Needlet Analysis of Spherical Random Fields

A thesis presented

by

Xiaohong Lan

Submitted to Universitá degli Studi di Roma, Tor Vergata

in partial fulfillment of the requirements

for the degree of

PhD

in

Mathematics

[05] [2009]

© [2009] by [LAN Xiaohong]
All rights reserved.

Acknowledgments

Sincere thanks to Prof. Domenico Marinucci who has given me the greatest help to finish my thesis. I have learned a lot in these 3 years. Thank you sir! Thank you for making me confident about myself. Thank you for making me realize what to do in my life and continuously helping me achieve my dream. Thank you sir!

Many thanks to the committee of Equal Opportunity, University of Rome, Tor Vergata to provide me the opportunity to study abroad in Rome, one of the greatest cities in the world; they are Marina Tesauro, Elisabetta Strickland and Velleda Baldoni. Especially thanks to Prof. Elisabetta Strickland, as well as my adviser in Beijing Prof. Li Jiayu, who have given me much support to study and live abroad well.

I would like to thank my family and relatives. Thanks to my grandparents, who are always the sweetest persons to me. Thank to my parents, who always encourage me whenever I met problem and give me the strength to go on. Thanks to my young sister, who takes the responsibility to take care of my parents and enables me to concentrate on my thesis.

Especially many thanks to my dearest friend Wu Jinsong, I owe you a lot. I also would like to take this chance to thank my other friends, they are Chen Fang, Huang Chao, Liu Beibei, Miccoli Lorenzo, Qian Zhonghui, Ruan Huiling, Tanimoto Yoh,

Wang Yingying, Xiao Jun, Zhang Ying, Zingoni Monica, Zingoni Silvano and many other friends.

Thank you all!

Contents

Abstract

This thesis is a collection of essays on spherical wavelets and their applications on statistical models. In particular, its aim is to investigate a form of second generation wavelets on the sphere called needlets, and their statistical applications to the analysis of isotropic spherical random fields. These methods are strongly motivated by many applications, especially from Cosmology and Astrophysics; in particular, the analysis of so-called Cosmic Microwave Background (CMB) radiation.

Chapter 1 is to introduce the physical background and motivations of this thesis. We will emphasize on the nature of CMB, and the statistical challenges in this field.

The concept of needlet is introduced in Chapter 2. This chapter is also preliminary to all the remaining work we have done. It includes the introduction of continuous and discrete spherical wavelets, isotropic spherical random fields, and the diagram formulae, i.e. the basic ingredients that will be needed in the following chapters.

Chapter 3 is devoted to the investigation of non-Gaussianity testing, focusing on a new statistical procedure which we label the needlets bispectrum. We also establish a central limit theorem and multivariate procedures for these statistics and investigate their power properties against non-Gaussian alternatives.

In Chapter 4, we consider an extension of the needlet ideas, leading to a new class of spherical wavelets, called Mexican needlets. In particular, we investigate the extent in which such Mexican needlets enjoy the same stochastic properties as the

standard construction. By means of these, we go on to establish some limit results for related statistics.

Some auxiliary material is collected in Appendices A-C.

Chapter 1
Motivations and Overview

The interest on isotropic random fields on the sphere has grown enormously over the last decade. Such interest is primarily due to strong motivations coming from the applied sciences; in particular, spherical random fields emerge naturally in fields as different as Image Analysis, Brain Mapping, Atmospheric Sciences, Geophysics, Astronomy and Astrophysics, and Cosmology. In particular, although we believe that our results have a broad mathematical interest and can be applied to many different areas of research, we recognize that our leading motivations arose in connection with the analysis of so-called Cosmic Microwave Background radiation (hereafter CMB). To ease the reader, in this preface we shall hence discuss some issues concerning our motivations; these issues are by no means necessary for the understanding of our mathematical results, but we hope to provide a clue to the importance of this area of research.

CMB radiation provides the evidence from an earlier, hotter, and denser period of the Universe, providing pictures of the latter at the time of so-called recombination. The existence of CMB was predicted by G.Gamow in a series of papers in the forties of last century; it was then discovered in 1965 by Arno Penzias and Robert Wilson at the Bell Telephone Laboratories, and this discovery granted them the No-

bel Prizes for Physics in 1978. For several years, further experiments were only able to confirm the existence of the radiation, and to test its adherence to the Planckian curve of blackbody emission, as predicted by theorists (see [24]). A major breakthrough occurred with NASA satellite mission COBE, which was launched in 1989 and publicly released the first full-sky maps of radiation in 1992; these maps granted the Nobel Prize for Physics to Smoot and Mather in 2006 ([77]).

Our basic understanding of CMB can be loosely explained as follows (see again [24] for a textbook account). In nits primordial state, matter was completely ionized, i.e. electrons and protons did not form any stable nuclei. The so-called "cross-section" of a free electron, i.e. the probability of interaction with a photon, is so high that at such stage the Universe can be considered completely "opaque". As the Universe expands, the mean energy content decreases, i.e. the fluid of matter and radiation cools down; the mean kinetic energy of the electrons decreases as well, so that they are captured into orbits around the protons; stable (and neutral) hydrogen atoms are then formed. This change of state occurs at the so-called "age of recombination", which is 3.7×10^5 years after the Big Bang, i.e. when the Universe had only the 0.003% of its current age. During this period, the probability of photon-electrons interactions became negligible and the former could start to travel freely through space. Neglecting second order effects, we can assume they had no further interaction up to the present epoch. The remarkable consequence of this mechanism is that the Uni-

verse is embedded in a uniform radiation that provides pictures of its state nearly 1.37×10^{10} years ago; this is exactly the above-mentioned CMB radiation.

CMB is distributed in remarkably uniform fashion over the sky, with deviations in the order of 10^{-4} with respect to the mean value (corresponding to 2.731 Kelvin degrees). The attempts to understand this uniformity have lead to very important developments in cosmology, primarily the inflationary scenario which now dominates the theoretical landscape. Even more important, though, are that the tiny fluctuations, i.e. the one-in-10^4 variations observed in the CMB, are exactly the amplitude to form the large-scale structures such as stars and galaxies, which are formed out by the denser areas under the attractive influence of gravity. Measuring and understanding the nature of these fluctuations has then been the core of an enormous amount of experimental and theoretical research. In particular, their stochastic properties yield a goldmine of information on a variety of extremely important issues on astrophysics and cosmology, and on many problems at the frontier of fundamental physics (just to mention a few, the matter content of the Universe, its global geometry, the existence and nature of (non-baryonic) dark matter, the existence and nature of dark energy, which is related to Einstein's cosmological constant, and many others).

Much more refined maps have been made available by another NASA satellite experiments, *WMAP* (http://wmap.gsfc.nasa.gov/); still more refined data are expected from the ESA mission *Planck,* which is due to be launched in Spring 2009. These huge collaborations involve hundreds of scientists and are expected to provide

invaluable information on Physics and Cosmology. At the same time, these massive data sets have called for huge statistical challenges, ranging from power spectrum estimation (see for instance [36, 31]) to outlier detection ([85, 86]), testing for isotropy ([18, 19]), efficient denoising and map-making ([21, 48]), handling with missing data ([64, 83]), density estimation ([9, 37]), testing for non-Gaussianity ([62, 70]) and many others.

Among these issues, particular interest has been driven by efficient testing for non-Gaussianity. This is due to strong physical motivations (the leading paradigm for the Big Bang dynamics predicts (very close to) Gaussian fluctuations) and difficulties in finding a proper statistical procedure. There is by now a wide consensus that the most efficient procedures to probe non-Gaussianity are based upon the so-called bispectrum (see for instance [60] for an alternative method of testing for non-Gaussianity), in the idealistic circumstances where the spherical random field is fully observed: see for instance [10], [14], [52], [56], [57], [88]. However, cosmological maps are usually provided with large parts of missing observations, and the localization properties of wavelets in the real domain are expected to restore at least some properties that would be lost by a naive application of Fourier methods. There is currently a rapidly growing literature on the construction of wavelets systems on the sphere, owning to the simultaneous space and frequency localization, see for instance [2], [89] and the references therein. Among spherical wavelets, particular attention has been devoted to so-called needlets, which were introduced in the Func-

tional Analysis literature by [66, 67]; their statistical properties were first considered by [7], [49], [8] and [53]. Needlets enjoy one main property which seems especially worth recalling: they are quasi-exponentially localized in the real domain and compactly supported in the harmonic domain. A further, quite unexpected property is as follows: under regularity conditions for the angular power spectrum, random needlets coefficients are asymptotically uncorrelated at the highest frequencies and hence, in the Gaussian case, independent, see again [7]. This latter feature is rather surprising in a compact domain and makes asymptotic theory possible even in the presence of a single realization of a spherical random field.

In this thesis, we shall first review some basic features of Fourier analysis on the sphere, spending some time to highlight connections between harmonic analysis and group representation theory. Such connections are important for the understanding of research in this area, both from the point of view of probability characterizations and for statistical applications. We shall then review spherical wavelets, with particular attention to the needlet construction we shall adopt, and some other background material, such as the diagram formula and spectral representations.

This background material is then exploited in our first line of research, which borrows ideas from the bispectrum and the needlets literature to propose and analyze a needlets bispectrum, where the random coefficients in the needlets expansion are combined alike to the bispectrum construction. The aim is to obtain a procedure which mimics the ability of the bispectrum to search for non-Gaussianity at the

most efficient combination of frequencies, at the same time providing a much more robust construction in the presence of missing data, as typical of the needlets. We shall provide asymptotic results both in Gaussian and non-Gaussian circumstances, suggesting the procedure we suggest may indeed be useful when dealing with CMB data; indeed, simulated results and applications to WMAP data have been provided in [76].

Another line of research we pursued is to consider the correlation structure of the random coefficients for a wider class of wavelet systems on the sphere. Indeed, the uncorrelation property we discussed for standard needlets coefficients is so crucial for applications that it is natural to investigate to what extent it is unique or shared by other systems. To this aim, we focus on an extension of the needlet idea (the so-called Mexican needlets approach) which was recently introduced in the literature by [34]. This approach not only has considerable mathematical interest, but it also embeds as a special case a good approximation to the so-called Spherical Mexican Hat Wavelet construction, which is extremely popular among practitioners. In this thesis, we shall provide necessary and sufficient conditions for the Mexican needlet coefficients to be asymptotic uncorrelated in the real and in the frequency domain; these conditions are completely explicit, depending on a combination of user-chosen features and assumptions on the correlation structure of the underlying field. We believe this result has a considerable practical interest for the design of further wavelets, and at the same time it highlights the rationale behind the uncorrelation result: in particular,

it makes clear that the later (somewhat counterintuitively) does *not* follow from the localization properties (in some sense, quite the reverse is true).

Throughout the thesis, the asymptotic theory is developed in the high resolution sense, i.e. we assume we are observing a single realization of a continuous spherical random field, with observations at higher and higher resolutions. We believe this framework fits suitably the actual scientific experience of Cosmologists and other applied scientists.

Chapter 2
Preliminaries

2.1 Introduction

Let $\mathbb{S}^2 = \{x \in R^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ denote as usual the unit sphere in $\mathbb{R}^3$. In this thesis, we shall be concerned with zero-mean, mean square continuous and isotropic random fields $T(x)$ on the sphere $\mathbb{S}^2$; here, by isotropic we mean that the law of the field is invariant with respect to the action of the group of rotations in $\mathbb{R}^3$, i.e. $T(gx) \stackrel{d}{=} T(x)$ for all $g \in SO(3)$, where $\stackrel{d}{=}$ denotes identity in distribution as fields. Our main instrument for the analysis of such fields will be a spectral representation of the following form [1, 5, 36] : for $x \in \mathbb{S}^2$,

$$T(x) = \sum_{l,m} a_{lm} Y_{lm}(x) \,, \tag{2.1}$$

where $\{a_{lm}\}$ is a set of random coefficients such that they are zero-mean, orthogonal, complex-valued random variables with variance $\mathbb{E}|a_{lm}|^2 = C_l$. The $\{Y_{lm}\}$ are a triangular array of orthonormal functions whose properties we shall discuss below, and the representation (2.1) holds in the $L^2(\mathbb{S}^2)$ sense, i.e. we have

$$\lim_{L\to\infty} \mathbb{E} \left| \int_{S^2} \left\{ T(x) - \sum_{l=1}^{L} \sum_{m=-l}^{l} a_{lm} Y_{lm}(x) \right\}^2 \mu(dx) \right|^2 = 0 \,,$$

$\mu(dx)$ denoting the uniform measure on the sphere. There are many routes for establishing this fact, usually by means of Karhunen-Loeve arguments, the Spectral

Representation Theorem, or the Stochastic Peter-Weyl theorem, see for instance [1], [5] and [58]. Our purpose in this preliminary chapter is to review some of these basic concepts related to Fourier analysis on the sphere. For the sake of brevity, we shall not provide full details, and we refer instead to textbooks on random fields and harmonic analysis for a more detailed treatment. In particular, we will follow mainly the definitions and notations from [46]; other recommended references include for instance ([1], [40]). We then start to review briefly some basic background facts.

2.2 Fourier Analysis on $\mathbb{S}^2$

Let us first recall that the standard Laplacian in $\mathbb{R}^3$ with Euclidean metric is represented by

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \, .$$

Let us now consider the change of variables from rectangular to spherical coordinates

$$x_1 = r \sin\vartheta \sin\phi \, , \tag{2.2}$$

$$x_2 = r \sin\vartheta \cos\phi \, , \tag{2.3}$$

$$x_3 = r \cos\vartheta \, . \tag{2.4}$$

To proceed with our arguments, we need to recall a few basic facts from differential geometry. We will omit most details and all proofs, and we refer to standard textbooks such as [12] for a much more complete analysis.

In spherical coordinates (r, ϑ, ϕ), the Euclidean metric tensor transforms into

$$\begin{aligned} ds^2 &= dx_1^2 + dx_2^2 + dx_3^2 \\ &= dx_1 \otimes dx_1 + dx_2 \otimes dx_2 + dx_3 \otimes dx_3 \\ &= \otimes^2 (\sin\vartheta \sin\phi dr + r\cos\vartheta \sin\phi d\theta + r\sin\vartheta\cos\phi d\phi) \\ &\quad + \otimes^2 (\sin\vartheta\cos\phi dr + r\cos\vartheta\cos\phi d\theta - r\sin\vartheta\sin\phi d\phi) \\ &\quad + \otimes^2 (\cos\vartheta dr - r\sin\vartheta d\vartheta) \end{aligned}$$

where $\otimes^2$denotes that for any tensor ω, $\otimes^2\omega := \omega \otimes \omega$. Thus

$$\begin{aligned} ds^2 &= dr \otimes dr + r^2 d\vartheta \otimes d\vartheta + r^2 \sin^2\vartheta d\phi \otimes d\phi \\ &= dr^2 + r^2 d\vartheta^2 + r^2 \sin^2\vartheta d\phi^2. \end{aligned}$$

The metric can also be written in matrix form as

$$g = (g_{ij})_{i,j=1,2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\vartheta \end{pmatrix}.$$

Now we recall that the intrinsic Laplacian under a transformed coordinate system y is given by (see [1, 12])

$$\frac{1}{\sqrt{|g|}} \frac{\partial}{\partial y^i} \left(\sqrt{|g|} g^{ij} \frac{\partial f}{\partial y^j} \right),$$

where (g^{ij}) and $|g|$ are the inverse matrix and determinant of g, respectively. Here we are using the standard convention (due to Einstein) that repeated indexes are summed. In our case, it is easily seen that this leads to the well-known expression for the Laplacian in spherical coordinates, that is

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \frac{\partial}{\partial r}) + \frac{1}{r^2\sin\vartheta}\frac{\partial}{\partial\vartheta}(\sin\vartheta \frac{\partial}{\partial\vartheta}) + \frac{1}{r^2\sin^2\vartheta}\frac{\partial^2}{\partial\phi^2} \tag{2.5}$$

This expression can also be derived more directly, but much more tediously, working directly with the change of variables in (2.2)-(2.4) and standard rules of multivariate calculus.

It is then natural to decompose the Laplacian into a radial part $\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial}{\partial r})$ and a spherical part

$$\Delta_{S^2} = \frac{1}{r^2\sin\vartheta}\frac{\partial}{\partial\vartheta}(\sin\vartheta\frac{\partial}{\partial\vartheta}) + \frac{1}{r^2\sin^2\vartheta}\frac{\partial^2}{\partial\phi^2}\ ,$$

where the latter is usually labelled as the spherical Laplacian. Now let us consider the harmonic homogeneous polynomials $a_{\alpha\beta}x_1^\alpha x_2^\beta x_3^{l-(\alpha+\beta)}$ of degree l satisfying

$$\Delta\left(a_{\alpha\beta}x_1^\alpha x_2^\beta x_3^{l-(\alpha+\beta)}\right) = 0,$$

i.e. under spherical coordinates,

$$\Delta\left(r^l H_l\left(\vartheta,\phi\right)\right) = 0\ , \tag{2.6}$$

with

$$H_l\left(\vartheta,\phi\right) = \sum_{\alpha,\beta\geq 0,\alpha+\beta\leq l} a_{\alpha\beta}\left(\sin\vartheta\sin\phi\right)^\alpha\left(\sin\vartheta\cos\phi\right)^\beta\left(\cos\vartheta\right)^{l-(\alpha+\beta)}\ .$$

In view of (2.5), (2.6) and easy manipulations, we obtain

$$r^l l\left(l+1\right)H_l\left(\vartheta,\phi\right) + r^{l-2}\Delta_{S^2}H_l\left(\vartheta,\phi\right) = 0\ ,$$

from which it follows immediately that

$$\Delta_{S^2}H_l\left(\vartheta,\phi\right) = -l\left(l+1\right)H_l\left(\vartheta,\phi\right)\ , \tag{2.7}$$

by taking $r = 1$, i.e. considering the restrictions of harmonic polynomials to the unit sphere. We have then established a very important fact, that is, $H_l\left(\vartheta,\phi\right)$ must

be included in the space $\mathcal{H}_l(S^2)$ spanned by the eigenfunctions of the spherical Laplacian, with eigenvalue $-l(l+1)$. From an analytic point of view, the chain of ideas which leads to the spectral representation (2.1) can now be explained as follows. We shall show below that the spherical harmonics $\{Y_{lm}\}_{m=-l,\ldots,l}$ are indeed a system of orthonomal bases for the spaces of homogeneous harmonic polynomials $\mathcal{H}_l(\mathbb{S}^2)$; moreover, it is a standard fact in functional analysis that the polynomial space $\oplus_{l=0}^{\infty}\mathcal{H}_l(\mathbb{S}^2)$ is dense in $L^2(\mathbb{S}^2)$, from which it will follow that the spherical harmonics provide a complete orthonormal systems for $L^2(\mathbb{S}^2)$.

More precisely, it is well-known that the space of continuous functions is dense in L^2 space. By Weierstrass theorem, a continuous function $f(x,y,z)$ which is defined in the compact domain $-1 \leq x,y,z \leq 1$, can be uniformly approximated by the polynomial

$$f_n = \sum_{\alpha,\beta,\gamma=0}^{n} a_{\alpha\beta\gamma} x^{\alpha} y^{\beta} z^{\gamma}.$$

The monomials $x^{\alpha}y^{\beta}z^{\gamma}$ therefore form a basis for the space of square-integrable functions over this domain; that is, they form a complete set of functions. We can group the monomials together as homogeneous polynomials of degree $l = \alpha+\beta+\gamma$ and note that there are $\frac{1}{2}(l+1)(l+2)$ linearly independent monomials of degree l. On the unit sphere they are subject to the constraint $x^2+y^2+z^2=1$, so that they are no longer all linearly independent.

Using this constraint to eliminate x^2, each term $x^{\alpha}y^{\beta}z^{\gamma}$ can be reduced to terms of the form $y^{\beta}z^{\gamma}$ where $\beta+\gamma=l$, or $xy^{\beta}z^{\gamma}$ where $\beta+\gamma+1=l$ together with lower

degree monomials. There are $l+1$ possible combinations in the first case and l in the second, giving a total of $2l+1$ linearly independent monomials of the l-th degree. This establishes the dimension of $\mathcal{H}_l\left(\mathbb{S}^2\right)$.

We shall now derive explicitly the basis of $\mathcal{H}_l\left(\mathbb{S}^2\right)$ as a set of spherical harmonics $\{Y_{lm}\}_{m=-l,\ldots l}$. To do so, we need to solve the differential equations (2.7); we use the method of separation of variables, and look for solutions of the form $Y_{lm}\left(\theta,\phi\right) = \Theta_{lm}(\theta)\Phi_{lm}(\phi)$; for $\Phi_{lm}(\phi)$, we shall take the functions $\exp(im\phi)$, which are easily seen to form an orthogonal system over the sphere $\mathbb{S}^2$. With this choice for Φ_{lm} , we find that Θ_{lm} satisfies the equation

$$\begin{aligned}
&\left\{\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right)+\frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial\phi^2}\right\}\Theta_{lm}(\vartheta)\exp(im\phi)+l(l+1)\Theta_{lm}(\vartheta)\exp(im\phi)\\
=\;&\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right)\Theta_{lm}(\vartheta)\exp(im\phi)+\frac{P_{lm}(\vartheta)}{\sin^2\vartheta}\frac{\partial^2}{\partial\varphi^2}\exp(im\phi)+l(l+1)\Theta_{lm}(\vartheta)\exp(im\phi)\\
=\;&\left\{\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right)+\left[l(l+1)-\frac{m^2}{\sin^2\vartheta}\right]\right\}\Theta_{lm}(\vartheta)\exp(im\phi)\,,
\end{aligned}$$

which leads to

$$\left\{\frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right)+\left[l(l+1)-\frac{m^2}{\sin^2\vartheta}\right]\right\}\Theta_{lm}(\vartheta)=0\,,$$

i.e., after the change of variable $\mu=\cos\vartheta$,

$$\left\{\frac{d}{d\mu}\left((1-\mu^2)\frac{d}{d\mu}\right)+\left[l(l+1)-\frac{m^2}{1-\mu^2}\right]\right\}\Theta_{lm}(\arccos\mu)=0\,. \tag{2.8}$$

Hence, for $m=0$ we have

$$\frac{d}{d\mu}\left((1-\mu^2)\frac{d}{d\mu}\right)\Theta_{l0}(\arccos\mu)+l(l+1)\Theta_{l0}(\arccos\mu)=0\,,$$

and a solution is given by

$$\Theta_{l0}(\arccos\mu) = P_l(\mu) = \frac{1}{2^l l!}\frac{d^l}{d\mu^l}(\mu^2-1)^l \ , \tag{2.9}$$

which is called the Legendre polynomial of the l-th degree (the normalizing factor is chosen so that $P_l(\cos 0) = 1$). It may be noted that Legendre polynomials are a special case of so-called Jacobi polynomials, which play a very important role in the theory of special functions (see [87]).

