascarenhas, J.E.T. Taylor, Inference for eigenvalues and eigenvectors of gaussian symmetric matrices, Annals of Statistics 36 (2008) 2886–2919.
[61] U. Seljak, M. Zaldarriaga, An all-sky analysis of polarization in the microwave background, Physical Review D 55 (4) (1997) 1830–1840.
[62] D.A. Varshalovich, A.N. Moskalev, V.K. Khersonskii, Quantum Theory of Angular Momentum, World Scientific, Singapore, 1988.
[63] N.Ja. Vilenkin, A.U. Klimyk, Representation of Lie Groups and Special Functions, Kluwer, Dordrecht, 1991.
[64] Y. Wiaux, L. Jacques, P. Vandergheynst, Fast spin ± 2 spherical harmonics and applications in cosmology, Journal of Computational Physics 226 (2007) 2359–2371.