We note that $P_l(\mu)$ is independent of longitude-ϕ over the unit sphere $\mathbb{S}^2$, so that the Legendre polynomials are naturally applicable to problems with axial symmetry. For non-symmetrical problems, it is necessary to introduce the associated Legendre function, which is defined in terms of the Legendre polynomial by the equation

$$P_{lm}(\mu) = (-1)^m(1-\mu^2)^{m/2}\frac{d^m}{d\mu^m}P_l(\mu) \ , \ m = -l, ..., l, \ l = 0, 1, 2, ...$$

and by a careful calculation, we see that the normalized form of $\Theta_{lm}\,(\vartheta)$ can be taken as $P_{lm}(\cos\vartheta)$. We shall recall here two important properties of the $P_{lm}(x)$, both of which we shall require later; see for instance [84] for a more complete treatment:

Orthogonality

The $P_{lm}(x)$ form an orthogonal set of functions on the interval $[-1, 1]$. Thus

$$\int_{-1}^{1} P_{l'm}(\mu)P_{lm}(\mu)d\mu = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}\delta^l_{l'} \ ,$$

where $\delta^l_{l'}$ is the Kronecker delta, such that $\delta^l_{l'} = 0$ if $l' \neq l$ and $\delta^l_{l'} = 1$, if $l = l'$.

Recurrence relation

The Legendre polynomials not only satisfy the associated differential equation (2.8), they also satisfy following equations:

$$(2l+1)\,\mu P_{lm} = (l-m+1)\,P_{l+1,m} + (l+m)\,P_{l-1,m}\ ,$$
$$\left(1-\mu^2\right)\frac{dP_{lm}}{d\mu} = (l+1)\,\mu P_{lm} - (l-m+1)\,P_{l+1,m}\ .$$

We are now in the position to provide an explicit definition for the spherical harmonics $\{Y_{lm}\}$, see for instance [46, 84] for some textbook analysis. The definition we shall adopt is

$$\begin{aligned} Y_{lm}(\theta,\phi) &= \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_{lm}(\cos\theta)\exp(im\phi),\ \ for\quad m\geq 0, \\ Y_{lm}(\theta,\phi) &= (-1)^m Y^*_{l,-m}(\theta,\phi),\ \ for\quad m<0\ . \end{aligned}$$

It is immediate to see that the spherical harmonics are orthogonal for different values of m, m'; indeed we have

$$\begin{aligned} \int_{S^2} Y_{lm}(x)Y^*_{lm'}(x)dx &= \int_0^\pi\int_{-\pi}^\pi Y_{lm}(\vartheta,\varphi)Y^*_{lm'}(\vartheta,\varphi)\sin\vartheta d\vartheta d\phi \\ &= \frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\int_0^\pi P^2_{lm}(\cos\vartheta)\sin\vartheta d\vartheta\int_{-\pi}^\pi \exp(i(m-m')\phi)d\phi \end{aligned}$$

and the latter integral can be different from zero only if $m=m'$. Thus it follows that the set $\{Y_{lm}\}_{m=-l,\ldots,l}$ for fixed l consists of $2l+1$ orthogonal functions over the unit sphere, and hence it makes up an orthogonal basis. To establish the orthogonality across different multipoles l, we use a standard argument for eigenfunctions of the spherical Laplacian operator. More precisely, we recall that, by Stokes' Theorem and standard manipulations on the sphere (see for instance [78]) , it is possible to show

that

$$\int_{S^2} \Delta_{S^2} Y_{lm} Y^*_{l'm'} dx = \int_{S^2} Y_{lm} \Delta_{S^2} Y^*_{l'm'} dx \; .$$

Now recall that the $\{Y_{lm}(\vartheta, \phi)\}$ are the eigenfunctions of spherical Laplacian operator, i.e. they satisfy $\Delta_{S^2} Y_{lm} = -l\,(l+1)\,Y_{lm}$, whence

$$\begin{aligned} -l\,(l+1) \int_{S^2} Y_{lm} Y^*_{l'm'} dx &= \int_{S^2} \Delta_{S^2} Y_{lm} Y^*_{l'm'} dx \\ &= \int_{S^2} Y_{lm} \Delta_{S^2} Y^*_{l'm'} dx \\ &= -l'\,(l'+1) \int_{S^2} Y_{lm} Y^*_{l'm'} dx \; . \end{aligned}$$

The orthogonality property for $l \neq l'$ now follows immediately.

2.3 Connection with Group Representation Properties for SO(3)

In this Section, we shall explore some background material on representation theory for the group of rotations $SO(3)$, and we shall discuss its relationships with spherical harmonics and the spectral representation.

It is a standard fact from elementary geometry that any element of $SO(3)$ can be represented in terms of three Euler angles, i.e., for all $g \in SO(3)$ we have

$$g \begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_z(\alpha) R_x(\beta) R_z(\gamma) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \; ,$$

where

$$R_z(\phi) := \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\theta & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \; , R_x(\phi) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \; ,$$

$0 \leq \alpha, \gamma < 2\pi$, $0 \leq \beta \leq \pi$. More explicitly, each rotation can be realized by rotating first by an angle γ around the z axis, then by an angle β around the y axis, than again around the z axis by an angle α; the angles α, β, γ are called the Euler angles of g. We recall that a *matrix representation* (of dimension n) of a group (see for instance [87, 84, 29]) is a homomorphism H from the group to a matrix space $H : \mathcal{G} \rightarrow \mathcal{M}_n$ which preserves the group structure, i.e. for $g_1, g_2 \in \mathcal{G}$, $H(g_1) = M_1 \in \mathcal{M}_n$, $H(g_2) = M_2 \in \mathcal{M}_n$ implies that $H(g_1 g_2) = M_1 M_2$. The representation M acts on an $n-$dimensional vector space V; a representation is called *irreducible* if there is no invariant subspace other than the null space and V itself.

For the group $SO(3)$, a well-known family of irreducible representations is provided by the so-called Wigner's matrices $D^l(\alpha, \beta, \gamma)$, with elements

$$D^l_{mm'}(\alpha, \beta, \gamma) = e^{-im\alpha} d^l_{mm'}(\beta)\, e^{-im'\gamma}, \tag{2.10}$$

where

$$\begin{aligned} d^l_{mm'}(\beta) &= (-1)^{l-m'} \left[(l+m)!\,(l-m)!\,(l+m')!\,(l-m')!\right]^{1/2} \\ &\times \sum_{k=0}^{\max(l-m, l-m')} \frac{\left(\cos\frac{\beta}{2}\right)^{2k-m-m'} \left(\sin\frac{\beta}{2}\right)^{2l-2k-m-m'}}{k!(l-m-k)!(l-m'-k)!(m+m'+k)!} \,. \end{aligned}$$

The matrices D^l are $(2l+1) \times (2l+1)$ and unitary, i.e.

$$\overline{(D^l)}^t D^l = \mathbb{I}_{2l+1} \,.$$

We recall now briefly how the spectral representation (2.1) can be obtained, at least in the deterministic case, by means of group representation results. In particular, the Peter-Weyl Theorem (see [29, 87]) ensures that the space of square in-

tegrable functions on a compact group such as $SO(3)$ can be decomposed in terms of the elements of the group representation matrices. More explicitly, denoting by dg the uniform (Haar) measure on the group, given any function $h(g)$ belonging to $L^2(SO(3), dg)$, we have, in the L^2 sense

$$h(\alpha, \beta, \gamma) = \sum_{l=0}^{\infty} \sum_{m,m'} b^l_{m'm} D^l_{mm'}(\alpha, \beta, \gamma) \ ,$$

where

$$b^l_{m'm} = \int_{SO(3)} h(g) \overline{D^l_{mm'}(g)} dg \ .$$

Now we identify the sphere $\mathbb{S}^2$ with the quotient space $SO(3)/SO(2)$; in other words, we consider functions of the form $f(g) = f(gR_z(\gamma))$, i.e. the functions which are constant along the commutative subgroup given by the rotations $R_z(\gamma)$. For such functions, we have

$$\begin{aligned} b^l_{m'm} &= \int_{-\pi}^{\pi} \int_{0}^{\pi} \int_{-\pi}^{\pi} f(\alpha, \beta, \gamma)) e^{im\alpha} d^l_{mm'}(\beta) \, e^{im'\gamma} \sin\beta d\alpha d\beta d\gamma \\ &= \int_{-\pi}^{\pi} \int_{0}^{\pi} f(\alpha, \beta, 0) e^{im\alpha} d^l_{mm'}(\beta) \left\{ \int_{-\pi}^{\pi} e^{im'\gamma} \sin\beta d\gamma \right\} d\alpha d\beta \\ &= 0 \text{ unless } m' = 0 \ , \end{aligned}$$

i.e. for functions on the sphere we shall only focus on the coefficients corresponding to $m' = 0$. In particular, we have that,

$$Y_{lm}(\vartheta, \varphi) = \sqrt{\frac{2l+1}{4\pi}} D^l_{m0}(\varphi, \vartheta, \gamma) \ ,$$

so that the spectral representation (2.1) is really a special instance of the Peter-Weyl Theorem. See [71] for a discussion of a stochastic version of Peter-Weyl Theorem.

Remark The different normalization is due to the fact that for the spherical harmonics we require orthonormality, i.e.

$$\int Y_{lm}\overline{Y_{l'm'}}dx = \delta_l^{l'}\delta_m^{m'},$$

while for the elements of D^l we require the unitary property

$$\sum_m \left|D^l_{mm'}(\alpha,\beta,\gamma)\right|^2 = 1\,.$$

From the group representation properties, we have immediately

$$Y_{lm}(gx) = \sum_{m'} D^l_{m'm}(g)\, Y_{lm'}(x)\,. \tag{2.11}$$

The Integral of the triple harmonic product

In the analysis of the higher order moments of spherical random fields and connected invariant statistics such as the bispectrum, we shall frequently need multiple integrals of spherical harmonics. The integral of the triple harmonic product is also known as the Gaunt formulae. With fully normalized spherical harmonics Y_{lm}, it may be written (see [84], [56] for details)

$$\begin{aligned}
&\int_{S^2} Y_{l_1m_1}(x)\, Y_{l_2m_2}(x)\, Y_{l_3m_3}(x)\, dx \\
=\;& 2\pi\sqrt{\frac{2l_1+1}{4\pi}}\sqrt{\frac{2l_2+1}{4\pi}}\sqrt{\frac{2l_3+1}{4\pi}} \\
&\times\int_0^{\pi}\int_0^{2\pi} D^{l_1}_{m_10}(\varphi,\vartheta,\gamma)\, D^{l_2}_{m_20}(\varphi,\vartheta,\gamma)\, D^{l_3}_{m_30}(\varphi,\vartheta,\gamma)\sin\theta d\varphi d\theta \\
=\;& \left[\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}\right]^{1/2}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3\end{pmatrix}\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0\end{pmatrix}
\end{aligned} \tag{2.12}$$

where

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3\end{pmatrix}$$

are called the Wigner's 3-j symbols, and algebraically defined by $(A.1)$. From Appendix A it is clear that the integral vanishes when any of the following relations are not satisfied.

(i) $m_3 = m_1 + m_2$;

(ii) the so called triangle relation in which l_1, l_2 and l_3 form a triangle with integral sides:

$$|l_1 - l_2| \leq l_3 \leq l_1 + l_2;$$

(iii) $(l_1 + l_2 + l_3)$ is even.

Wigner's 3-j coefficients arise in many different instances, especially in the quantum theory of angular momentum (see [84], where explicit expressions are also provided). Apart from their group-theoretical significance, they are very useful in the manipulation of complicated algebraic expressions. They can be efficiently computed, so that no further algebraic reduction is necessary although, in certain cases, simplification is possible. Up to a normalization factor, they are equivalent to the so-called Clebsch-Gordan coefficients-$C^{l_3 m_3}_{l_1 m_1 l_2 m_2}$, which play an important role in representation theory for the group of rotations $SO(3)$. The relation beween these two symbols is given by

$$C^{l_3 m_3}_{l_1 m_1 l_2 m_2} = (-1)^{l_1 - l_2 + m_3} \sqrt{2l_3 + 1} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}$$

see [58, 59] for a much more detailed discussion and probabilistic applications.

Some examples

The following table provides some explicit formulae for the spherical harmonics of degree less than 5.

$TABLE1.The\ Spherical\ Harmonics\ for\ l \leq 4.$

l	m	$Y_{lm}(\theta, \phi)$
0	0	$1/\sqrt{4\pi}$
1	0	$\sqrt{\frac{3}{4\pi}}\cos\theta$
1	± 1	$\mp\sqrt{\frac{3}{8\pi}}\sin\theta\exp(\pm i\phi)$
2	0	$\frac{1}{2}\sqrt{\frac{5}{4\pi}}(3\cos^2\theta - 1)$
2	± 1	$\mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\exp(\pm i\phi)$
2	± 2	$\sqrt{\frac{15}{32\pi}}\sin\theta\exp(\pm 2i\phi)$
3	0	$\frac{1}{2}\sqrt{\frac{7}{4\pi}}(5\cos^3\theta - 3\cos\theta)$
3	± 1	$\mp\sqrt{\frac{21}{64\pi}}\sin\theta(5\cos^2\theta - 1)\exp(\pm i\phi)$
3	± 2	$\sqrt{\frac{105}{32\pi}}\sin^2\theta\cos\theta\exp(\pm 2i\phi)$
3	± 3	$\mp\sqrt{\frac{35}{64\pi}}\sin^3\theta\exp(\pm 3i\phi)$
4	0	$\frac{1}{8}\sqrt{\frac{9}{4\pi}}(35\cos^4\theta - 30\cos^2\theta + 3)$
4	± 1	$\mp\frac{3}{4}\sqrt{\frac{5}{4\pi}}\sin\theta(7\cos^3\theta - 3\cos\theta)\exp(\pm i\phi)$
4	± 2	$\frac{3}{4}\sqrt{\frac{5}{8\pi}}\sin^2\theta(7\cos^2\theta - 1)\exp(\pm 2i\phi)$
4	± 3	$\mp\frac{3}{4}\sqrt{\frac{35}{4\pi}}\sin^3\theta\cos\theta\exp(\pm 3i\phi)$
4	± 4	$\frac{3}{4}\sqrt{\frac{5}{8\pi}}\sin^2\theta(7\cos^2\theta - 1)\exp(\pm 2i\phi)$

2.4 The Spectral Representation, Isotropy and Joint Moments

The previous two sections were meant to highlight some algebraic and analytic aspects related to Fourier analysis on the sphere in the deterministic case. In this sec-

tion, we shall discuss briefly some issues concerning spectral properties in a stochastic setting; again, our presentation will be rather quick for the sake of brevity, and we refer to textbooks and recent papers for much more details (see for instance [1, 5, 40, 58, 59]).

We shall first review the construction of the spectral representation (2.1) under Gaussian circumstances. In some sense, this derivation is redundant, as (2.1) can be viewed as a direct stochastic counterpart of the previous deterministic constructions (such as the Peter-Weyl Theorem); nevertheless, we discuss it briefly to provide one more point of view. Given a mean zero, mean-square continuous Gaussian field $T(x)$, we can define its covariance kernel as

$$K(x,y) = \mathbb{E}T(x)\overline{T(y)} \; ;$$

here, it is convenient to consider the case where $T(x)$ may be complex-valued, the bar as usual denoting conjugation. For fixed x, the function $K(x,.)$ is in $L^2(S^2)$ and we have the decomposition

$$K(x,y) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm} Y_{lm}(x) \overline{Y_{lm}}(y) \; ;$$

the sequence $\{c_{lm}\}$ is necessarily nonnegative because the covariance kernel $K(x,y)$ is positive definite. Now consider the space of functions

$$S = \left\{ f(.) = \sum_{j=1}^{n} a_j K(x_j,.) \; , a_j \in R, \; n \in N, \; x_j \in S^2 \text{ for } j = 1,2,...,n \right\}$$

with inner product $\langle .,. \rangle_H$ given by

$$\left\langle \sum_{j_1} a_{j_1} K(x_{j_1},.), \sum_{j_2} a_{j_2} K(x_{j_2},.) \right\rangle_H := \sum_{j_1 j_2} a_{j_1} \overline{a_{j_2}} K(x_{j_1}, x_{j_2}) \; .$$

The *reproducing kernel property* is readily verified, i.e.

$$\begin{aligned} \langle f(.), K(x,.) \rangle_H &= \left\langle \sum_{j_1} a_{j_1} K(x_{j_1},.), K(x,.) \right\rangle_H \\ &= \sum_{j_1} a_{j_1} K(x_{j_1}, x) = f(x) \; . \end{aligned}$$

It can be shown by standard arguments that the closure of S defines a complete and separable Hilbert space $H(K)$, labeled the *Reproducing Kernel Hilbert Space* (RKHS) of K.

The standard procedure for the derivation of the spectral representation is then the following (see again [1]); consider the space spanned by the random variables $T(x), x \in S^2$

$$\mathcal{H}(K) = span \left\{ T(x), \; x \in S^2 \right\} \; ,$$

with the inner product induced by the covariance function, $\langle T(x), T(y) \rangle_{\mathcal{H}} := K(x,y)$; next define the linear mapping $\Phi : S \rightarrow \mathcal{H}(K)$ such that

$$\Phi(f) = \Phi \left(\sum_j a_j K(x_j,.) \right) = \sum_j a_j T(x_j) \; ;$$

the application clearly is norm-preserving, indeed

$$\langle \Phi(f), \Phi(f) \rangle_{\mathcal{H}} = \sum_{j_1, j_2} a_{j_1} \overline{a_{j_2}} K(x_{j_1}, x_{j_2}) = \langle f, f \rangle_H \; .$$

The application makes $\Phi(f)$ Gaussian with all limits, and can then be extended to all of $H(K)$ with range that covers all of $\mathcal{H}$; this is known as the canonical isomorphism

between H and $\mathcal{H}$. As we have mentioned earlier, H is separable, and hence so it is $\mathcal{H}$; moreover the projection through Φ of an orthonomal basis on H will yield an orthonormal basis on $\mathcal{H}$. An orthonormal basis on H is readily seen to be given by $\sqrt{c_{lm}}Y_{lm}$, whence $\alpha_{lm} := \Phi(\sqrt{c_{lm}}Y_{lm}), m = -l, ..., l, l = 0, 1, 2, ...$ In particular, for a random function T, we have

$$T(x) = \sum_{lm} \alpha_{lm} \langle T(.), \alpha_{lm} \rangle_{\mathcal{H}} \ ,$$

where the identity must be understood in terms of the L^2 convergence. Now, in view of the isometry property

$$\begin{aligned} \langle T(x), \alpha_{lm} \rangle_{\mathcal{H}} &= \langle \Phi \left(K(< x, . >) \right), \Phi(\sqrt{c_{lm}}Y_{lm}) \rangle_{\mathcal{H}} \\ &= \langle K(< x, . >), \sqrt{c_{lm}}Y_{lm}(.) \rangle_{H} \\ &= \sqrt{c_{lm}}Y_{lm}(x) \ , \end{aligned}$$

for all l, m, in view of the reproducing kernel property. Hence

$$T(x) = \sum_{lm} \alpha_{lm}\sqrt{c_{lm}}Y_{lm}(x) = \sum_{lm} a_{lm}Y_{lm}(x) \ ,$$

with $E\left|\alpha_{lm}\right|^2 = 1$, $a_{lm} = \alpha_{lm}\sqrt{c_{lm}}$, and (2.1) is established. Note that here we did not assume isotropy, whence the variances of the random coefficients depend also on m.

We can now focus more directly on the consequences of isotropy, i.e. we recall we shall deal with random fields such that

$$T\left(gx\right) \stackrel{d}{=} T\left(x\right) \ ,$$

for all $g \in SO(3)$, where $\stackrel{d}{=}$ denotes identity in distribution in the sense of stochastic processes. We recall that the spherical harmonic coefficients $\{a_{lm}\}$ can be obtained from $T(\theta, \phi)$ through the inversion formula

$$a_{lm} = \int_0^{2\pi} \int_0^{\pi} T(\theta, \phi) \overline{Y_{lm}}(\theta, \phi) \sin\theta d\theta d\phi, m = 0, \pm 1, ..., \pm l, l = 1, 2, ... \quad (2.13)$$

We have already seen that the triangular array $\{a_{lm}\}, m = 1, ..., l$, is made up of zero-mean and orthogonal, complex-valued random variables. Under isotropy, as a simple consequence of (2.11) we have that

$$a_{l.} \stackrel{d}{=} D^l(g) a_{l.} \text{ , for all } l = 0, 1, 2, ...$$

where $a_{l.} = (a_{l\,-l}, ..., a_{ll})$. Hence, a simple consequence of Schur's Lemma (see [87]) is that the variance of these coefficients is constant over l, $\mathbb{E}|a_{lm}|^2 = C_l$, (the angular power spectrum of the random field), see [5] for details; also, for $m < 0$ we have $a_{lm} = (-1)^m \overline{a_{l\,-m}}$,, whereas a_{l0} is real with the same mean and variance. Moreover, because the random spherical harmonic coefficients are uncorrelated, it is obvious that they are also independent under Gaussianity. It is more surprising that the reverse also holds, i.e. under isotropy Gaussianity and independence are equivalent; indeed, in [5], it is proved the following

Proposition 1 *For an isotropic random field $T(x)$ satisfying $\mathbb{E}|a_{l.}|^2 < \infty$, coefficients $(a_{l0}, a_{l1}, ..., a_{ll})$ are independent if and only if they are Gaussian.*

Further constraints that must be satisfied by the joint moments and cumulants of the spherical harmonic coefficients under isotropy are discussed for instance by

[45] and [59]. In particular, in these references it is shown that, for isotropic random fields with finite third-order moment, the following identity must hold

$$\begin{aligned} & \mathbb{E}\left[a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\right] \\ = & \left[\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}\right]^{1/2} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 b_{l_1l_2l_3}\,, \end{aligned}$$

where $b_{l_1l_2l_3}$ is known as the reduced bispectrum of the random field. This is a special case of a more general characterization of joint moments of arbitrary order. In terms of statistical applications, a natural way to investigate the possible existence of non-Gaussianities in a spherical random field is hence the estimation of the reduced bispectrum, an idea which has been investigated for instance by [45] and [56, 57] et al. In particular, it is immediate to focus on the statistics

$$I_{l_1l_2l_3} = \sum_{m_1m_2m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \left[a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\right]\,,$$

for which we have immediately

$$\begin{aligned} \mathbb{E}I_{l_1l_2l_3} &= \sum_{m_1m_2m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} E\left[a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\right] \\ &= \sum_{m_1m_2m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 \left[\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}\right]^{1/2} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} b_{l_1l_2l_3} \\ &= \left[\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}\right]^{1/2} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} b_{l_1l_2l_3}\,, \end{aligned}$$

in view of the orthogonality properties of Clebsch-Gordan matrices. Moreover, we can show that $I_{l_1 l_2 l_3}$ is an invariant statistic with respect to the action of $SO(3)$; indeed

$$\begin{aligned} a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} &= \prod_{\alpha=1}^{3} \int_{S^2} T(g x_\alpha) \sum_{m'_\alpha = -l_\alpha}^{l_\alpha} D^{l_\alpha}_{m_\alpha m'_\alpha}(g)\, \overline{Y_{l_\alpha m'_\alpha}}(x_\alpha) dg x_\alpha \\ &= \int_{S^2} \prod_{\alpha=1}^{3} T(g x_\alpha) \prod_{\alpha=1}^{3} \sum_{m'_\alpha = -l_\alpha}^{l_\alpha} D^{l_\alpha}_{m_\alpha m'_\alpha}(g)\, \overline{Y_{l_\alpha m'_\alpha}}(x_\alpha) dg x_\alpha \\ &= \int_{S^2} \prod_{\alpha=1}^{3} T(x_\alpha) \prod_{\alpha=1}^{3} \sum_{m'_\alpha = -l_\alpha}^{l_\alpha} D^{l_\alpha}_{m_\alpha m'_\alpha}(g)\, \overline{Y_{l_\alpha m'_\alpha}}(x_\alpha) dx_\alpha \\ &= \sum_{m'_1 m'_2 m'_3} D^{l_1}_{m'_1 m_1}(g) D^{l_2}_{m'_2 m_2}(g) D^{l_3}_{m'_3 m_3} a_{l_1 m'_1} a_{l_2 m'_2} a_{l_3 m'_3} \ . \end{aligned}$$

The third equation follows from isotropy of T and the invariance properties of the Haar measure under a rotation $g \in SO(3)$, $dg x_\alpha = dx_\alpha$. Hence we have

$$\sum_{m_1 m_2 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3}$$

$$\begin{aligned} &= \sum_{m_1 m_2 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \sum_{m'_1 m'_2 m'_3} D^{l_1}_{m'_1 m_1}(g) D^{l_2}_{m'_2 m_2}(g) D^{l_3}_{m'_3 m_3}(g) \left\{ a_{l_1 m'_1} a_{l_2 m'_2} a_{l_3 m'_3} \right\} \\ &= \sum_{m'_1 m'_2 m'_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} a_{l_1 m'_1} a_{l_2 m'_2} a_{l_3 m'_3} \ , \end{aligned}$$

as required, in view of the identity

$$\sum_{m_1 m_2 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} D^{l_1}_{m'_1 m_1}(g) D^{l_2}_{m'_2 m_2}(g) D^{l_3}_{m'_3 m_3}(g) = \begin{pmatrix} l_1 & l_2 & l_3 \\ m'_1 & m'_2 & m'_3 \end{pmatrix} \ .$$

2.5 Some Background on the Existing Literature for Wavelet Analysis on the Sphere

In the previous Sections we have discussed at length Fourier analysis on $\mathbb{S}^2$, which is based on expansions of the random fields into a spherical harmonics basis. As it is always the case with Fourier analysis, this ensures excellent localization properties in harmonic space, while at the same time no form of localization can be established in the real space. The latter can make up a serious problem for experiments when the sky is only partially observed, as it is typically the case with CMB experiments (for instance, a vast region of CMB is masked by the presence of the Milky Way emissions). Moreover, in many cases of interest it is extremely important to be able to localize cosmological features over the celestial sphere - for instance, when testing for anisotropies/asymmetries, a topic which has become very relevant after the latest data releases (see [28, 41, 70]). From the mathematical point of view, it is well-known that Fourier analysis lacks in general robustness properties when considering functional norms other than L^2, for instance when one is interested in general approximation properties over L^p spaces. For all these reasons, it is natural to expect that wavelet techniques may prove extremely useful when dealing with spherical random fields. As well-known, wavelets have experienced an enormous success over the last two decades, with an enormous range of applications in physics, signal processing and applied mathematics in the last two decades (see for instance [43, 20, 42] for some standard references). The basic idea of the wavelet transform is well-known,

i.e. to decompose a signal locally into contributions living at different frequencies. This is a marked contrast with the Fourier components, which are sinusoidal waves repeating themselves indefinitely and do not have any local concentration in real domain. In this section, we will mainly follow the definitions and notations in [3, 43]; for spherical wavelets, we will focus especially on the recent papers [2, 64, 89].

2.5.1 A Quick Review on Continuous Wavelet Transforms

It is customary to distinguish two different versions of the wavelet transform (WT), the continuous (CWT) and the discrete (DWT). In this thesis, we shall only be concerned with discrete wavelets; however for historical reasons we find it convenient to provide a very quick discussion on the continuous case, first in one dimension, then in two and higher dimensions.

As we said before, the basic idea behind wavelet analysis is to aim at a space-frequency representation. As a consequence, we shall typically need two parameters, the first (a, say) which characterizes the frequency, while the second (b, say) characterizes localization in the real domain. Let $s(x)$ be a square integrable function $s \in L^2(R, dx)$ (a finite energy signal, in the language of signal analysis). A general space-frequency transform of the signal s will take the form:

$$s(x) \longmapsto S(b,a) = |a|^{-1/2} \int_{-\infty}^{\infty} s(x)\, \overline{\psi\left(\frac{x-b}{a}\right)} dx\,, \tag{2.14}$$

where $a \neq 0$ is a scale parameter, $b \in \mathbb{R}$ a translation parameter, and the function ψ, the analyzing wavelet, is assumed to be well localized both in the space (or time)

domain and in the frequency domain. In addition ψ must satisfy the following admissibility condition, which guarantees the invertibility of the WT:

$$c_\psi = 2\pi \int_{-\infty}^{\infty} \frac{\left|\widehat{\psi}(\xi)\right|^2}{|\xi|} d\xi < \infty \, .$$

In most cases, this condition may be reduced to the (only slightly weaker) requirement that ψ has zero mean:

$$\widehat{\psi}(0) = 0 \Longleftrightarrow \int_{-\infty}^{\infty} \psi(x)\, dx = 0 \, .$$

Intuitively, it express the fact that a wavelet must be an oscillating function, real or complex ("little wave").

Notice that, instead of (2.14), which defines the WT as the scalar product of the signal s with the transformed wavelet $\psi_{b,a} := |a|^{-1/2} \psi\left(\frac{x-b}{a}\right)$, $S(b,a)$ may also be seen as the convolution of s with the scaled, flipped and conjugated wavelet $\psi_a^{\#}(x) := |a|^{-1/2} \overline{\psi(-x/a)}$:

$$S(b,a) = \left(\psi_a^{\#} * s\right)(b) = \int_{-\infty}^{\infty} s(x)\, \psi_a^{\#}(b-x)\, dx \, .$$

In other words, the CWT acts as a filter with a function of zero mean.

Heuristically, the idea behind continuous wavelet transforms is to ensure that both ψ and its Fourier transform $\widehat{\psi}$ enjoy "good" localization properties; as well-known, there is a sort of trade-off between properties in the real and harmonic domains, provided by the so-called Fourier (or Heisenberg, in the physical literature) uncertainty principle [32], which implies the products of the (suitably defined) variances of ψ, $\widehat{\psi}$ is bounded from below. More precisely, one defines the centers of

gravity (which may in fact be normalized to zero by a suitable redefinition of the coordinates):

$$x_0 = \int_{-\infty}^{\infty} x \left|\psi\left(x\right)\right|^2 dx,\ \xi_0 = \int_{-\infty}^{\infty} \xi \left|\widehat{\psi}\left(\xi\right)\right|^2 d\xi,$$

and the corresponding variances

$$\left(\triangle\ x\right)^2 = \left\|\psi\right\|^{-2} \int_{-\infty}^{\infty} \left(x - x_0\right)^2 \left|\psi\left(x\right)\right|^2 dx;$$
$$\left(\triangle\ \xi\right)^2 = \left\|\psi\right\|^{-2} \int_{-\infty}^{\infty} \left(\xi - \xi_0\right)^2 \left|\widehat{\psi}\left(\xi\right)\right|^2 d\xi;$$

it is then possible to prove the lower bound [32]

$$\triangle\ x\ \triangle\ \xi \geq 1/2\ .$$

Let us now focus on the two- and higher-dimensional wavelet analysis. In one-dimension, the CWT (2.14) amounts to projecting the function onto the wavelet $\psi_{b,a}$, obtained by the action of translations and dilations of the *mother* wavelet ψ. Thus the transform is fully determined by these elementary operations on $\mathbb{R}$. Likewise, to discuss CWT in higher dimension, a good starting point is to consider first the elementary operations we want to apply to our signals.

For instance, given a square integrable function $s \in L^2\left(\mathbb{R}^n, d^n\overrightarrow{x}\right)$, as a transformation group we may focus on the similitude group, which taking the same notation as in [2] can be written as $SIM(n) = \mathbb{R}^n \rtimes \left(\mathbb{R}_*^+ \times SO(n)\right)$. From the abstact algebra point of view, this group has a (unique) natural unitary irreducible representation in $L^2\left(\mathbb{R}^n, d^n\overrightarrow{x}\right)$, given by [2, 3]

$$s_{\overrightarrow{b},a,\varrho}\left(\overrightarrow{x}\right) := a^{-n/2} s\left(a^{-1}\varrho^{-1}\left(\overrightarrow{x} - \overrightarrow{b}\right)\right),$$

where $\overrightarrow{b} \in \mathbb{R}^n$ is a translation, $a > 0$ is a global dilation and $\varrho \in SO(n)$ is a rotation. This representation is square integrable, with admissibility condition

$$\int_{\mathbb{R}^n} \left|\widehat{\psi}\left(\overrightarrow{k}\right)\right|^2 \left|\overrightarrow{k}\right|^{-n} d^n\overrightarrow{k} < \infty,$$

implying as usual to the necessary condition $\int_{\mathbb{R}^n} \psi\left(\overrightarrow{x}\right) d^n\overrightarrow{x} = 0$. The $n-$dimensional wavelet transform is then as before,

$$S\left(\overrightarrow{b}, a, \varrho\right) = a^{-n/2} \int_{\mathbb{R}^n} s\left(\overrightarrow{x}\right) \overline{\psi\left(a^{-1}\varrho^{-1}\left(\overrightarrow{x} - \overrightarrow{b}\right)\right)} d^n\overrightarrow{x}. \tag{2.15}$$

2.5.2 Some Examples of Spherical Wavelets

Over the last decade, a number of attempts have been spent to extend wavelet constructions on the sphere. We shall not even try to review all these efforts, and we refer instead to the existing literature; for instance, *The Journal of Fourier Analysis and Applications* published two special issues *Volume 13, No.4 and No.5 (2007)* entirely focussed on these themes. Some attempts have just tried to lift to the sphere the most classical construction on $\mathbb{R}$; a classical example is provided by the Haar system, which was extended to the sphere by Sweldens [82]**,** and applied to CMB data analysis by Tenorio et al. [83]. More precisely, it is assumed that we have a grid partition of the sphere (such as the one provided by HEALPix ([39]), highly popular among the cosmologists); consider the scaling (”father”) function $\phi_{j,k}$ and three wavelet (”mother”) functions $\psi_{m,j,k}$ at each resolution level j and position on the grid k, such that each elements of the partition (”pixel”) $S_{j,k}$ is divided into four smaller

ones S_{j+1,k_0},..., S_{j+1,k_3} at resolution $j+1$. The scaling and wavelet function are given by (see also [62])

$$\phi_{j,k}(x) = \begin{cases} 1 & if\ x \in S_{j,k} \\ 0 & otherwise, \end{cases}$$

$$\psi_{1,j,k} = \frac{\phi_{j+1,k_0} + \phi_{j+1,k_2} - \phi_{j+1,k_1} - \phi_{j+1,k_3}}{4\mu_{j+1}},$$

$$\psi_{2,j,k} = \frac{\phi_{j+1,k_0} + \phi_{j+1,k_1} - \phi_{j+1,k_2} - \phi_{j+1,k_3}}{4\mu_{j+1}},$$

$$\psi_{3,j,k} = \frac{\phi_{j+1,k_0} + \phi_{j+1,k_3} - \phi_{j+1,k_1} - \phi_{j+1,k_2}}{4\mu_{j+1}}.$$

This proposal is simple and has enjoyed some success among practitioners, but is not difficult to find many reasons why it could not be considered satisfactory, indeed the Haar system has now been largely abandoned even in the case of the real line. In other fields of applications, for instance geomagnetic or gravity field modeling, Poisson wavelets have been quite often considered. More precisely, and with the same notations as in [44], we denote by Ω_R the sphere of radius R, $\Omega_R = \{x : x/R \in \mathbb{S}^2\}$, and by $Int\Omega_R$ the interior and $Ext\Omega_R$ the exterior, respectively

$$Int\,(Ext)\,\Omega_R = \left\{x \in \mathbb{R}^3, |x| < (>)\, R\right\}.$$

We then define the exterior Poisson wavelet of degree d as follows:

$$W_x^{ext,d}(y) = \frac{R}{|y|}\sum_{l=0}^{\infty} l^d \left(\frac{|x|}{|y|}\right)^l (2l+1)\, P_l\left(\langle \widehat{x}, \widehat{y}\rangle\right), x \in Int\Omega_R,\ y \in \overline{Ext\Omega_R},$$

where for $z \neq 0$, we write $\widehat{z} = z/|z|$. P_l is the Legendre Polynomial of order l defined by (2.9).

The interior Possion wavelet is defined analogously,

$$W_x^{int,d}(y) = \frac{R}{|x|}\sum_{l=0}^{\infty} l^d \left(\frac{|y|}{|x|}\right)^l (2l+1)\, P_l\left(\langle \widehat{x}, \widehat{y}\rangle\right), x \in Ext\Omega_R,\ y \in \overline{Int\Omega_R}.$$

Very little is known on the mathematical properties of this proposal. In the cosmological framework, a huge success has been met instead by the so-called Spherical Mexican-hat wavelets (SMHW), see for instance [16, 62, 85, 86]; a detailed treatment may be found in [2, 64, 89]. A naive way to introduce SMHW is to consider the stereographic projection (on the tangent plane at each point of $\mathbb{S}^2$) of the usual Mexican-hat wavelets on the Euclidean plane,

$$\psi(x;R) = \frac{1}{(2\pi)^{1/2} R}\left(2 - \frac{|x|^2}{R^2}\right)\exp\left(-\frac{|x|^2}{2R^2}\right).$$

Standard manipulations yield the the following close form expression, depending on some positive scale parameter R,

$$\psi_R(\theta) \propto \left(1+\tan^2(\theta/2)\right)^2 \left(1 - 2R^2\tan^2(\theta/2)\right)\exp\left\{-2R^2\tan^2(\theta/2)\right\}.$$

A more refined derivation can be entertained as follows. As mentioned earlier, it is natural to introduce a CWT by considering translations and dilations of a single function. Of course, the sphere is a compact manifold, and hence it does not admit any form of global dilations. Several authors have addressed this problem, with various techniques. The solution proposed by [2] is the following: consider the 2-sphere $\mathbb{S}^2$, with polar spherical coordinates $\zeta = (\theta, \varphi)$. As usual, let us focus on square integrable functions on the sphere, $s \in L^2(S^2, d\mu)$, where $d\mu(\zeta) \equiv d\mu(\theta,\varphi) = \sin\theta d\theta d\varphi$ is the usual (rotation invariant) measure on $\mathbb{S}^2$.

The first step for constructing a CWT on $\mathbb{S}^2$ is to identify the natural operations on such functions. These are of two types:

(i) Motions or displacements, given by elements of the rotation group $SO(3)$, which indeed acts transitively on $\mathbb{S}^2$, and on the homogeneous space $\mathbb{S}^2 \backsimeq SO(3)/SO(2)$.

(ii) Dilations, that may be derived in two steps. First, dilations around the North Pole are obtained by considering usual dilations in the tangent plane and lifting them to $\mathbb{S}^2$ by inverse stereographic projection from the South Pole. This gives:

$$D_a^{(N)}(\theta, \varphi) = (\theta_a, \varphi), \text{ with } \tan\frac{\theta_a}{2} = a \tan\frac{\theta}{2}.$$

Then a dilation around any other point $\zeta \in \mathbb{S}^2$ is obtained by moving ζ to the North Pole by a rotation $\varrho \in SO(3)$, performing a dilation $D_a^{(N)}$ as before and going back by the inverse rotation: $D_a^{(N)} = \varrho^{-1} D_a^{(N)} \varrho$. Clearly the dilations act also transitively on $\mathbb{S}^2$.

The wavelets that are so-derived can be viewed, from the group theoretic point of view, as unitary irreducible representation of group $SIM(3) = \mathbb{R}^3 \rtimes (\mathbb{R}_*^+ \times SO(3))$ acting on the Hilbert space $L^2(\mathbb{S}^2, d\mu)$. This can lead to a link with the theory of so-called coherent states, which has important connections with quantum mechanics - we refer to [50] for more details. We shall not discuss further these efforts, as the direction we shall undertake in this thesis is radically different: i.e. focus on discrete

wavelets and their stochastic properties, which have not at all been investigated (to the best of our knowledge) in the framework of continuous transforms.

2.6 Second-generation Spherical Wavelets: Needlets

Needlets are a new kind of second-generation spherical wavelets, which were introduced in the Functional Analysis literature by [66, 67], whereas the investigation of their properties from a stochastic point of view is due to [6, 7] and [8]; see also [30, 31, 33, 49]. They were shown to make up a tight frame (see below) with excellent localization properties in both the real and the harmonic domains. In this section we shall briefly recall the main features of the needlets construction. We need first to introduce some notation and definitions, which are largely identical to those in [7, 67].

2.6.1 Frames

Let us start with the definition of *frame.*

Definition 1 *A countable family of vectors* $\{\psi_n\}$ *in a Hilbert space* $\mathcal{L}$ *is called a (discrete) frame if there are two positive constants A, B, with* $0 < A \leq B < \infty$, *such that*

$$A \|f\|^2 \leq \sum_{n=1}^{\infty} |\langle \psi_n, f \rangle|^2 \leq B \|f\|^2, \ \forall \ f \in \mathcal{L}. \tag{2.16}$$

The two constants A, B are the frame bounds. If $A = B \geq 1$, the frame is said to be tight. if $A = B = 1$, and $\|\psi_n\| = 1, \forall\ n$, the set $\{\psi_n\}$ is simply an orthonormal basis.

The properties of a frame are best discussed in terms of the frame operator $F : \mathcal{L} \longrightarrow \ell^2$, defined by

$$F : f \longmapsto \{\langle \psi_n, f \rangle\}.$$

As we can see, the upper bound in (2.16) simply means that F is a bounded operator, whereas the left inequality guarantees the numerical stability for the recovery of the signal f from its frame coefficients $\{\langle \psi_n, f \rangle\}$ – in other words, it gives an estimation of the inverse operator F^{-1}. As for the frame bounds A, B, they measure the redundancy of the representation of the signal in terms of its coefficients. Further information about properties between frame and wavelets may be found in [3].

2.6.2 Voronoi Cells

In the sequel, we shall find it often useful to partition the sphere $\mathbb{S}^2$ into some special disjoint subsets with roughly the same dimension, called Voronoi Cells. In a sense to be made rigorous later, we need first to introduce some notation and definitions, which are largely identical to those in [7, 67]. Given any two positive sequences $\{a_j\}, \{b_j\}$, we write $a_j \approx b_j$ if there exist a real number $c > 0$ such that $c^{-1}a_j \leq b_j \leq ca_j$ for all j. The standard (closed and open) balls in $\mathbb{S}^2$ are given as always by $B(a, \alpha) = \{x,\ d(a, x) \leq \alpha\}$, $B^{\circ}(a, \alpha) = \{x,\ d(a, x) < \alpha\}$. For a general

$A \subset \mathbb{S}^2$ we will denote by $|A|$ the spherical measure of A. Now fix $\epsilon > 0$ and $x_1, ...x_N$ in $\mathbb{S}^2$ such that $\forall i \neq j$, $d(x_i, x_j) > \epsilon$ and the set $\{x_1, ..., x_N\} = \Xi_\epsilon$ is maximal in the sense that: $\forall x \in \mathbb{S}^2$, $d(x, \Xi_\epsilon) \leq \epsilon$, $\cup_{x_i \in \Xi_\epsilon} B(x_i, \epsilon) = \mathbb{S}^2$ and $\forall i \neq j$, $B(x_i, \epsilon/2) \cap B(x_j, \epsilon/2) = \emptyset$. It follows from Lemma 5 in [7] that

$$\frac{4}{\epsilon^2} \leq N \leq \frac{4}{\epsilon^2}\pi^2 \tag{2.17}$$

We can now recall the definition of Voronoi cells, for all $x_i \in \Xi_\epsilon$:

Definition 2 *Let Ξ_ϵ be a maximal $\epsilon-$ net. The associated family of Voronoi cells is defined by:*

$$\mathcal{V}(x_i) = \{x \in \mathbb{S}^2, \ \ \forall j \neq i, \ \ d(x, x_i) \leq d(x, x_j)\} .$$

We recall that $B(x_i, \epsilon/2) \subset \mathcal{V}(x_i) \subset B(x_i, \epsilon)$. Also, if two Voronoi cells are adjacent, i.e. $\mathcal{V}(x_i) \cap \mathcal{V}(x_j) \neq \emptyset$, then by necessity $d(x_i, x_j) \leq 2\epsilon$. It is proved in [7] that there are at most $6\pi^2$ adjacent cells to any given cell.

2.6.3 The Cubature Points and Cubature Weights

For the construction of needlets, we should first start to define $\mathcal{K}_l$ as the space of the restrictions to the sphere $\mathbb{S}^2$ of polynomials of degree less than l. The next ingredient are the set of *cubature points* and *cubature weights;* indeed, it is now a standard result (see for instance [65]) that for all $j \in \mathbb{N}$, there exists a finite subset $\mathcal{X}_j$ of $\mathbb{S}^2$ and positive real numbers $\lambda_{jk} > 0$, indexed by the elements of $\mathcal{X}_j$, such that

$$\forall f \in \mathcal{K}_l, \quad \int_{\mathbb{S}^2} f(x)dx = \sum_{\xi_{jk} \in \mathcal{X}_j} \lambda_{jk} f(\xi_{jk}) . \tag{2.18}$$

Given a fixed $B > 1$, we shall denote by $\{\xi_{jk}\}$ the cubature points corresponding to the space $\mathcal{K}_{[3B^{j+1}]}$, where $[.]$ represents as usual integer part. It is known that $\{\mathcal{X}_j\}_{j=0}^{\infty}$ can be taken s.t. the cubature points for each j are almost uniformly ϵ_j-distributed with $\kappa^{-1}B^{-j} \leq \epsilon_j \leq \kappa B^{-j}$, i.e. $\epsilon_j \approx B^{-j}$, and the coefficients $\{\lambda_{jk}\}$ are such that $c^{-1}B^{-2j} \leq \lambda_{jk} \leq cB^{-2j}$, where c is a finite real number, and $card\,\{\mathcal{X}_j\} \approx B^{2j}$.

2.6.4 Construction of Needlets

We are now in the position to start the construction of the needlets system. Let ϕ be a C^{α} $(4 \leq \alpha \leq \infty)$ function supported in $|\xi| \leq 1$, such that $0 \leq \phi(\xi) \leq 1$ and $\phi(\xi) = 1$ if $|\xi| \leq B^{-1}$, $B > 1$. Following again [66, 67], we define

$$b^2(\xi) = \phi(\frac{\xi}{B}) - \phi(\xi) \geq 0 \text{ so that } \forall |\xi| \geq 1, \sum_j b^2(\frac{\xi}{B^j}) = 1\,. \tag{2.19}$$

It is immediate to verify that

(i) $b^2(\cdot)$ has support in $[\frac{1}{B}, B]$, and hence $b(\frac{l}{B^j})$ has support in $l \in [B^{j-1}, B^{j+1}]$;

(ii) we have $\sum_{j=0}^{\infty} b^2(\frac{l}{B^j}) \equiv 1$ for all $l \geq 1$.

To explore more explicit recipes for the construction of $b(\cdot)$, see Appendix B.

The spherical needlets is defined in the following way

$$\varphi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_l b(\frac{l}{B^j}) \sum_{m=-l}^{l} Y_{lm}(\xi_{jk}) \overline{Y_{lm}}(x)\,. \tag{2.20}$$

The (possibly random) spherical needlet coefficients are then defined as

$$\beta_{jk} = \int_{\mathbb{S}^2} T(x)\varphi_{jk}(x)dx = \sqrt{\lambda_{jk}} \sum_l b(\frac{l}{B^j}) \sum_{m=-l}^{l} a_{lm} Y_{lm}(\xi_{jk})\,. \tag{2.21}$$

We have immediately

$$\sum_k \beta_{jk}\sqrt{\lambda_{jk}} = 0\ , \tag{2.22}$$

i.e., the (weighted) sample mean of the needlets coefficients is identically zero at any level j. The proof is trivial, because

$$\begin{aligned}\sum_k \beta_{jk}\sqrt{\lambda_{jk}} &= \sum_{l=B^{j-1}}^{B^{j+1}}\sum_{m=-l}^{l} b(\frac{l}{B^j})a_{lm}\left[\sum_k \lambda_{jk}Y_{lm}(\xi_{jk})\right] \\ &= \sum_{l=B^{j-1}}^{B^{j+1}}\sum_{m=-l}^{l} b(\frac{l}{B^j})a_{lm}\left[\int_{\mathbb{S}^2} Y_{lm}(x)dx\right] = 0\ .\end{aligned}$$

In the stochastic case, the variance of the needlets coefficients is given by

$$\begin{aligned}\mathbb{E}\beta_{jk}^2 &= \lambda_{jk}\sum_{l=B^{j-1}}^{B^{j+1}} b^2(\frac{l}{B^j})C_l\sum_{m=-l}^{l} Y_{lm}(\xi_{jk})\overline{Y_{lm}(\xi_{jk})} = \lambda_{jk}\sum_{l=B^{j-1}}^{B^{j+1}} b^2(\frac{l}{B^j})C_l\frac{2l+1}{4\pi}P_l(\cos 0) \\ &= \lambda_{jk}\sum_{l=B^{j-1}}^{B^{j+1}}\frac{2l+1}{4\pi}b^2(\frac{l}{B^j})C_l =: \sigma_{jk}^2 > 0\ .\end{aligned}$$

Note that we have $\sigma_{jk}^2 \approx: \sigma_j^2$ uniformly over k, where

$$\sigma_j^2 := \frac{4\pi}{N_j}\sum_{l=B^{j-1}}^{B^{j+1}}\frac{2l+1}{4\pi}b^2(\frac{l}{B^j})C_l\ ,\ N_j = card\,\{\mathcal{X}_j\}\ .$$

From now on, we shall typically focus on the normalized needlets coefficients, defined as $\widehat{\beta}_{jk} := \beta_{jk}/\sigma_j$.

Needlets enjoy a series of very important properties, which we recall below:

1) Needlets make up a tight frame system. These ensures in particular that the following reconstruction formula holds:

$$T(x) = \sum_{jk} \beta_{jk} \varphi_{jk}(x) \ .$$

This property (see [66, 67]) is crucial for most applications and it is not shared by other existing spherical wavelets frames.

2) Needlets do not rely on any tangent plane approximation, while on the contrary they are naturally embedded into the intrinsic geometry of the sphere.

3) Needlets are quasi-exponentially (i.e.faster than any polynomial) concentrated in pixel space, i.e. their tails decay very fast .This main localization property is first established in [66], where it is shown that for any $M \in \mathbb{N}$ there exists a constant $c_M > 0$, s.t. for every $\xi \in \mathbb{S}^2$:

$$\left|\varphi_{jk}(\xi)\right| \leq \frac{c_M B^j}{(1 + B^j \arccos\langle \xi_{jk}, \xi \rangle)^M} \text{ uniformly in } (j,k) \ .$$

In the appendix B, we provide some computations to establish (upper bounds on) the values of c_M, for alternative choices of the function $b(.)$.

4) Needlets are compactly supported in the frequency domain. Indeed, their support is exactly localized on a finite number of multipoles, its width is explicitly known and can be specified as an input parameter by the practitioner. This property is again crucial for physical applications (see [75, 61]) and lies at the heart of the uncorrelation results we shall discuss below.

5) Random needlets coefficients can be shown to be asymptotically uncorrelated (and hence, in the Gaussian case, independent) at any fixed angular distance, as the frequency increases (see [6, 7] for detail). As we shall see further on in this thesis, this is the crucial property to be exploited in several statistical procedures: in some sense, it allows to treat normalized needlets coefficients as (approximately) a sample of independent and identically distributed coefficients on small scales, at least under the Gaussianity assumption.

To investigate the last point further, we need first to introduce the same, mild regularity conditions on the angular power spectrum C_l of the random field $T(x)$ as in [6, 7, 15].

Condition A The random field $T(x)$ is Gaussian and isotropic with angular power spectrum such that, for all $B > 1$, there exist $\alpha > 2$, and $\{g_j(.)\}_{j=1,2,...}$ a sequence of functions such that

$$C_l = l^{-\alpha} g_j(\frac{l}{B^j}) > 0 \text{ , for } B^{j-1} < l < B^{j+1} \text{ , } j = 1, 2, ... \tag{2.23}$$

where $c_0^{-1} \leq g_j \leq c_0$ for all $j \in \mathbb{N}$, and $\sup_j \sup_{B^{-1} \leq u \leq B} |\frac{d^r}{du^r} g_j(u)| \leq c_r$, $c_0, c_1, ..., c_M > 0$, $M \in \mathbb{N}$.

Condition (2.23) entails a weak smoothness requirement on the behaviour of the angular power spectrum, which is trivially satisfied by some cosmologically relevant models (where the angular power spectrum usually behaves as an inverse polynomial, see again [24], pp.243-244). For instance, considering positive constants d_j,

$j = 1, ..., p$, $p > 2$, for all $B > 1$ condition A holds for

$$\begin{aligned} C_l &= \frac{1}{\sum_{k=0}^{p} d_k l^k} = l^{-p} \frac{l^p}{d_0 + d_1 l + ... d_p l^p} \\ &= l^{-p} \frac{(l/B^j)^p}{\sum_{k=0}^{p} d_k B^{jk-jp} (l/B^j)^k} = l^{-p} g_j(\frac{l}{B^j}) \text{ , for } B^{j-1} < l < B^{j+1} \text{ , } j = 1, 2, ... \\ g_j(u) &: = \frac{u^p}{\sum_{k=0}^{p} d_k B^{jk-jp} u^k} \text{ .} \end{aligned}$$

The uncorrelation property can then be stated as follows ([7]): under Condition A

$$|Corr(\beta_{jk}, \beta_{jk'})| \leq \frac{c_M}{(1 + B^j d(\xi_{jk}, \xi_{jk'}))^M}, \tag{2.24}$$

where $d(\xi_{jk}, \xi_{jk'}) = \arccos(\langle \xi_{jk}, \xi_{jk'} \rangle)$. In words, (2.24) is stating that as the frequency increases, spherical needlets coefficients are asymptotically uncorrelated, for any given angular distance. As mentioned before, in a loose sense (2.24) is stating that it is possible to derive an infinitely growing array of asymptotically independent "observations" (the needlets coefficients) out of a single realization of a Gaussian field on a compact domain.

An important fact to remark is that, somewhat counterintuitively, (2.24) is *not* by any means a consequence of the localization properties of the needlets frame. Indeed, in Chapter 4 we shall investigate the stochastic properties of a modified form of needlets (so-called Mexican needlets), which may have better localization properties than standard needlets, but for which uncorrelation holds only under further conditions (Mexican needlets were first introduced by [34] and are then considered also in [35] and [63]). It is easy to construct also other examples of spherical frames hav-

ing bounded support in real space, whereas the corresponding random coefficients are not at all uncorrelated.

Finally, we note that, from the computational point of view, needlets are not only feasible, but indeed extremely convenient. The implementation can be performed with a minimal effort by means of standard packages for the analysis of spherical random fields such as HEALPix or GLESP ([39] and [27]); we refer for further details to [61], where plots and numerical evidence on localization and uncorrelation are also provided, see also [23, 74, 75].

To conclude this Section, we recall briefly two further results that will be used often in the sequel. First, we note that for $M > 2$, $j \in \mathbb{N}, \forall\, k, k'$,([7], *Lemma 10*)

$$\sum_{\xi_{jk} \in \chi_j} \frac{1}{(1 + B^j d(\xi_{jk}, \xi_{jk'}))^M} \leq C_M \text{ some } C_M > 0\,. \tag{2.25}$$

Also ([67], *Lemma 4.8*), for some C'_M depending only on M we have the inequality

$$\sum_{\xi_{jk} \in \chi_j} \frac{1}{(1 + B^j d(\xi_{jk}, \xi_{jk'}))^M} \frac{1}{(1 + B^j d(\xi_{jk}, \xi_{jk''}))^M} \leq \frac{C'_M}{(1 + B^j d(\xi_{jk'}, \xi_{jk''}))^M}\,. \tag{2.26}$$

2.7 Diagram Formula

To complete our background, we need a quick review on the diagram formula. This is material which can now be found at a textbook level (see for instance [73, 80]); nevertheless, we need a brief overview to fix notation. Denote by H_q the $q-$th order

Hermite polynomials, defined as

$$H_q(u) = (-1)^q e^{u^2/2} \frac{d^q}{du^q} e^{-u^2/2}.$$

We now introduce *diagrams*, which are basically mnemonic devices for computing the moments and cumulants of polynomial forms in Gaussian random variables. Our notation is the same as for instance in [56, 57]. Let p and $l_{ij} = 1, ..., p$ be given integers. A diagram γ of order $(l_1, ..., l_p)$ is a set of points $\{(j, l) : 1 \leq j \leq p, 1 \leq l \leq l_j\}$ called vertices, viewed as a table $W = \overrightarrow{l_1} \otimes \cdots \otimes \overrightarrow{l_p}$ and a partition of these points into pairs

$$\{((j, l), (k, s)) : 1 \leq j \leq k \leq p; 1 \leq l \leq l_j, 1 \leq s \leq l_k\},$$

called edges. We denote by $\Gamma(W)$ the set of diagrams of order $(l_1, ..., l_p)$. If the order is $l_1 = \cdots = l_p = q$, for simplicity, we also write $\Gamma(p, q)$ instead of $\Gamma(W)$. We say that:

$a)$ A diagram has a flat edge if there is at least one pair $\{(i, j)(i', j')\}$ such that $i = i'$; we write Γ_F for the set of diagrams that has at least one flat edge, and $\Gamma_{\overline{F}}$ otherwise.

$b)$ A diagram is *connected* if it is not possible to partition the rows $\overrightarrow{l_1} \cdots \overrightarrow{l_p}$ of the table W into two parts, i.e. one cannot find a partition $K_1 \cup K_2 = \{1, ..., p\}$ that, for each member V_k of the set of edges $(V_1, ..., V_r)$ in a diagram γ, either $V_k \in \cup_{j \in K_1} \overrightarrow{l_j}$, or $V_k \in \cup_{j \in K_2} \overrightarrow{l_j}$ holds; we write Γ_C for connected diagrams, and $\Gamma_{\overline{C}}$ otherwise.

c) A diagram is paired if, considering any two sets of edges $\{(i_1, j_1)(i_2, j_2)\}$ $\{(i_3, j_3)(i_4, j_4)\}$, then $i_1 = i_3$ implies $i_2 = i_4$; in words, the rows are completely coupled two by two. We write Γ_P for the set of paired diagrams, and $\Gamma_{\overline{P}}$ otherwise.

Example 1

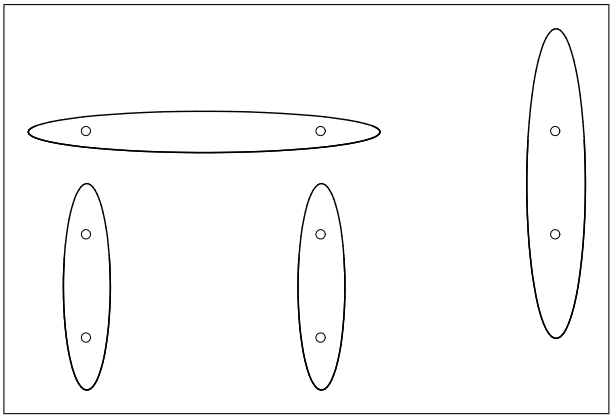

fig1

In fig1, The diagram has a flat edge $\{(1,1)\,(1,2)\}$ *, and also it is connected as the first row is connected with the second row via edge* $\{(1,3)\,(2,3)\}$ *, while the second row is connected with the third row via edges* $\{(2,1)\,(3,1)\}$ *and* $\{(2,2)\,(3,2)\}$ *.*

Example 2

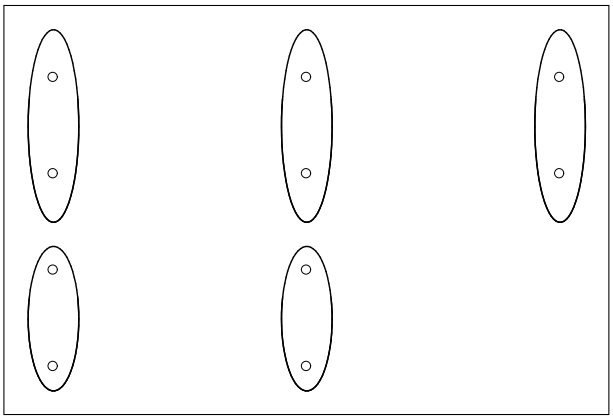

fig2

The diagram in fig2 is paired, as obviously the first row and second row are coupled while the third and fourth are also coupled.

Here we show how a multigraph can be derived from a diagram above. This representation of diagrams (Gaussian) are always used in the computation of moments and cumulants of polynomials in Gaussian random variables (see [56, 73]).

We denote the multigraph as $\widehat{\Gamma}(W)$. It can be obtained from $\Gamma(W)$ as follows:

1. Identify each row of $\Gamma(W)$ with a vertex of $\widehat{\Gamma}(W)$, in such a way that the ith row of $\Gamma(W)$ corresponds to the ith vertex v_i of $\widehat{\Gamma}(W)$.

2. Draw an edge linking v_{i_1} and v_{i_2} for every pair $\{(i_1, j_1)(i_2, j_2)\} \in \Gamma(W)$ such that (i_1, j_1) (i_2, j_2) each comes from the i_1th row and i_2th row, respectively.

Example 3 *(i)The multigraph obtained from the diagram in fig1 is*

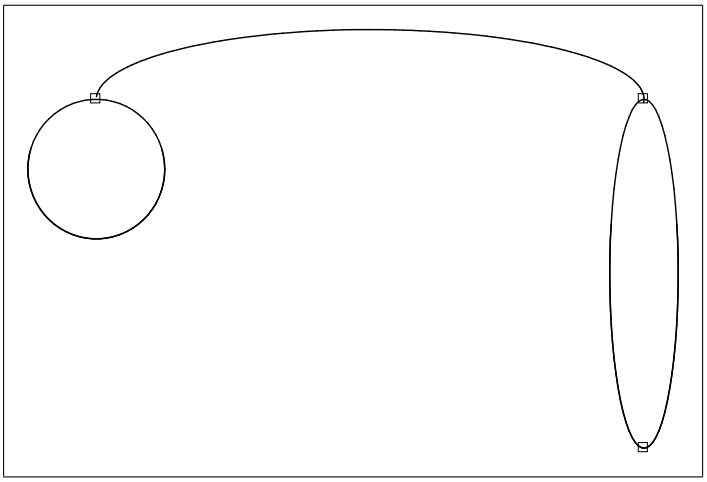

fig3

(ii) The multigraph comes from the diagram in fig2 is

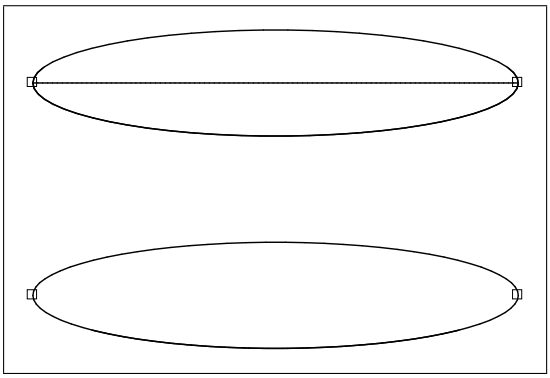

fig4

The following result (see [73] and [80]) plays a key role in evaluating the higher moments of the Gaussian subordinated variables..

Proposition 2 (Diagram Formula) *Let* $(X_1, ..., X_p)$ *be a centered Gaussian vector, and let* $\gamma_{ij} = E[X_i X_j], i, j = 1, ..., p$ *be their covariances,* $H_{l_1}, ..., H_{l_p}$ *be Hermite polynomials of degree* $l_1, ..., l_p$ *respectively. Let* L *be a table consisting of* p *rows* $l_1, ...l_p$*, where* l_j *is the order of Hermite polynomial in the variable* X_j*. Then*

$$\begin{aligned} E[\prod_{j=1}^{p} H_{l_j}(X_j)] &= \sum_{G \in \Gamma(l_1,...,l_p)} \prod_{1 \le i \le j \le p} \gamma_{ij}^{\eta_{ij}(G)} \\ Cum(H_{l_1}(X_1), ..., H_{l_p}(X_p)) &= \sum_{G \in \Gamma_c(l_1,...,l_p)} \prod_{1 \le i \le j \le p} \gamma_{ij}^{\eta_{ij}(G)} \end{aligned} \tag{2.27}$$

where, for each diagram G*,* $\eta_{ij}(G)$ *is the number of edges between rows* l_i, l_j *and* $Cum(H_{l_1}(X_1), ..., H_{l_p}(X_p))$ *denotes the* p*-*th *order cumulant.*

Here the cumulant is defined by

$$Cum\,(X_1, ..., X_n) = \left(-\sqrt{-1}\right)^n \partial^n \log E \exp\left\{\sqrt{-1}\sum_{i=1}^{n} \nu_i X_i\right\} / \partial\nu_1 ... \partial\nu_n|_{\nu_1 = ... \nu_n = 0}\,, \tag{2.28}$$

where $X_j, j = 1, ..., n$ are random variables satisfying $E|X_j|^n < \infty$. It is easy to see that $Cum(X_1) = EX_1$ and $Cum(X_1, X_2) = EX_1X_2 - EX_1X_2 = Cov(X_1, X_2)$. In particular, we have $Cum(X_1, ..., X_n) = 0$, if $\{X_j\}_j$ is Gaussian and $n \geq 3$. Now by *Leonov-Shiryaev formulas* [55],

$$E[X_1...X_n] = \sum_{(V_1,...,V_r)\in\Gamma(\underbrace{1,...,1}_{n})} Cum\left(X^{\cdot V_1}\right)...Cum\left(X^{\cdot V_r}\right)$$

and in the centered Gaussian case, all the other components disappeared except the diagram $(V_1, ..., V_r)$ with $|V_1| = ... |V_r| = 2$, from which we have

$$\mathbb{E}[X_1...X_n] = \sum_{G\in\Gamma(\underbrace{1,...,1}_{n})} \prod_{1\leq i\leq j\leq p} \gamma_{ij}^{\eta_{ij}(G)}$$

For the general case, see [81] for detail.

To show how this formula can be used, we provide here a few examples.

Example 4 *Assume X, Y are two random variables following standard Normal distribution $N(0, 1)$, with covariance $\mathbb{E}[XY] = \rho$, H_m, H_n are Hermite polynomials with degrees m and n respectively, then by (2.27), we have*

$$\mathbb{E}H_m(X) H_n(Y) = m!\rho^m\delta_m^n. \tag{2.29}$$

Example 5 *Recall the Hermite polynomials of degree 2 and 4 are $H_2(X) = X^2-1$, $H_4(Y) = Y^4 - 6Y^2 + 3$, respectively, with the same random variables X, Y defined*

as in the above example. Then in view of (2.29), we obtain

$$\begin{aligned}\mathbb{E}X^2Y^4 &= \mathbb{E}H_2(X)H_4(Y) + 6\mathbb{E}H_2(X)H_2(Y) - 3\mathbb{E}H_2(X) + \mathbb{E}H_4(Y) \\ &= 12\rho^2,\end{aligned}$$

by the facts that $\mathbb{E}H_2(X)H_4(Y) = 0$, $\mathbb{E}H_2(X)H_2(Y) = 2!\rho^2$, $\mathbb{E}H_2(X) = 0$ *and* $\mathbb{E}H_4(Y) = 0$.

We refer to [69, 68] for very important recent developments in the use of diagram formula in the derivation of Central Limit Theorem and Stein bounds.

Chapter 3
The Needlets Bispectrum

[1]The purpose of this chapter is to join two different threads of the recent literature on random fields on the sphere, namely the statistical analysis of higher order angular power spectra on one hand, and the construction of second-generation wavelets on the sphere on the other hand. More precisely, the aim is to borrow ideas from the bispectrum and the needlets literature to propose and analyze a needlets bispectrum, where the random coefficients in the needlets expansion are combined in a similar way to the bispectrum construction. The purpose is to obtain a procedure which mimics the ability of the bispectrum to search for non-Gaussianity at the most efficient combination of frequencies, at the same time providing a much more robust construction in the presence of missing data, as typical of the needlets. The plan of the chapter is as follows: In Section 1, we introduce the needlets bispectrum and we establish a central limit theorem, in the high resolution sense. In Section 2, we go on to establish a functional central limit theorem for the integrated needlets bispectrum; in Section 3, we provide some preliminary discussion on the behaviour under non-Gaussian assumptions and discuss some possibilities for applications and further research.

---

[1] This chapter is based on the paper [53].

3.1 A Central Limit Theorem for the Needlets Bispectrum

3.1.1 Bispectrum

The angular bispectrum (see [56, 84]) can be viewed as the harmonic transformation of the three-point angular correlation function, where the angular power spectrum is the Legendre transformation of the (two-point) angular correlation function. Write $\Omega_i = (\theta_i, \varphi_i)$, for $i = 1, 2, 3$; we have

$$ET(\Omega_1) T(\Omega_2) T(\Omega_3) = \sum_{l_1,l_2,l_3=1}^{\infty} \sum_{m_1 m_2 m_3} B_{l_1 l_2 l_3}^{m_1 m_2 m_3} Y_{l_1 m_1}(\Omega_1) Y_{l_2 m_2}(\Omega_2) Y_{l_3 m_3}(\Omega_3), \tag{3.1}$$

where the bispectrum $B_{l_1 l_2 l_3}^{m_1 m_2 m_3}$ is given by

$$B_{l_1 l_2 l_3}^{m_1 m_2 m_3} = E(a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3}). \tag{3.2}$$

Here, and in the sequel, the sums over m_i run from $-l_i$ to l_i, unless otherwise indicated. Both (3.1) and (3.2) are clearly equal to zero for zero-mean Guassian random fields. Moreover, the assumption that the CMB random field is statistacally isotropic entails that the right- and left-hand sides of (3.1) should be left unaltered by a rotation of the coordinate system. Therefore $B_{l_1 l_2 l_3}^{m_1 m_2 m_3}$ must take values ensuring that the three-point correlation function on the left-hand side of (3.1) remains unchanged if the three directions Ω_1, Ω_2 and Ω_3 are rotated by the same angle. Careful choices of the orientations entail that the angular bispectrum of an isotropic field can be non-zero only $l_i \leq l_j + l_k$ for all choices of $i, j, k = 1, 2, 3; l_1 + l_2 + l_3$ is even; and

$m_1 + m_2 + m_3 = 0$. More generally, as mentioned in the previous chapter, [45] shows that a necessary and sufficient condition for $B_{l_1 l_2 l_3}^{m_1 m_2 m_3}$ to represent the angular bispectrum of an isotropic random field is that there exist a real symmetric function of l_1, l_2, l_3, which we denote $b_{l_1 l_2 l_3}$, such that we have the identity

$$B_{l_1 l_2 l_3}^{m_1 m_2 m_3} = \mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3} b_{l_1 l_2 l_3}; \tag{3.3}$$

$b_{l_1 l_2 l_3}$ is labeled the reduced bispectrum, and $\mathcal{G}_{l_1 l_2 l_3}^{m_1 m_2 m_3}$ is the Gaunt integral (2.12). In view of (3.3), the dependence of the bispectrum ordinates on m_1, m_2, m_3 does not carry any physical information if the field is isotropic; hence it can be eliminated by focussing on the angular averaged bispectrum, defined by

$$\begin{aligned} B_{l_1 l_2 l_3} &:= \sum_{m_1 m_2 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} B_{l_1 l_2 l_3}^{m_1 m_2 m_3} \\ &= \left[\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \right]^{1/2} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} b_{l_1 l_2 l_3}, \end{aligned}$$

where we have used the orthogonality condition

$$\sum_{m_1, m_2, m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 = 1,$$

The minimum mean square error estimator of the bispectrum is provided by (Hu [45]),

$$\widehat{B}_{l_1 l_2 l_3} = \sum_{m_1 m_2 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \left(a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \right).$$

The statistic $\widehat{B}_{l_1 l_2 l_3}$ is called the (sample) angular averaged bispectrum; by any realization of the random field $T(x)$, it is a real-valued scalar, which does not depend on the choice of the coordinates, and it is invariant with respect to permutation of its arguments l_1, l_2, l_3.

Under Gaussianity, the bispectrum can be easily made model-independent, namely we can focus on the normalized bispectrum, which is defined by

$$I_{l_1 l_2 l_3} := \frac{\widehat{B}_{l_1 l_2 l_3}}{\sqrt{C_{l_1} C_{l_2} C_{l_3}}} = \sum_{m_1 m_2 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \frac{a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3}}{\sqrt{C_{l_1} C_{l_2} C_{l_3}}}. \tag{3.4}$$

Following [52, 88], an alternative definition of the (normalized) bispectrum can be considered, namely

$$\widetilde{I}_{l_1 l_2 l_3} = \int_{S^2} \frac{T_{l_1}(x) T_{l_2}(x) T_{l_3}(x)}{\sqrt{Var(T_{l_1}(x)) Var(T_{l_2}(x)) Var(T_{l_3}(x))}} dx \tag{3.5}$$

where

$$T_l(x) := L_l * T = \int_{S^2} T(y) \sum_m Y_{lm}(x) Y^*_{lm}(y) dy = \sum_m a_{lm} Y_{lm}(x) \,,$$

i.e. we focus on the Fourier projections of the random fields at multipoles (l_1, l_2, l_3). Indeed,

$$\begin{aligned} \widetilde{I}_{l_1 l_2 l_3} &= \int_{S^2} \sum_{m_1 m_2 m_3} \frac{a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \times Y_{l_1 m_1}(x) Y_{l_2 m_2}(x) Y_{l_3 m_3}(x)}{\sqrt{(2l_1+1)(2l_2+1)(2l_3+1)} \sqrt{C_{l_1} C_{l_2} C_{l_3}}} dx \\ &= \sum_{m_1 m_2 m_3} \frac{a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3}}{\sqrt{C_{l_1} C_{l_2} C_{l_3}}} \frac{\int_{S^2} Y_{l_1 m_1}(x) Y_{l_2 m_2}(x) Y_{l_3 m_3}(x) dx}{\sqrt{(2l_1+1)(2l_2+1)(2l_3+1)}} \\ &= \sum_{m_1 m_2 m_3} \frac{a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3}}{\sqrt{C_{l_1} C_{l_2} C_{l_3}}} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

So it is the same as (3.4) up to $\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}$. Both versions of the bispectrum have been shown to be extremely powerful against non-Gaussian alternatives, indeed there is now a widespread consensus that they make up the most efficient statistical procedures to search for non-Gaussianity, at least in the presence of fully observed spherical maps.

3.1.2 The Needlets Bispectrum

As mentioned in the introduction, the recent literature suggests that the most powerful statistic to search for non-Gaussianity in fully observed spherical random fields is the (normalized) angular bispectrum, defined as (3.5).

On the other hand, it is also well-known that the performance of Fourier methods in general, and the bispectrum in particular, deteriorates quite clearly in the presence of missing observations/partially observed maps ([14],[52]). A natural idea is then to explore the localization properties of the needlets in harmonic domain, together with their robustness in the presence of gaps, in order to devise a statistic which might mimic the positive features of the bispectrum, at the same time coping with the difficulties brought in by missing observations.

To this aim, we shall consider the (normalized) *needlets bispectrum*

$$I_{j_1j_2j_3} := \sum_{k_3} \widehat{\beta}_{j_1k_1}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_3k_3}\delta_{j_1j_2j_3}(k_1,k_2,k_3)h_{j_1j_2j_3}(k_1,k_2,k_3)\ , j_1 \leq j_2 \leq j_3\ , \tag{3.6}$$

where

$$\begin{aligned} \delta_{j_1j_2j_3}(k_1,k_2,k_3) &= \mathbb{I}(\xi_{j_3k_3} \in \mathcal{V}_{j_2k_2})\mathbb{I}(\xi_{j_2k_2} \in \mathcal{V}_{j_1k_1})\ , \\ h_{j_1j_2j_3}(k_1,k_2,k_3) &= \begin{cases} B^{j_2-j_1}\frac{\sqrt{\lambda_{j_1k_1}}}{\#\{k_2:k_2\in\mathcal{V}_{j_1k_1}\cap\mathcal{X}_{j_2}\}}, & j_1 < j_2 = j_3 \\ \sqrt{\lambda_{j_3k_3}}, & otherwise \end{cases}\ , \end{aligned}$$

$\mathbb{I}(.)$ denotes the indicator function and $\mathcal{V}_{jk}$ is the Voronoi Cell that corresponds to ξ_{jk}; note that for $j_2 = j_3$ we have $h_{j_1j_2j_3} \approx \sqrt{\lambda_{j_2k_2}}$. It is immediate to see that (3.6) can be seen as a natural development of (3.5), where the convolution with the orthonor-

mal projector operator L_l is replaced by the (discretized) convolution with the frame operator projection $\sqrt{\lambda_{jk}}\sum_l b(l/B^j)L_l$. Of course, in practice (3.5) is unfeasible and requires discretization to be implemented. In words, we are considering a version of the bispectrum where the exact identification of the multipoles is blurred by a form of suitable smoothing, with the purpose of a better robustness against missing observations.

The summation convention in (3.6) needs some further discussion. In practice, for applications the Voronoi tessellation is chosen in such a way to be nested across different scales (this is the case, for instance, for HEALPix ([39]), which is the standard package for CMB applications). Under such circumstances, our procedure can be described more explicitly as follows: we take

$$I_{j_1j_2j_3} := \sum_{k_3} \widehat{\beta}_{j_1k_1}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_3k_3}h_{j_1j_2j_3}(k_1,k_2,k_3), \tag{3.7}$$

where $k_2 = k_2(k_3)$ is the (unique) value of k_2 such that $\xi_{j_3k_3} \in \mathcal{V}_{j_2k_2}$, and $k_1 = k_1(k_3)$ is the (unique) value of k_1 such that $\xi_{j_2k_2}, \xi_{j_3k_3} \in \mathcal{V}_{j_1k_1}$. In other words, the "finest grid" $\mathcal{X}_{j_3}$ is the one which leads the summation, whereas smaller frequency terms are identified with those *centres* whose corresponding Voronoi cells include the points being summed. Note, however, that in the general non-nested case the centre of $\mathcal{V}_{j_1k_1}$ needs not belong to $\mathcal{X}_{j_3}$. In the sequel, for notational simplicity we write k_1, k_2 rather than $k_1(k_3), k_2(k_3)$, when no ambiguity is possible.

To investigate the asymptotic behaviour of the needlets bispectrum, we shall make an extensive use of the Diagram Formula which was introduced in the previous

chapter. A crucial point, of course is the determination of the frequencies where the needlets bispectrum is evaluated. We distinguish three cases, namely:

- **Case 1)** $j_1 + 1 < j_2 < j_3 - 1$;
- **Case 2)** $j_1 + 1 < j_2 = j_3$, or $j_1 = j_2 < j_3 - 1$;
- **Case 3)** $j_1 = j_2 = j_3$.

Case 1 corresponds to the situation where all three frequencies are different. Case 2) is basically the "squeezed" or collapsed configuration which is considered by [4], [56], and many other cosmological references; in words, one frequency is (much) smaller than the other two. It has been widely argued in the physical literature that this configuration corresponds to the highest power region for so-called local models of non-Gaussianity. Case 3 corresponds to so-called equilateral configurations; this case, however, can be largely investigated by means of results in ([7]) and we report it only for completeness, omitting many details in the proof. It should be noted that for case 1) and 2) we focus on frequencies that are at least two steps apart, in order to exploit the semiorthogonality properties of the needlets systems. Relaxing this assumption implies no new ideas and would only make the thesis notationally more complicated.

In each of the three cases we have trivially $\mathbb{E}I_{j_1 j_2 j_3} = 0$. We now focus on the asymptotic behaviour; here, asymptotic must be understood in the high resolution sense, i.e. we focus on a single realization of an isotropic random field, and we in-

vestigate the behaviour of our statistics at higher and higher bands. The first task is to ensure the statistics are non-degenerate and do not exhibit an explosive behaviour; this is the aim of the next Lemma. While the bound from above is quite straightforward, the lower bound is much more complicated and settles a question which was left open in ([6]), where the lower limit was simply assumed to be strictly larger than zero even for the simple case where $j_1 = j_2 = j_3$. As it is evident from the proof, the integer K depends on the choice of cubature points and of kernel function $b(.)$; more explicit expressions can be found below. Note also that unless the three bands are equal, condition b) cannot be satisfied for $B = 2$; indeed in CMB applications values of order $B \simeq 1.2, 1.3$ are likely to be favoured.

Lemma 3 *Under Condition A, as* $j_3 \to \infty$,

a) For all $j_1 \leq j_2 \leq j_3$, $\mathbb{E}I^2_{j_1 j_2 j_3} = O(1)$*; Also, there exists a positive integer* K *such that*

b) For $\{j_1 = j_2 = j_3\}$, *or* $\{j_1 + K < j_2,\ B^{j_1} + B^{j_2} \geq B^{j_3}\}$, $\left\{\mathbb{E}I^2_{j_1 j_2 j_3}\right\}^{-1} = O(1)$.*The bounds are uniform with respect to* j_1, j_2.

Proof. **(a)** In the sequel, we shall use the fact that the set of the cubature points of polynomial spaces with degree less than B^j are a $\kappa B^{-j}-$net; we also define $\rho := \max_{j,k}\{B^{-2j}\lambda_{j,k}\}$. The proof of all three cases are similar; we shall focus on case 2)

$j_1 + 1 < j_2 = j_3$. We have

$$\begin{aligned}
\mathbb{E}I^2_{j_1 j_2 j_2} &= \sum_{k_2, k_2' \in \mathcal{X}_{j_2}} \frac{\mathbb{E}\beta_{j_1 k_1}\beta_{j_1 k_1'}\mathbb{E}\left(\beta^2_{j_2 k_2}\beta^2_{j_2 k_2'}\right)}{\left(\sigma_{j_1}\sigma_{j_2}\sigma_{j_2}\right)^2} h_{j_1 j_2 j_3}(k_1, k_2, k_3) h_{j_1 j_2 j_3}(k_1', k_2', k_3') \\
&= B^{2j_2 - 2j_1} \sum_{k_1, k_1' \in \mathcal{X}_{j_1}} \sum_{\substack{k_2 \in \mathcal{V}_{j_1 k_1} \cap \mathcal{X}_{j_2}, \\ k_2' \in \mathcal{V}_{j_1 k_1'} \cap \mathcal{X}_{j_2}}} (\mathbb{E}\widehat{\beta}_{j_1 k_1}\widehat{\beta}_{j_1 k_1'} + 2\mathbb{E}\widehat{\beta}_{j_1 k_1}\widehat{\beta}_{j_1 k_1'}\left(\mathbb{E}\widehat{\beta}_{j_2 k_2}\widehat{\beta}_{j_2 k_2'}\right)^2) \\
&\quad \times \frac{\sqrt{\lambda_{j_1 k_1}}}{\#\{k_2, k_2 \in \mathcal{V}_{j_1 k_1} \cap \mathcal{X}_{j_2}\}} \frac{\sqrt{\lambda_{j_1 k_1'}}}{\#\{k_2', k_2' \in \mathcal{V}_{j_1 k_1'} \cap \mathcal{X}_{j_2}\}}
\end{aligned}$$

$$= B^{2j_2 - 2j_1} \sum_{k_2 k_2'} \frac{\mathbb{E}\beta_{j_1 k_1}\beta_{j_1 k_1'}\left(\mathbb{E}\beta_{j_2 k_2}\beta_{j_2 k_2'}\right)^2}{\left(\sigma_{j_1}\sigma_{j_2}\sigma_{j_2}\right)^2} \frac{\sqrt{\lambda_{j_1 k_1}}}{\#\{k_2, k_2 \in \mathcal{V}_{j_1 k_1} \cap \mathcal{X}_{j_2}\}} \frac{\sqrt{\lambda_{j_1 k_1'}}}{\#\{k_2', k_2' \in \mathcal{V}_{j_1 k_1'} \cap \mathcal{X}_{j_2}\}} .$$

Since $\lambda_{jk} \simeq B^{-2j}$ for every ξ_{jk}, and $\#\{k_2, k_2 \in \mathcal{V}_{j_1 k_1} \cap \mathcal{X}_{j_2}\} \simeq B^{2j_2 - 2j_1}$, the sum can be readily bounded by

$$\begin{aligned}
\frac{C}{B^{2j_2}} \sum_{k_2 k_2'} \mathbb{E}\widehat{\beta}_{j_1 k_1}\widehat{\beta}_{j_1 k_1'}\left(\mathbb{E}\widehat{\beta}_{j_2 k_2}\widehat{\beta}_{j_2 k_2'}\right)^2 &\leq \frac{C'}{B^{2j_2}} \sum_{k_2 k_2'} \left(\mathbb{E}\widehat{\beta}_{j_2 k_2}\widehat{\beta}_{j_2 k_2'}\right)^2 \\
&= O(1) \text{ as } j_2 \to \infty ,
\end{aligned}$$

in view of (2.25). This completes the proof of part a).

(b) The proof of the lower bound on the variance is considerably more delicate. We recall the correlation of the needlets coefficient is provided by

$$\mathbb{E}\widehat{\beta}_{jk}\widehat{\beta}_{jk'} \approx \frac{\sum_{l=B^{j-1}}^{B^{j+1}} b^2(\frac{l}{B^j}) C_l P_l(\cos\theta)}{\sum_{l=B^{j-1}}^{B^{j+1}} b^2(\frac{l}{B^j}) C_l \frac{2l+1}{4\pi}}$$

where $\theta = \arccos\langle \xi_{jk}, \xi_{jk'} \rangle$. The idea of our argument is to replace the needlets coefficients in the coarsest grid $\mathcal{X}_{j_1}$ by coefficients with the same resolution but evaluated over the pixels $\mathcal{X}_{j_2} \cap V_{j_1}(\xi_{j_1 k_1})$. This will allow us to circumvent the fact that the cubature weights at the smaller frequencies stay constant over a Voronoi cell,

whereas those corresponding to the highest j's may vary. We shall hence consider

$$\beta^*_{jk}(\xi_{j'k'}) = \sqrt{\lambda_{jk}} \sum_{l=B^{j-1}}^{B^{j+1}} \sum_{m=-l}^{l} b(\frac{l}{B^j}) a_{lm} Y_{lm}(\xi_{j'k'}) \ .$$

For fixed (j', k'), $j' < j$, $\beta^*_{jk}(\xi_{j'k'})$ varies over the pixels in $\mathcal{X}_j \cap V_{j'}(\xi_{j'k'})$. Let us now recall a few properties of Legendre polynomials, that we shall use extensively soon (see ([84]) for details); we have

$$\sup_{\theta\in[0,\pi]} P_l(\cos\theta) = P_l(\cos 0) = 1 \ , \text{ and } \sup_{\theta\in[0,\pi]} |\frac{d}{d\theta} P_l(\cos\theta)| \leq 3l \ .$$

See Appendix C for the proof for the second inequality (The inequality could also be obtained by Bernstein inequality. In fact, with different algorithm, the constant in the bound may even be improved to 2). As a consequence, for any positive $\epsilon < 1$, there exists a $\delta > 0$, s.t. if $0 \leq \theta < \delta \leq \epsilon/(3l)$, then,

$$|P_l(\cos(\theta_0 + \theta)) - P_l(\cos\theta_0)| \leq 3l\theta \leq \epsilon,$$

because

$$0 \leq \cos(\theta_0 + \theta) - \cos\theta_0 = 2\sin\frac{2\theta_0 + \theta}{2} \sin\frac{\theta}{2} \leq \theta \ .$$

Now fix an integer K such that $K \geq \log_B 6\kappa/\epsilon$; if we let $\{\xi_{jk}\}$ be the cubature points of polynomial space with degree less than B^{j+K+1}, (note that we can assume all of these sets are $\kappa B^{-j}-$nets), then for any $j_2 > j_1, \xi_{j_2k_2} \in \mathcal{V}_{j_1k_1}, \xi_{j_2k'_2} \in \mathcal{V}_{j_1k'_1}$,

$$\begin{aligned} \left|\left\langle \xi_{j_1k_1}, \xi_{j_1k'_1} \right\rangle - \left\langle \xi_{j_2k_2}, \xi_{j_2k'_2} \right\rangle\right| &= \left|d(\xi_{j_1k_1}, \xi_{j_1k'_1}) - d\left(\xi_{j_2k_2}, \xi_{j_2k'_2}\right)\right| \\ &\leq 2\kappa B^{-(j_1+K+1)} \leq B^{-(j_1+1)}\epsilon/3 \ . \end{aligned}$$

It follows that

$$\begin{aligned}
&\left|\mathbb{E}\beta_{j_1k_1}\beta_{j_1k_1'} - \mathbb{E}[\beta^*_{j_1k_1}(\xi_{j_2k_2})\beta^*_{j_1k_1'}(\xi_{j_2k_2'})\right| \\
\leq\ & \sqrt{\lambda_{jk}\lambda_{jk'}} \sum_{l=B^{j_1-1}}^{B^{j_1+1}} b^2(\frac{l}{B^j})C_l \left| \sum_{m=-l}^{l} Y_{lm}(\xi_{j_1k_1})\overline{Y_{lm}(\xi_{j_1k_1'})} - \sum_{m=-l}^{l} Y_{lm}(\xi_{j_2k_2})\overline{Y_{lm}(\xi_{j_2k_2'})}\right| \\
=\ & \sqrt{\lambda_{jk}\lambda_{jk'}} \sum_{l=B^{j_1-1}}^{B^{j_1+1}} b^2(\frac{l}{B^j})C_l \frac{2l+1}{4\pi} \left| P_l(\left\langle \xi_{j_1k_1}, \xi_{j_1k_1'} \right\rangle) - P_l(\left\langle \xi_{j_2k_2}, \xi_{j_2k_2'} \right\rangle)\right| \\
\leq\ & \epsilon\sqrt{\lambda_{jk}\lambda_{jk'}} \sum_{l=B^{j_1-1}}^{B^{j_1+1}} b^2(\frac{l}{B^j})C_l \frac{2l+1}{4\pi} \leq \epsilon C\sigma^2_{j_1}\ , \qquad (3.8)
\end{aligned}$$

where C is a constant, $C = C\,(\kappa)$ (see [65]). We are now in the position to establish our lower bound. By simple algebraic manipulations, we have

$$\begin{aligned}
\mathbb{E}I^2_{j_1j_2j_3} \approx\ & \sum_{k_3,k_3'\in\mathcal{X}_{j_3}} \frac{\mathbb{E}\beta^*_{j_1k_1}(\xi_{j_3k_3})\beta^*_{j_1k_1'}(\xi_{j_3k_3'})\mathbb{E}\beta^*_{j_2k_2}(\xi_{j_3k_3})\beta^*_{j_2k_2'}(\xi_{j_3k_3'})}{(\sigma_{j_1}\sigma_{j_2}\sigma_{j_3})^2} \\
& \times\mathbb{E}\beta_{j_3k_3}\beta_{j_3k_3'}\sqrt{\lambda_{j_3k_3}}\sqrt{\lambda_{j_3k_3'}} \\
& + \sum_{k_3,k_3'\in\mathcal{X}_{j_3}} \frac{\left(\mathbb{E}\beta_{j_1k_1}\beta_{j_1k_1'} - \mathbb{E}\beta^*_{j_1k_1}(\xi_{j_3k_3})\beta^*_{j_1k_1'}(\xi_{j_3k_3'})\right)}{\sigma^2_{j_1}} \\
& \times\mathbb{E}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_2k_2'}\mathbb{E}\widehat{\beta}_{j_3k_3}\widehat{\beta}_{j_3k_3'}\sqrt{\lambda_{j_3k_3}}\sqrt{\lambda_{j_3k_3'}} \\
& + \sum_{k_3,k_3'\in\mathcal{X}_{j_3}} \frac{\mathbb{E}\beta^*_{j_1k_1}(\xi_{j_3k_3})\beta^*_{j_1k_1'}(\xi_{j_3k_3'})\left(\mathbb{E}\beta_{j_2k_2}\beta_{j_2k_2'} - \mathbb{E}\beta^*_{j_2k_2}(\xi_{j_3k_3})\beta^*_{j_2k_2'}(\xi_{j_3k_3'})\right)}{(\sigma_{j_1}\sigma_{j_2})^2} \\
& \mathbb{E}\widehat{\beta}_{j_3k_3}\widehat{\beta}_{j_3k_3'}\sqrt{\lambda_{j_3k_3}}\sqrt{\lambda_{j_3k_3'}}.
\end{aligned}$$

In view of (3.8), (2.25) and standard manipulations, the second and third components can be made arbitrarily small. To bound the first component, we recall the Gaunt

integral (2.12),

$$\begin{aligned} & \int_{S^2} Y_{l_1m_1}(x)Y_{l_2m_2}(x)Y_{l_3m_3}(x)dx \\ = & \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

where for the Wigner's 3j coefficients introduced on the right-hand side we recall the properties (see Appendix A)

$$\sum_{m_1m_2m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 \equiv 1, \tag{3.9}$$

and

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \geq \frac{C}{l_3^2} \mathbb{I}\left\{l_1+l_2 \geq l_3, l_1+l_2+l_3 = even\right\} . \tag{3.10}$$

Using the fact that $\{\lambda_{jk}\}$ are the cubature weights corresponding to the space $\mathcal{K}_{3B^{j+1}}$, easy manipulations yield

$$\begin{aligned} & \frac{1}{\sigma_{j_1}^2\sigma_{j_2}^2\sigma_{j_3}^2} B^{2j_3} \sum_{l_1l_2l_3} \sum_{m_1m_2m_3} b^2(\frac{l_1}{B^{j_1}}) b^2(\frac{l_2}{B^{j_2}}) b^2(\frac{l_3}{B^{j_3}}) C_{l_1}C_{l_2}C_{l_3} \\ & \times \sum_{k_3} Y_{l_1m_1}(\xi_{j_3k_3}) Y_{l_2m_2}(\xi_{j_3k_3}) Y_{l_3m_3}(\xi_{j_3k_3}) \lambda_{j_3k_3} \\ & \times \sum_{k_3'} Y_{l_1m_1}(\xi_{j_3k_3'}) Y_{l_2m_2}(\xi_{j_3k_3'}) Y_{l_3m_3}(\xi_{j_3k_3'}) \lambda_{j_3k_3'} \\ = & \frac{1}{\sigma_{j_1}^2\sigma_{j_2}^2\sigma_{j_3}^2} B^{2j_3} \sum_{l_1l_2l_3} \sum_{m_1m_2m_3} b^2(\frac{l_1}{B^{j_1}}) b^2(\frac{l_2}{B^{j_2}}) b^2(\frac{l_3}{B^{j_3}}) C_{l_1}C_{l_2}C_{l_3} \\ & \times (\int_{S^2} Y_{l_1m_1}(x)Y_{l_2m_2}(x)Y_{l_3m_3}(x)dx)^2 \\ = & \frac{1}{\sigma_{j_1}^2\sigma_{j_2}^2\sigma_{j_3}^2} B^{2j_3} \sum_{l_1l_2l_3} b^2(\frac{l_1}{B^{j_1}}) b^2(\frac{l_2}{B^{j_2}}) b^2(\frac{l_3}{B^{j_3}}) C_{l_1}C_{l_2}C_{l_3} \\ & \times \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \sum_{m_1m_2m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \end{aligned}$$

and using $\sigma_j^2 \approx B^{2j} C_{B^j}$, $B^{j_i} \approx l_i, i = 1, 2, 3$ we get

$$\begin{aligned} & \frac{1}{\sigma_{j_1}^2 \sigma_{j_2}^2 \sigma_{j_3}^2} B^{2j_3} \sum_{l_1 l_2 l_3} b^2(\frac{l_1}{B^{j_1}}) b^2(\frac{l_2}{B^{j_2}}) b^2(\frac{l_3}{B^{j_3}}) C_{l_1} C_{l_2} C_{l_3} \\ & \times \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \\ \approx \quad & B^{-j_1-j_2-j_3} \sum_{\substack{l_1+l_2 \geq l_3 \\ l_1+l_2+l_3 = even}} b^2(\frac{l_1}{B^{j_1}}) b^2(\frac{l_2}{B^{j_2}}) b^2(\frac{l_3}{B^{j_3}}) > c > 0 \, . \end{aligned}$$

In the previous argument, we have taken $c := \inf_{j_1 j_2 j_3} c_{j_1 j_2 j_3}$, where

$$c_{j_1 j_2 j_3} := B^{-j_1-j_2-j_3} \sum_{\substack{l_1+l_2 \geq l_3 \\ l_1+l_2+l_3 = even}} b^2(\frac{l_1}{B^{j_1}}) b^2(\frac{l_2}{B^{j_2}}) b^2(\frac{l_3}{B^{j_3}}) \, .$$

It is simple to show that $c > 0$. Indeed, under (2.19), we have

$$\begin{aligned} c_{j_1 j_2 j_3} \quad & \geq \quad \frac{1}{2} \int_{\substack{\{u_1+u_2 \geq u_3\}, \\ [0,2]^3 \cap \{u_1 \leq u_2 \leq u_3\}}} b^2(u_1) \, b^2(u_2) \, b^2(u_3) \, du_1 du_2 du_3 \\ & \geq \quad \frac{1}{2} \int_{1-\delta_0}^{1+\delta_0} du_1 \int_{u_1}^{1+\delta_0} du_2 \int_{1+\delta_0}^{2-2\delta_0} du_3 \left\{ b^2(u_1) \, b^2(u_2) \, b^2(u_3) \right\} \\ & \geq \quad \frac{c_0}{8} \delta_0^3 . \end{aligned}$$

where $\delta_0 \leq 1/4$, and $b(x) \geq 1/2$ for any $x \in [1-\delta_0, 1+\delta_0]$, $c_0 = \inf_{x \in [1+\delta_0, 2-2\delta_0]} b^2(x)$.

The same argument as before could be used to establish lower bounds when $j_1 = j_2 < j_3$ or $j_1 = j_2 = j_3$. To conclude our proof, we consider the case when

$j_1 < j_2 = j_3$. We obtain

$$
\begin{aligned}
& B^{2j_2-2j_1} \sum_{k_2 k_2'} \mathbb{E}\widehat{\beta}_{j_1k_1}\widehat{\beta}_{j_1k_1'} \left(\mathbb{E}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_2k_2'}\right)^2 \\
& \times \frac{\sqrt{\lambda_{j_1k_1}}}{\#\{k_2, k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}\}} \frac{\sqrt{\lambda_{j_1k_1'}}}{\#\{k_2', k_2' \in \mathcal{V}_{j_1k_1'} \cap \mathcal{X}_{j_2}\}} \\
\geq \; & C(\kappa) \frac{B^{2j_2-2j_1}}{B^{4j_2-4j_1}} \sum_{k_1} \sum_{k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}} \left(\mathbb{E}\widehat{\beta}^2_{j_2k_2}\right)^2 \lambda_{j_1k_1} \\
& + C'(\kappa, \rho) \frac{B^{2j_2-2j_1}}{B^{4j_2-4j_1}} \sum_{k_1 \neq k_1'} \sum_{\substack{k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}, \\ k_2' \in \mathcal{V}_{j_1k_1'} \cap \mathcal{X}_{j_2}}} \left(\mathbb{E}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_2k_2'}\right)^2,
\end{aligned}
$$

the two summands corresponding to the cases where k_2, k_2' belong to the same or to different Voronoi cells, respectively. The first part is equal to

$$
C(\kappa) B^{2j_1-2j_2} \sum_{k_1} \sum_{k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}} \lambda_{j_1k_1} \geq C(\kappa) \sum_{k_1} \lambda_{j_1k_1} = 4\pi C(\kappa),
$$

while the second part is equal to

$$
\begin{aligned}
& C'(\kappa, \rho) B^{-2j_2} \Big\{ \sum_{\substack{k_1 \neq k_1'; \\ d\left(\mathcal{V}_{j_1k_1}, \mathcal{V}_{j_1k_1'}\right)=0}} \sum_{\substack{k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}, \\ k_2' \in \mathcal{V}_{j_1k_1'} \cap \mathcal{X}_{j_2}}} + \sum_{\substack{k_1 \neq k_1'; \\ d\left(\mathcal{V}_{j_1k_1}, \mathcal{V}_{j_1k_1'}\right)>0}} \sum_{\substack{k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}, \\ k_2' \in \mathcal{V}_{j_1k_1'} \cap \mathcal{X}_{j_2}}} \Big\} \left(\mathbb{E}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_2k_2'}\right)^2 \\
\leq \; & C'(\kappa, \rho) \frac{6\pi}{B^{2j_2}} \sum_{k_1} \sum_{\substack{k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}, \\ k_2' \in \mathcal{V}_{j_1k_1'} \cap \mathcal{X}_{j_2}}} \left(\mathbb{E}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_2k_2'}\right)^2 \\
& + \frac{C'(\kappa, \rho)}{B^{2j_2}} \sum_{k_1} \sum_{k_2 \in \mathcal{V}_{j_1k_1} \cap \mathcal{X}_{j_2}} \Big\{ \sum_{k_1'} \sum_{\substack{k_2' \in \mathcal{V}_{j_1k_1'} \cap \mathcal{X}_{j_2}; \\ d\left(\mathcal{V}_{j_1k_1}, \mathcal{V}_{j_1k_1'}\right)>0}} \left(\mathbb{E}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_2k_2'}\right)^2 \Big\} \\
\leq \; & C'(\kappa, \rho) \frac{6\pi}{B^{2j_2}} \sum_{k_1} B^{j_2-j_1} + \frac{C'(\kappa, \rho)}{B^{2j_2}} \sum_{k_2} \sum_{k_2'} \frac{C_M}{\left(1 + B^{j_2}(B^{-j_1} + d(k_2', \mathcal{V}_{j_1k_1(k_2)}))\right)^M} \\
\leq \; & C'(\kappa, \rho) B^{-(j_2-j_1)} + C'(\kappa, \rho, C_M) \int_{B^{-j_1}}^{\infty} \frac{B^{4j_2} \sin\theta}{(1 + B^{j_2}\theta)^M} d\theta = C(\kappa, \rho, C_M) B^{-(j_2-j_1)})
\end{aligned}
$$

By taking $j_2 - j_1 \geq K \geq \max\left\{[\log_B\{2\kappa/\epsilon\}], \frac{1}{2}c_1/c_2\right\}$, where $c_1 = C(\kappa), c_2 = C(\kappa, \rho, C_M)$, and $\epsilon \leq \frac{1}{2}c/C_M$, the argument is completed. The proof for case (3) is similar (actually slightly easier). ■

The following weak convergence theorem is the main result of this Section. We stress that the statement could be easily extended to a multivariate Central Limit Theorem; however, because this extension would not entail any substantial novelty, at the same time making the notation much more cumbersome, we prefer to stick to the univariate case.

Theorem 4 *Let $T(x)$ be a zero-mean, mean square continuous and isotropic Gaussian random field, with angular power spectrum that satisfies Condition A. As $j_1 \to \infty$, we have*

$$\frac{I_{j_1 j_2 j_3}}{\sqrt{EI^2_{j_1 j_2 j_3}}} \to_d N(0,1)\,.$$

Proof. In view of the results in [69, 72], to establish a Central Limit Theorem for a multilinear form in Gaussian random variables, it is enough to investigate the asymptotic behaviour of fourth order moments (or equivalently cumulants), see also [22]. Our aim is then to show that, as $j_1 \to \infty$,

$$\mathbb{E}I^4_{j_1 j_2 j_3} = 3(\mathbb{E}I^2_{j_1 j_2 j_3})^2 + O(B^{-j_1/2})\,.$$

For notational simplicity, we shall write

$$\rho_j(k', k) := \mathbb{E}\widehat{\beta}_{j,k'}\widehat{\beta}_{j,k}.$$

By the diagram formula we have, for all index sets I:

$$\mathbb{E}\{\prod_{i\in I}\prod_{l=1}^{3}\widehat{\beta}_{j_l k_l^i}\} = \sum_{\gamma\in\Gamma(I,3)}\prod_{\{(i,l)(i',l')\}\in\gamma}\delta_{j_l}^{j_{l'}}\rho_{j_l}(k_l^{(i)},k_{l'}^{(i')})\,.$$

Similarly to [56], we define

$$\begin{aligned}\rho(\gamma;j_1,j_2,j_3) &= \sum_{k_3^i\in\mathcal{X}_{j_3}}\prod_{\{(i,l)(i',l')\}\in\gamma}\delta_{j_l}^{j_{l'}}\rho_{j_l}(k_l^{(i)},k_{l'}^{(i')})\\ &\quad\times\prod_{i\in I}\delta_{j_1j_2j_3}(k_1^{(i)},k_2^{(i)},k_3^{(i)})h_{j_1j_2j_3}(k_1^{(i)},k_2^{(i)},k_3^{(i)})\end{aligned}$$

so that

$$\begin{aligned}\mathbb{E}I^4_{j_1j_2j_3} &= \mathbb{E}\{\sum_{k_3}\widehat{\beta}_{j_1k_1}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_3k_3}\delta_{j_1j_2j_3}(k_1,k_2,k_3)h_{j_1j_2j_3}(k_1,k_2,k_3)\}^4\\ &= \sum_{k_3^{(1)}}\cdots\sum_{k_3^{(4)}}\mathbb{E}\{\prod_{i=1}^{4}\prod_{l=1}^{3}\widehat{\beta}_{j_lk_l^i}\delta_{j_1j_2j_3}(k_1^{(i)},k_2^{(i)},k_3^{(i)})h_{j_1j_2j_3}(k_1^{(i)},k_2^{(i)},k_3^{(i)})\}\\ &= \{\sum_{\gamma\in\Gamma_C(4,3)}+\sum_{\gamma\in\Gamma_{\overline{C}}(4,3)}\}\rho(\gamma;j_1,j_2,j_3)\end{aligned}$$

where Γ_C is the set of all connected diagrams. To conclude our argument, we only need to show that

$$\sum_{\gamma\in\Gamma_C(4,3)}\rho(\gamma;j_1,j_2,j_3) = O(B^{-j_1/2})\text{ , as }j_1\to\infty\,, \tag{3.11}$$

and

$$\sum_{\gamma\in\Gamma_{\overline{C}}(4,3)}\rho(\gamma;j_1,j_2,j_3) = \sum_{\gamma\in\Gamma_P(4,3)}\rho(\gamma;j_1,j_2,j_3) = 3(\mathbb{E}I^2_{j_1j_2j_3})^2. \tag{3.12}$$

(3.12) is an immediate consequence of the definition of $\mathbb{E}I^2_{j_1j_2j_3}$ and trivial combinatorial manipulations(see [80]). The result in (3.11) is proved by splitting connected diagrams into those with or without flat edges. Diagrams with flat edges are dealt with in Lemma 5, while those without are dealt with in Lemma 6. We stress that, on

the contrary of what is often the case when the diagram formula is applied, terms with flat edges do not vanish, due to correlation among different locations in the spherical needlets coefficients. This completes the proof of the main result. ■

Lemma 5 *For a connected diagram with flat edges, $\gamma \in \Gamma_{CF}(4,3)$, we have*

$$\rho(\gamma; j_1, j_2, j_3) = O(B^{-2j_1}) \text{ , as } j_1 \to \infty \, .$$

Proof. We write $\{(k_b^{(a)}, j_b)\}_{a=1,\dots,4, b=1,2,3}$ for the elements in our diagram, a and b denoting the row and column indexes, respectively. We recall also that $k_2 = k_2(k_3)$, $k_1 = k_1(k_3)$, as explained earlier. For Case 1), i.e. $j_1 < j_2 - 1 < j_3 - 2$, since $E\{\beta_{j_3 k_3} \beta_{j_2 k_2}\} = 0$ for every $k_2 \in \mathcal{X}_{j_2}, k_3 \in \mathcal{X}_{j_3}$,it is easy to see that $\rho(\gamma; j_1, j_2, j_3) \equiv 0$. For Case 2), i.e. $j_1 + 1 < j_2 = j_3$, we assume (without loss of generality) that a flat edge is present in the first row of the diagram, i.e. we let $\{(k_2^{(1)}, j_2)(k_3^{(1)}, j_3)\} \in \gamma$. By the same argument as in (2.22) we obtain immediately

$$B^{4j_2 - 4j_1} \prod_{i \in \{2,..,4\}} \sum_{k_2^{(i)}, i \neq 1} \prod_{\substack{\{(i,l)(i',l')\} \in \gamma, \\ i, i' \neq 1}} \delta_{j_l}^{j_{l'}} \rho_{j_l}(k_l^{(i)}, k_{l'}^{(i')}) \sum_{k_2^{(1)}} \frac{\rho_{j_1}(k_1^{(1)}, k_1^{(2)}) \sqrt{\lambda_{j_1 k_1^{(1)}}}}{\#\{k_2^{(1)}, k_2^{(1)} \in \mathcal{V}_{j_1 k_1^{(1)}} \cap \mathcal{X}_{j_2}\}} = 0 \, .$$

On the other hand, under $j_1 = j_2 < j_3 - 1$, again we assume a flat edge $\{(k_1^{(1)}, j_1)(k_1^{(1)}, j_2)\} \in \gamma$; by necessity, there should exist another flat edge in the graph, and w.l.o.g. we take it be in the fourth row, i.e. $\{(k_1^{(4)}, j_1)(k_2^{(4)}, j_2)\} \in \gamma$. Then we have $\xi_{j_1 k_1^{(i)}} = \xi_{j_2 k_2^{(i)}}$, and the resulting term can be bounded by

$$
\begin{aligned}
& \frac{1}{B^{4j_3}} \prod_{i\in\{1,..,4\}} \sum_{k_3^{(i)}} \left|\rho_{j_3}(k_3^{(1)},k_3^{(2)})\rho_{j_3}(k_3^{(3)},k_3^{(4)})\right| \rho_{j_1}^2(k_1^{(2)},k_1^{(3)}) \\
= & \frac{1}{B^{4j_3}} \sum_{k_3^{(2)},k_3^{(3)}} \rho_{j_1}^2(k_1^{(2)},k_1^{(3)}) \{\sum_{k_3^{(1)}} \left|\rho_{j_3}(k_3^{(1)},k_3^{(2)})\right| \sum_{k_3^{(4)}} \left|\rho_{j_3}(k_3^{(3)},k_3^{(4)})\right|\} \\
\le & \frac{1}{B^{4j_3}} \sum_{k_3^{(2)},k_3^{(3)}} \rho_{j_1}^2(k_1^{(2)},k_1^{(3)}) C_M \times C_M \\
\le & \frac{C}{B^{4j_3}} \sum_{k_1^{(2)},k_1^{(3)}} \rho_{j_2}^2(k_1^{(2)},k_1^{(3)}) \left[\max_{k_1^{(2)}} \#\left\{\xi_{j_3k_3}\in\mathcal{V}_{j_1}(k_1^{(2)})\right\}\right] \left[\max_{k_1^{(3)}} \#\left\{\xi_{j_3k_3}\in\mathcal{V}_{j_1}(k_1^{(3)})\right\}\right] \\
= & \frac{CB^{2j_1}}{B^{4j_3}} B^{4j_3-4j_1} = O(B^{-2j_1}) \,.
\end{aligned}
$$

Finally for Case 3), the argument is analogous; more precisely, components with diagrams in $\Gamma_{CF}(4,3)$ can be bounded by

$$
\begin{aligned}
|\rho(\gamma;j,j,j)| &= \frac{1}{B^{4j}} \prod_{i\in I} \sum_{k^{(i)}} \left|\rho_j(k^{(1)},k^{(2)})\rho_j(k^{(3)},k^{(4)})\right| \rho_j^2(k^{(2)},k^{(3)}) \\
&= \frac{1}{B^{4j}} \sum_{k^{(2)},k^{(3)}} \rho_j^2(k^{(2)},k^{(3)}) \{\sum_{k^{(1)}} \left|\rho_j(k^{(1)},k^{(2)})\right| \sum_{k^{(4)}} \left|\rho_j(k^{(3)},k^{(4)})\right|\} \\
&\le \frac{1}{B^{4j}} C_M \sum_{k^{i_2},k^{i_3}} \rho_j^2(k^{(2)},k^{(3)}) = O(B^{-2j}) \,,
\end{aligned}
$$

where in the last equality we used (2.25). Thus the proof is completed. ■

Lemma 6 *For a connected diagram without a flat edge, $\gamma \in \Gamma_{C\overline{F}}(4,3)$, we have*

$$
\rho(\gamma;j_1,j_2,j_3) = O(B^{-j_1/2}) \text{ , as } j_1 \to \infty \,.
$$

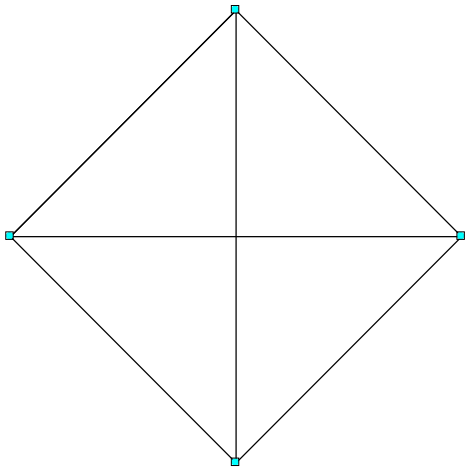

Figure 3.1:

Proof. Connected diagrams without flat edges and with four nodes can be partitioned into two classes, i.e. so-called *cliques,* where each vertex is connected to all three others (see fig3.1), and terms with loops of order 2 (see fig3.2).We focus on the first class; without loss of generality, we can express $\rho(\gamma; j_1, j_2, j_3)$ by

$$\prod_{i=1}^{4} \sum_{k_3^i} \rho_{j_1}(k_1^{(1)}, k_1^{(2)}) \rho_{j_1}(k_1^{(3)}, k_1^{(4)}) \rho_{j_2}(k_2^{(1)}, k_2^{(3)}) \rho_{j_2}(k_2^{(2)}, k_2^{(4)}) \rho_{j_3}(k_3^{(1)}, k_3^{(4)}) \rho_{j_3}(k_3^{(2)}, k_3^{(3)})$$
$$\times \prod_{i=1}^{4} \delta_{j_1 j_2 j_3}(k_1^{(i)}, k_2^{(i)}, k_3^{(i)}) h_{j_1 j_2 j_3}(k_1^{(i)}, k_2^{(i)}, k_3^{(i)}) \, .$$

By means of (2.24) we readily obtain the bound

$$\frac{C}{B^{4j_3}} \prod_{i=1}^{4} \sum_{k_3^i} \frac{1}{(1 + B^{j_1} d(k_1^{(1)}, k_1^{(2)}))^M} \frac{1}{(1 + B^{j_1} d(k_1^{(3)}, k_1^{(4)}))^M} \frac{1}{(1 + B^{j_2} d(k_2^{(1)}, k_2^{(3)}))^M}$$
$$\times \quad \frac{1}{(1 + B^{j_2} d(k_2^{(2)}, k_2^{(4)}))^M} \frac{1}{(1 + B^{j_3} d(k_3^{(1)}, k_3^{(4)}))^M} \frac{1}{(1 + B^{j_3} d(k_3^{(2)}, k_3^{(3)}))^M} \, . \qquad (3.13)$$

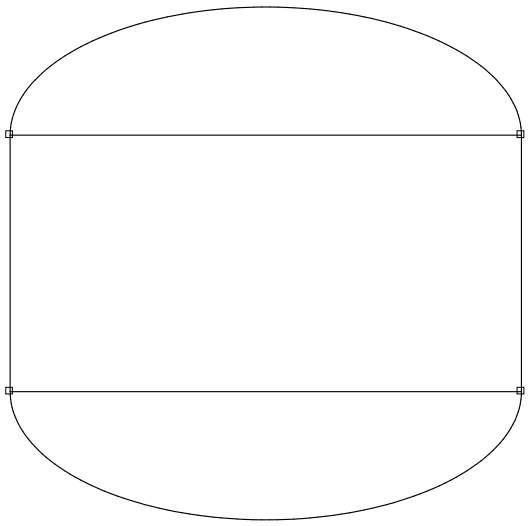

Figure 3.2:

Consider first Case 2), with $j_1 + 1 < j_2 = j_3$. Using inequality (2.26), (3.13) can be replaced by

$$
\begin{aligned}
& B^{4j_2-4j_1} \sum_{k_2^{(1)},..,k_2^{(I)} \in \mathcal{X}_{j_2}} \frac{1}{(1+B^{j_2} d(k_2^{(1)}, k_2^{(3)}))^M} \frac{1}{(1+B^{j_2} d(k_2^{(2)}, k_2^{(4)}))^M} \\
& \times \frac{1}{(1+B^{j_2} d(k_2^{(1)}, k_2^{(4)}))^M} \frac{1}{(1+B^{j_2} d(k_2^{(2)}, k_2^{(3)}))^M} \prod_{i=1}^{4} \frac{\sqrt{\lambda_{j_2 k_2^{(i)}}}}{\#\{k_2^{(i)}, k_2^{(i)} \in \mathcal{V}_{j_1 k_1^{(i)}} \cap \mathcal{X}_{j_2}\}} \\
\leq \; & B^{4j_2} \frac{C}{B^{8j_2}} \sum_{k_2^{(3)}, k_2^{(4)}} \{ \sum_{k_2^{(1)}} \frac{1}{(1+B^{j_2} d(k_2^{(1)}, k_2^{(3)}))^M} \frac{1}{(1+B^{j_2} d(k_2^{(1)}, k_2^{(4)}))^M} \} \\
& \times \{ \sum_{k_2^{(2)}} \frac{1}{(1+B^{j_2} d(k_2^{(2)}, k_2^{(4)}))^M} \frac{1}{(1+B^{j_2} d(k_2^{(2)}, k_2^{(3)}))^M} \} \\
\leq \; & \frac{C}{B^{4j_2}} \sum_{k_2^{(3)}, k_2^{(4)}} \{ \frac{C_M}{(1+B^{j_2} d(k_2^{(3)}, k_2^{(4)}))^{2M}} \} \leq \frac{C}{B^{4j_2}} C_M C_{2M} B^{2j_2} = O(B^{-2j_2}) = o(B^{-2j_1}) \, .
\end{aligned}
$$

Likewise, for $j_1 = j_2 < j_3 - 1$, we have the bound

$$
\begin{aligned}
& \frac{C}{B^{4j_3}} \prod_{i=1}^{4} \sum_{k_3^{(i)}} \frac{1}{(1+B^{j_1} d(k_1^{(1)}, k_1^{(2)}))^M} \frac{1}{(1+B^{j_1} d(k_1^{(3)}, k_1^{(4)}))^M} \frac{1}{(1+B^{j_2} d(k_1^{(1)}, k_1^{(3)}))^M} \\
& \times \frac{1}{(1+B^{j_2} d(k_1^{(2)}, k_1^{(4)}))^M} \frac{1}{(1+B^{j_3} d(k_3^{(1)}, k_3^{(4)}))^M} \frac{1}{(1+B^{j_3} d(k_3^{(2)}, k_3^{(3)}))^M} \\
\leq \; & \frac{C}{B^{4j_3}} \prod_{i=1}^{4} \sum_{k_1^{(i)}} \frac{1}{(1+B^{j_1} d(k_1^{(1)}, k_1^{(2)}))^M} \frac{1}{(1+B^{j_1} d(k_1^{(3)}, k_1^{(4)}))^M} \\
& \times \frac{1}{(1+B^{j_2} d(k_1^{(1)}, k_1^{(3)}))^M} \frac{1}{(1+B^{j_2} d(k_1^{(2)}, k_1^{(4)}))^M} \\
& \times \sum_{k_3^1 \in \mathcal{V}_{j_1}(k_1^{(1)}) \cap \mathcal{X}_{j_3}} \frac{1}{(1+B^{j_3} d(k_3^{(1)}, k_3^{(4)}))^M} \sum_{k_3^2 \in \mathcal{V}_{j_1}(k_1^{(2)}) \cap \mathcal{X}_{j_3}} \frac{1}{(1+B^{j_3} d(k_3^{(2)}, k_3^{(3)}))^M} \\
\leq \; & \frac{C B^{4j_3 - 4j_1}}{B^{4j_3}} \sum_{k_1^{(2)}, k_1^{(3)}} \{ \sum_{k_1^{(1)}} \frac{1}{(1+B^{j_1} d(k_1^{(1)}, k_1^{(2)}))^M} \frac{1}{(1+B^{j_2} d(k_1^{(1)}, k_1^{(3)}))^M} \} \\
& \times \{ \sum_{k_1^{(4)}} \frac{1}{(1+B^{j_1} d(k_1^{(3)}, k_1^{(4)}))^M} \frac{1}{(1+B^{j_2} d(k_1^{(2)}, k_1^{(4)}))^M} \} \\
\leq \; & \frac{C B^{4j_3 - 4j_1}}{B^{4j_3}} \sum_{k_1^{(2)}, k_1^{(3)}} \frac{C_M^2}{(1+B^{j_2} d(k_1^{(2)}, k_1^{(3)}))^{2M}} \leq C B^{2j_1 - 4j_1} = O(B^{-2j_1}) \, .
\end{aligned}
$$

This concludes the proof for Case 2); the proof for Case 3) could be implemented along identical lines.

The analysis for case 1) $j_1 + 1 < j_2 < j_3 - 1$, i.e. the case where all three frequencies differ, requires considerably more care. As before, let $\mathcal{V}_{j_u}(k_u^{(i)}, r)$ be the Voronoi cells associated to $k_u^{(i)}$, and we recall it satisfies $B(x_i, r/2) \subset \mathcal{V}_{j_u}(x_i) \subset$

$B(x_i, r), k_u^{(i)} \in \mathcal{X}_{j_u}$. Our idea is to partition (3.13) into four elements, as follows:

$$
\begin{aligned}
& (3.13) \\
\leq\ & \frac{C}{B^{4j_3}} \prod_{i=1}^{4} \sum_{k_3^i} \{\mathbb{I}(d(k_1^{(1)}, k_1^{(4)} > 3r), d(k_1^{(2)}, k_1^{(3)} > 3r)) + \mathbb{I}(d(k_1^{(1)}, k_1^{(4)} > 3r), d(k_1^{(2)}, k_1^{(3)} \leq 3r)) \\
+\ & \mathbb{I}(d(k_1^{(1)}, k_1^{(4)} \leq 3r), d(k_1^{(2)}, k_1^{(3)} > 3r)) + \mathbb{I}(d(k_1^{(1)}, k_1^{(4)} \leq 3r), d(k_1^{(2)}, k_1^{(3)} \leq 3r))\} \\
\times\ & \frac{1}{(1 + B^{j_1} d(k_1^{(1)}, k_1^{(2)}))^M} \frac{1}{(1 + B^{j_1} d(k_1^{(3)}, k_1^{(4)}))^M} \frac{1}{(1 + B^{j_2} d(k_2^{(1)}, k_2^{(3)}))^M} \\
\times\ & \frac{1}{(1 + B^{j_2} d(k_2^{(2)}, k_2^{(4)}))^M} \frac{1}{(1 + B^{j_3} d(k_3^{(1)}, k_3^{(4)}))^M} \frac{1}{(1 + B^{j_3} d(k_3^{(2)}, k_3^{(3)}))^M}
\end{aligned}
$$

The first three summands are easy to bound, indeed it is enough to notice that

$$
\begin{aligned}
& \frac{C}{B^{4j_3}} \prod_{i=1}^{4} \sum_{k_3^i} \{\mathbb{I}(d(\left(k_1^{(1)}, k_1^{(4)} > 3r\right), d\left(k_1^{(2)}, k_1^{(3)} > 3r\right)\right) \\
\times\ & \frac{1}{(1 + B^{j_1} d(k_1^{(1)}, k_1^{(2)}))^M} \frac{1}{(1 + B^{j_1} d(k_1^{(3)}, k_1^{(4)}))^M} \frac{1}{(1 + B^{j_2} d(k_2^{(1)}, k_2^{(3)}))^M} \\
\times\ & \frac{1}{(1 + B^{j_2} d(k_2^{(2)}, k_2^{(4)}))^M} \frac{1}{(1 + B^{j_3} d(k_3^{(1)}, k_3^{(4)}))^M} \frac{1}{(1 + B^{j_3} d(k_3^{(2)}, k_3^{(3)}))^M}
\end{aligned}
$$

$$
\begin{aligned}
\leq\ & \frac{C}{B^{4j_3}} \prod_{i=1}^{4} \sum_{k_3^i} \{\mathbb{I}(d\left(k_1^{(1)}, k_1^{(4)} > 3r\right))\} \frac{1}{(1 + B^{j_3} d(k_3^{(1)}, k_3^{(4)}))^M} \frac{1}{(1 + B^{j_3} d(k_3^{(2)}, k_3^{(3)}))^M} \\
\leq\ & \frac{CB^{2j_3}}{B^{4j_3}} \sum_{k_3^{(1)}, k_3^{(4)}} \mathbb{I}\left(d(k_1^{(1)}, k_1^{(4)} > 3r)\right) \frac{1}{(1 + B^{j_3} d(k_3^{(1)}, k_3^{(4)}))^M} \\
\leq\ & \frac{C}{B^{2j_3}} B^{\frac{1}{2}j_1 - j_3} B^{2j_3} = O(B^{\frac{1}{2}j_1 - j_3}) = O(B^{-\frac{1}{2}j_1})
\end{aligned}
$$

because

$$
\begin{aligned}
\sum_{k' \in \mathcal{Z}_{j_3}} \mathbb{I}(d(k, k') \ >\ & r) \frac{1}{(1 + B^{j_3} d(k, k'))^M} \leq CB^{2j_3} \int_r^\pi \frac{\sin\theta}{(1 + B^{j_3}\theta)^M} d\theta \\
\leq\ & CB^{2j_3} \int_r^\pi \frac{\theta d\theta}{B^{Mj_3}\theta^M} \leq CB^{(M-2)(\frac{1}{2}j_1 - j_3)},
\end{aligned}
$$

by taking $r = B^{-j_1/2}$, compare *Lemma 10* in [7]. The argument for the second and third term is analogous. Concerning the last summand, we recall that

$$Card\{k' \in \mathcal{V}_{j_1/2}(x_i) \cap \mathcal{X}_{j_3}\} \approx B^{2(j_3 - j_1/2)}, \tag{3.14}$$

for every $k \in \mathcal{X}_{j_1/2}$. Now denote

$$\Omega(k; j_3) := \{(k_1, k_2, k_3, k_4) : k_1, k_2, k_3, k_4 \in \mathcal{X}_{j_3}, k_1 \in \mathcal{V}_{j_1/2}(k), d(k_1, k_4) \leq 3r, d(k_2, k_3) \leq 3r\} ,$$

where $r = \kappa B^{-\frac{1}{2}j_1}$. Heuristically, the idea is to split $\Omega(k)$ into regions where (k_1, k_2, k_3, k_4) are each "close" to all three others, and regions where they are close two by two but the two pairs need not belong to the same neighbourhood. More precisely, we take $\Omega(k; j_3) \subseteq \Delta_1 \cup \Delta_2$, where

$$\Delta_1 = \{(k_1, k_2, k_3, k_4) : k_1, \cdots, k_4 \in B_{6r}(k), k_1 \in \mathcal{V}_{j_1/2}(k)\}$$

and

$$\Delta_2 = \{(k_1, k_2, k_3, k_4) : k_1 \in \mathcal{V}_{j_1/2}(k) \ , \ k_2 \in S^2 / B_{6r}(k)\}.$$

We can hence define

$$\Omega_1(k; j_3) := \Omega(k; j_3) \cap \Delta_1 \ , \ \Omega_2(k; j_3) := \Omega(k; j_3) \cap \Delta_2 \ .$$

In the region $\Omega_1(k; j_3)$, the idea is to keep k fixed and then proceed by evaluating the cardinality of $B_{6r}(k)$; in view of (3.14), this leads to

$$\begin{aligned} (3.13) \ &\leq \ \frac{C}{B^{4j_3}} \sum_{k \in \mathcal{X}_{j_1/2}} \sum_{(k_1,k_2,k_3,k_4) \in \Omega_1(k;j_3),} \frac{1}{(1 + B^{j_3} d(k_1, k_4))^M} \frac{1}{(1 + B^{j_3} d(k_2, k_3))^M} \\ &= \ O(\frac{B^{4(j_3 - j_1/2) + j_1}}{B^{4j_3}}) = O(B^{-j_1}) \ , \end{aligned}$$

by (2.24) and *Lemma 4.8* in [67]. On the other hand, in the region $\Omega_2(k)$, we exploit the fact that $d(k_1,k_2), d(k_3,k_4) \geq 3r = 3\kappa B^{-j_1/2}$, so that we obtain the upper bound (for some $C>0$)

$$
\begin{aligned}
(3.13) \;\leq\; & \frac{C}{B^{4j_3}} \sum_{k\in\mathcal{X}_{j_1/2}} \sum_{(k_3^{(1)},\cdots,k_3^{(4)})\in\Omega_2(k;j_3)} \frac{1}{(1+B^{j_2}d(k_2^{(1)},k_2^{(2)}))^M} \frac{1}{(1+B^{j_2}d(k_2^{(3)},k_2^{(4)}))^M} \\
& \times \frac{1}{(1+B^{j_3}d(k_3^{(1)},k_3^{(4)}))^M} \frac{1}{(1+B^{j_3}d(k_3^{(2)},k_3^{(3)}))^M}
\end{aligned}
$$

$$
\begin{aligned}
\leq\; & \frac{C}{B^{4j_3}} \sum_{k\in\mathcal{X}_{j_1}} \sum_{(k_1,k_2,k_3,k_4)\in\Omega_2(k;j_3)} \frac{1}{(1+C(B^{j_2-j_1/2}))^{2M}} \frac{1}{(1+B^{j_3}d(k_1,k_4))^M} \frac{1}{(1+B^{j_3}d(k_2,k_3))^M} \\
=\; & O(\frac{B^{4j_3-2M(j_2-j_1/2)}}{B^{4j_3}}) = O(B^{-Mj_1}) \,.
\end{aligned}
$$

The proof for the remaining terms is very similar and hence omitted for brevity's sake.

Similarly, it can be shown that the same argument holds for the term corresponding to diagrams as those depicted in fig3.2.

$$
\begin{aligned}
& \frac{1}{B^{4j_3}} \Pi_{i=1}^4 \sum_{k_3^{(i)}} \frac{1}{(1+B^{j_1}d(k_1^{(1)},k_1^{(4)}))^3} \frac{1}{(1+B^{j_1}d(k_1^{(2)},k_1^{(3)}))^3} \frac{1}{(1+B^{j_2}d(k_2^{(1)},k_2^{(2)}))^3} \\
& \times \frac{1}{(1+B^{j_1}d(k_2^{(3)},k_2^{(4)}))^3} \frac{1}{(1+B^{j_3}d(k_3^{(1)},k_3^{(2)}))^3} \frac{1}{(1+B^{j_3}d(k_3^{(3)},k_3^{(4)}))^3} \\
\leq\; & \frac{1}{B^{4j_3}} \Pi_{i=1}^4 \sum_{k_3^{(i)}} \frac{1}{(1+B^{j_1}d(k_1^{(1)},k_1^{(4)}))^3} \frac{1}{(1+B^{j_1}d(k_1^{(2)},k_1^{(3)}))^3} \\
& \times \frac{1}{(1+B^{j_3}d(k_3^{(1)},k_3^{(2)}))^3} \frac{1}{(1+B^{j_3}d(k_3^{(3)},k_3^{(4)}))^3}
\end{aligned}
$$

Again we divide the region $\Omega(k)$ into two parts $\Omega_1(k)$ and $\Omega_2(k)$,taking $r = \kappa 2^{-\frac{1}{2}j_1}$

$$\begin{aligned}
& \frac{1}{B^{4j_3}} \sum_{k\in\mathcal{X}_{\frac{1}{2}j_1}} \sum_{(k_3^{(1)},\cdots,k_3^{(4)})\in\Omega_1(k)\cap\mathcal{X}_{j_3}} \frac{1}{(1+B^{j_1}d(k_1^{(1)},k_1^{(4)}))^3}\frac{1}{(1+B^{j_1}d(k_1^{(2)},k_1^{(3)}))^3} \\
& \times\frac{1}{(1+B^{j_3}d(k_3^{(1)},k_3^{(2)}))^3}\frac{1}{(1+B^{j_3}d(k_3^{(3)},k_3^{(4)}))^3}\delta... \\
\leq\ & \frac{1}{B^{4j_3}} C_M B^{4(j_3-\frac{1}{2}j_1)+2\times\frac{1}{2}j_1} \\
\approx\ & O(B^{-j_1})
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{B^{4j_3}} \sum_{k\in\mathcal{X}_{\frac{1}{2}j_1}} \sum_{(k_3^{(1)},\cdots,k_3^{(4)})\in\Omega_2(k)\cap\mathcal{X}_{j_3}} \frac{1}{(1+B^{j_1}d(k_1^{(1)},k_1^{(4)}))^3}\frac{1}{(1+B^{j_1}d(k_1^{(2)},k_1^{(3)}))^3} \\
& \times\frac{1}{(1+B^{j_3}d(k_3^{(1)},k_3^{(2)}))^3}\frac{1}{(1+B^{j_3}d(k_3^{(3)},k_3^{(4)}))^3}\delta... \\
\leq\ & \frac{1}{B^{4j_3}} C_M 2^{4j_3}\frac{1}{(1+B^{\frac{1}{2}j_1})^6} \\
\approx\ & O(B^{-3j_1})
\end{aligned}$$

Thus the proof is finished. ■

3.1.3 Unknown Angular Power Spectrum

As the final result of this Section, we wish to focus on the case where the variance of the needlets coefficients is unknown and estimated from the data. A natural estimator for σ_j is provided by

$$\widetilde{\sigma}_j^2 = \frac{1}{N_j}\sum_{\xi_{jk}\in\mathcal{X}_j} \left|\beta_{jk}\right|^2$$

where as before $N_j = card\{\mathcal{X}_j\} \approx B^{2j}$. We define our studentized statistics as $\widetilde{\beta}_{jk} := \beta_{jk}/\widetilde{\sigma}_j$ and we then consider

$$\widetilde{I}_{j_1j_2j_3} = \sum_{k_3} \widetilde{\beta}_{j_1k_1}\widetilde{\beta}_{j_2k_2}\widetilde{\beta}_{j_3k_3}\delta_{j_1j_2j_3}(k_1,k_2,k_3)h_{j_1j_2j_3}(k_1,k_2,k_3)\ .$$

Our next result shows that this studentization procedure has no effect on asymptotic behaviour.

Theorem 7 *As $j_1 \to \infty$, we have $\left\{\sigma^2_{j_1}\sigma^2_{j_2}\sigma^2_{j_3}\right\}^{-1}\widetilde{\sigma}^2_{j_1}\widetilde{\sigma}^2_{j_2}\widetilde{\sigma}^2_{j_3} \longrightarrow 1$ in probability, and hence*

$$\widetilde{I}_{j_1j_2j_3} \to_d N(0,1)\ .$$

Proof. By a standard application of Slutzki's theorem, the weak convergence result follows immediately from the consistency of the variance estimator. We provide a proof for the three cases separately.

For case 1), i.e. $j_1 < j_2 < j_3$ we have immediately $E\left[\left\{\sigma^2_{j_1}\sigma^2_{j_2}\sigma^2_{j_3}\right\}^{-1}\widetilde{\sigma}^2_{j_1}\widetilde{\sigma}^2_{j_2}\widetilde{\sigma}^2_{j_3}\right] = 1$. Moreover

$$Var[\widetilde{\sigma}^2_{j_1}\widetilde{\sigma}^2_{j_2}\widetilde{\sigma}^2_{j_3}] = \frac{1}{N^2_{j_1}N^2_{j_2}N^2_{j_3}}\prod_{i=1}^{3}\sum_{\xi_{j_ik_i},\xi_{j_ik'_i}\in\mathcal{X}_{j_i}} E[\left|\beta_{j_ik_i}\right|^2\left|\beta_{j_ik'_i}\right|^2] - \sigma^4_{j_1}\sigma^4_{j_2}\sigma^4_{j_3}\ ,$$

where

$$\begin{aligned}\sum_{\xi_{j_1k_1},\xi_{j_1k'_1}\in\mathcal{X}_{j_1}} E[|\beta_{j_1k_1}|^2|\beta_{j_1k'_1}|^2] &= (\sum_{\xi_{j_1k_1}\in\mathcal{X}_{j_1}} E|\beta_{j_1k_1}|^2)^2 + 2\sum_{\xi_{j_1k_1},\xi_{j_1k'_1}\in\mathcal{X}_{j_1}} |E[\beta_{j_1k_1},\beta_{j_1k'_1}]|^2 \\ &= N^2_{j_1}\sigma^4_{j_1} + O\left(N_{j_1}\sigma^4_{j_1}\right)\ .\end{aligned}$$

Hence

$$
\begin{aligned}
& Var[\widetilde{\sigma}_{j_1}^2 \widetilde{\sigma}_{j_2}^2 \widetilde{\sigma}_{j_3}^2] \\
= \; & \frac{1}{N_{j_1}^2}\frac{1}{N_{j_2}^2}\frac{1}{N_{j_3}^2}(N_{j_1}^2 + O\left(N_{j_1}\right))(N_{j_2}^2 + O\left(N_{j_2}\right))(N_{j_3}^2 + O\left(N_{j_3}\right))\sigma_{j_1}^4\sigma_{j_2}^4\sigma_{j_3}^4 - \sigma_{j_1}^4\sigma_{j_2}^4\sigma_{j_3}^4 \\
= \; & O\left(\frac{1}{N_{j_1}N_{j_2}N_{j_3}}\right)\sigma_{j_1}^4\sigma_{j_2}^4\sigma_{j_3}^4.
\end{aligned}
$$

For case 2), we focus on the case where $j_1 = j_2 < j_3$; the remaining part of the proof is nearly identical. First note that

$$
\begin{aligned}
E\left\{\frac{\widetilde{\sigma}_{j_1}^4\widetilde{\sigma}_{j_2}^2}{\sigma_{j_1}^4\sigma_{j_2}^2}\right\} &= \frac{\sum_{\xi_{j_2k_2}\in\mathcal{X}_{j_2}} E|\beta_{j_2k_2}|^2}{N_{j_2}\sigma_{j_1}^4\sigma_{j_2}^2}\{(\frac{4\pi}{N_{j_1}}\sum_{\xi_{j_1k_1}\in\mathcal{X}_{j_1}} E|\beta_{j_1k_1}|^2)^2 \\
&\quad + \frac{2(4\pi)^2}{N_{j_1}^2}\sum_{\xi_{j_1k_1},\xi_{j_1k_1'}\in\mathcal{X}_{j_1}} |E[\beta_{j_1k_1},\beta_{j_1k_1'}]|^2\} \\
&= \frac{1}{\sigma_{j_1}^4\sigma_{j_2}^2}(\sigma_{j_1}^4 + O\left(\frac{\sigma_{j_1}^4}{N_{j_1}}\right))\sigma_{j_2}^2 = 1 + O\left(\frac{1}{N_{j_1}}\right).
\end{aligned}
$$

Likewise

$$
\begin{aligned}
Var[\widetilde{\sigma}_{j_1}^4\widetilde{\sigma}_{j_2}^2] &= E(\frac{1}{N_{j_1}}\sum_{\xi_{j_1k_1}\in\mathcal{X}_{j_1}}|\beta_{j_1k_1}|^2)^4 E(\frac{1}{N_{j_2}}\sum_{\xi_{j_2k_2}\in\mathcal{X}_{j_2}}|\beta_{j_2k_2}|^2)^2 \\
&\quad - \left(\sigma_{j_1}^4 + O\left(\frac{\sigma_{j_1}^4}{N_{j_1}}\right)\right)^2\sigma_{j_2}^4 \\
&= \frac{1}{N_{j_1}^4N_{j_2}^2}E[\prod_{i=1}^4\sum_{\xi_{j_1k_1^I}\in\mathcal{X}_{j_1}}|\beta_{j_1k_1^1}|^2|\beta_{j_1k_1^2}|^2|\beta_{j_1k_1^3}|^2|\beta_{j_1k_1^4}|^2]E(\sum_{\xi_{j_2k_2}\in\mathcal{X}_{j_2}}|\beta_{j_2k_2}|^2)^2 \\
&\quad -(\sigma_{j_1}^4 + O\left(\frac{\sigma_{j_1}^4}{N_{j_1}}\right))^2\sigma_{j_2}^4
\end{aligned}
$$

The first ingredient in the last equality is equivalent to

$$[(\sum_{\xi_{j_1k_1}\in\mathcal{X}_{j_1}} E|\beta_{j_1k_1}|^2)^4 + 12(\sum_{\xi_{j_1k_1}\in\mathcal{X}_{j_1}} E|\beta_{j_1k_1}|^2)^2 \sum_{\xi_{j_1k_1},\xi_{j_1k_1'}\in\mathcal{X}_{j_1}} |E[\beta_{j_1k_1},\beta_{j_1k_1'}]|^2$$

$$+8(\sum_{\xi_{j_1k_1}\in\mathcal{X}_{j_1}} E|\beta_{j_1k_1}|^2)(\prod_{i=1}^{3}\sum_{\xi_{j_1k_1^i}\in\mathcal{X}_{j_1}} E[\beta_{j_1k_1^1},\beta_{j_1k_1^2}]E[\beta_{j_1k_1^2},\beta_{j_1k_1^3}]E[\beta_{j_1k_1^1},\beta_{j_1k_1^3})]$$

$$\times\frac{\sigma_{j_2}^4(1+O\left(\frac{1}{N_{j_2}}\right))}{N_{j_1}^4}\ .$$

By (2.24) and [67] (Lemma 4.8) , we have

$$\prod_{i=1}^{3}\sum_{\xi_{j_1k_1^I}\in\mathcal{X}_{j_1}} E[\beta_{j_1k_1^1},\beta_{j_1k_1^2}]E[\beta_{j_1k_1^2},\beta_{j_1k_1^3}]E[\beta_{j_1k_1^1},\beta_{j_1k_1^3}$$

$$\leq \sigma_{j_1}^6\prod_{i=1}^{3}\sum_{\xi_{j_1k_1^I}\in\mathcal{X}_{j_1}} \frac{1}{(1+B^{j_1}d(k_1^1,k_1^2))^3}\frac{1}{(1+B^{j_1}d(k_1^2,k_1^3))^3}\frac{1}{(1+B^{j_1}d(k_1^1,k_1^3))^3}$$

$$\leq \sigma_{j_1}^6\sum_{\xi_{j_1k_1^1},\xi_{j_1k_1^3}\in\mathcal{X}_{j_1}} \frac{C_M}{(1+B^{j_2}d(k_1^1,k_1^3))^6} \leq C_M N_{j_1}\sigma_{j_1}^6\ .$$

Hence the variance is bounded by the absolute value of

$$[N_{j_1}^4\sigma_{j_1}^8 + 8\sigma_{j_1}^8 O\left(N_{j_1}^2\right) + 12\sigma_{j_1}^8 O\left(N_{j_1}^3\right)] \times \frac{\sigma_{j_2}^4(1+O\left(\frac{1}{N_{j_2}}\right))}{N_{j_1}^4}$$

$$-(\sigma_{j_1}^4 + O\left(\frac{\sigma_{j_1}^4}{N_{j_1}}\right))^2\sigma_{j_2}^4\ .$$

and we immediately get that

$$Var[\frac{\widetilde{\sigma}_{j_1}^4\widetilde{\sigma}_{j_2}^2}{\sigma_{j_1}^4\sigma_{j_2}^2}] = O\left((\frac{1}{N_{j_1}}+\frac{1}{N_{j2}})\right) = O(\frac{1}{N_{j_1}}).$$

The proof for case 3) is an easy consequence of results by [7, 8]. ■

3.2 Convergence to Multiparameter Gaussian Processes

Our aim in this Section is to extend the previous results to functional convergence theorems. The motivation for such an extension can be easily explained. Indeed, from the applications points of view, practitioners are typically interested not only at the possible existence of non-Gaussianity and/or other features, but also to their location in frequency space. If we focus for instance on cosmological applications, which are the main driving rationale behind our work, it is important to recall that the existence of possible non-Gaussianities takes a very different meaning according to the scales where they are located, so that a suitable statistical procedure should provide information not only on their existence, but also on their position in the frequency domain. As an example, a huge debate has arisen in the Cosmological literature on the possible existence of a non-Gaussian "Cold Spot" in CMB data, much of the related literature concerning the determination of the angular scale of such features, see for instance [18, 19]. Concerning this feature, it may even be of interest to test for Gaussianity only on a subspace of the sphere (this is indeed what happens in practice, because of missing observations). The modification of (3.6) under these circumstances is straightforward: we would simply restrict our sum to a subset of the cubature points. Our following discussion would be asymptotically unaltered.

In [14], [56], it was proposed to build alternative forms of partial sum process from the bispectrum $B_{l_1 l_2 l_3}$, and to use them as a probe of non-Gaussianity at various

scales. All different proposals were univariate, in the following sense. Assume the resolution of the experiment is such that frequencies up to $l = 1, ..., L$ are observed; the partial sums were then run only on a subset of configurations with cardinality L, whereas the number of multipole combinations (l_1, l_2, l_3) which would be available for statistical analysis is in the order of L^3. One of the reasons for this restriction had to do with computational complexity: the evaluation of even a single bispectrum statistic is extremely time consuming, so that the exploration of all possible configurations is likely to be unfeasible even on the greatest supercomputing facilities. On the contrary, needlets are extremely convenient from a computational point of view, and there is no obstacle in considering larger frequency configurations, provided of course that the triangle conditions are satisfied.

We shall hence focus on two possible partial sums processes, which correspond broadly to cases 1) and 2) of the previous section; more precisely

$$J_{1L}(r_1, r_2) = \frac{1}{L} \sum_{j_1=1}^{[Lr_1]} \sum_{m=0}^{[Lr_2]} \widehat{I}_{j_1, j_1+K+m, j_1+K+m} \ , \tag{3.15}$$

$$J_{2L}(r) = \frac{1}{\sqrt{L}\sqrt{\sum_{m_1=0}^{N_1}(N(m_1)+1)}} \sum_{j=1}^{[Lr]} \sum_{m_1=0}^{N_1} \sum_{m_2=0}^{N(m_1)} \widehat{I}_{j_1, j_1+K+m_1, j_1+2K+m_1+m_2} \ , \tag{3.16}$$

where

$$\widehat{I}_{j_1,j_2,j_3} := \frac{I_{j_1,j_2,j_3}}{\sqrt{EI^2_{j_1,j_2,j_3}}},$$

$K \geq 0$ is some integer satisfying the constraint determined in Lemma 3,

$$N_1 = \max\left\{m : 1 + B^{K+m} \geq B^{2K+m}\right\}$$

$$N(m) := \max\left\{m : \log_B\left(1 + B^{K+m_1}\right) - 2K - m_1 \geq m\right\} \ .$$

Theorem 8 *a) As* $L \to \infty$

$$J_{1L}(r_{1,r_2}) \Rightarrow W(r_1, r_2) \ , \ in \ D[0,1]^2 \ , \tag{3.17}$$

where $W(.)$ *is two-dimensional Brownian sheet, i.e. the zero mean Gaussian process with covariance function* $EW(r_1, r_2)W(s_1, s_2) = (r_1 \wedge s_1)(r_2 \wedge s_2)$.

b) As $L \to \infty$

$$J_{2L}(r) \Rightarrow W(r) \ , \ in \ D[0,1] \tag{3.18}$$

where $W(.)$ *is standard Brownian Motion.*

Here, $\Rightarrow$ *denotes weak convergence in the sense of* [11]*,* $D[0,1]^p, p \in \mathbb{N}$ *is the usual multidimensional Skorohod space.*

Proof. We start from (3.17); as usual, we need to establish convergence of the finite-dimensional distributions and tightness. Obviously,

$$EJ_{1L}(r_1, r_2) = 0 \ ,$$

and because $E[\hat{I}_{j_1,j_2,j_2}\hat{I}_{j_1',j_2',j_2'}] = \delta_{j_1}^{j_1'}\delta_{j_2}^{j_2'}$,

$$\begin{aligned} EJ_{1L}(r_1,r_2)J_{1L}(s_1,s_2) &= \frac{1}{L^2}\sum_{j_1=1}^{[Lr_1]}\sum_{m=0}^{[Lr_2]}\sum_{j_1'=1}^{[Ls_1]}\sum_{m'=0}^{[Ls_2]} E\hat{I}_{j_1,j_1+K+m,j_1+K+m}\hat{I}_{j_1,j_1+K+m',j_1+K+m'} \\ &= \frac{1}{L^2}\sum_{j_1=1}^{[Lr_1]\wedge[Ls_1]}\sum_{m=0}^{[Lr_2]\wedge[Ls_2]} E\widehat{I}^2_{j_1,j_1+K+m,j_1+K+m} \\ &= \sum_{j_1=1}^{[Lr_1]\wedge[Ls_1]}\frac{[Lr_2]\wedge[Ls_2]}{L^2} \to [r_2\wedge s_2]\,[r_1\wedge s_1] \ . \end{aligned}$$

To establish Gaussianity, we can again rely on the results by [69] and proceed with the bounds for the fourth-order cumulants. As the computations are very much the same as in the previous Section, we omit the details for brevity's sake. To consider tightness, we use the classical criteria given for instance in [79]. Define first the two-dimensional increments

$$J_{1L}((s_1,r_1]\times(s_2,r_2]) := J_{1L}(r_1,r_2) - J_{1L}(r_1,s_2) - J_{1L}(s_1,r_2) + J_{1L}(s_1,s_2) \ .$$

It is again a standard computation to show that, as $L\to\infty$,

$$\begin{aligned} EJ^2_{1L}((s_1,r_1]\times(s_2,r_2]) &= \frac{1}{L^2}\{\sum_{j_1=[Ls_1]+1}^{[Lr_1]}\sum_{m=[Ls_2]+1}^{[Lr_2]} E\hat{I}^2_{j_1,j_1+K+m,j_1+K+m}\} \\ &= \frac{1}{L^2}\{[Lr_1]-[Ls_1]\}\{[Lr_2]-[Ls_2]\} \\ &\le 4\{r_1-s_1\}\{r_2-s_2\} \ . \end{aligned}$$

We can then establish tightness by showing that

1) $s_i \le t_i \le r_i, i=1,2$

$$E[J^2_{1L}((s_1,t_1]\times(s_2,t_2])J^2_{1L}((t_1,r_1]\times(t_2,r_2])]$$

$$= \frac{1}{L^4}E\Big(\sum_{j_1,j_1'=[Ls_1]}^{[Lt_1]}\sum_{\substack{j_2=j_1+K+m,\\ m=[Ls_2]+1}}^{[Lt_2]}\sum_{\substack{j_2'=j_1'+K+m',\\ m'=[Ls_2]+1}}^{[Lt_2]}\hat{I}_{j_1j_2j_2}\hat{I}_{j_1'j_2'j_2'}\Big)\Big(\sum_{j_1,j_1'=[Lt_1]}^{[Lr_1]}\sum_{\substack{j_2=j_1+K+m,\\ m=[Lt_2]+1}}^{[Lr_2]}\sum_{\substack{j_2'=j_1'+K+m',\\ m'=[Lt_2]+1}}^{[Lr_2]}\hat{I}_{j_1j_2j_2}\hat{I}_{j_1'j_2'j_2'}\Big)$$

$$\begin{aligned}
&= \frac{1}{L^4}\Big(\sum_{j_1=[Ls_1]}^{[Lt_1]}\sum_{m=[Ls_2]}^{[Lt_2]}E\hat{I}^2_{j_1,j_1+K+m,j_1+K+m}\Big)\Big(\sum_{j_1'=[Lt_1]}^{[Lr_1]}\sum_{m'=[Lt_2]}^{[Lr_2]}E\hat{I}^2_{j_1',j_1'+K+m',j_1'+K+m'}\Big)\\
&\le 16(r_1-t_1)(r_2-t_2)(t_1-s_1)(t_2-s_2)\\
&\le 4(r_1-s_1)^2(r_2-s_2)^2.
\end{aligned}$$

2) $s_1 \le r_1, s_2 \le t_2 \le r_2$,

$$E[J_{1L}((s_1,r_1]\times(s_2,t_2])J_{1L}((s_1,r_1]\times(t_2,r_2])]^2$$

$$= \frac{1}{L^4}E\Big(\sum_{j_1,j_1'=[Ls_1]}^{[Lt_1]}\sum_{\substack{j_2=j_1+K+m,\\ m=[Ls_2]+1}}^{[Lt_2]}\sum_{\substack{j_2'=j_1'+K+m',\\ m'=[Ls_2]+1}}^{[Lt_2]}\hat{I}_{j_1j_2j_2}\hat{I}_{j_1'j_2'j_2'}\Big)\Big(\sum_{j_1,j_1'=[Ls_1]}^{[Lr_1]}\sum_{\substack{j_2=j_1+K+m,\\ m=[Lt_2]+1}}^{[Lr_2]}\sum_{\substack{j_2'=j_1'+K+m',\\ m'=[Lt_2]+1}}^{[Lr_2]}\hat{I}_{j_1j_2j_2}\hat{I}_{j_1'j_2'j_2'}\Big)$$

$$\begin{aligned}
&= \frac{1}{L^4}\Big(\sum_{j_1=[Ls_1]}^{[Lr_1]}\sum_{m=[Ls_2]}^{[Lt_2]}E\hat{I}^2_{j_1,j_1+K+m,j_1+K+m}\Big)\Big(\sum_{j_1'=[Ls_1]}^{[Lr_1]}\sum_{m'=[Lt_2]}^{[Lr_2]}E\hat{I}^2_{j_1',j_1'+K+m',j_1'+K+m'}\Big)\\
&\quad+\frac{1}{L^4}\sum_{\substack{j_1^{(i)}=[Ls_1]+1,\\ i=1,..,4}}^{[Lr_1]}\sum_{\substack{m^{(i)}=[Ls_2];\\ i=1,2}}^{[Lt_2]}\sum_{\substack{m^{(i)}=[Lt_2];\\ i=3,4}}^{[Lr_2]}\sum_{\gamma\in\Gamma_C}\rho(\gamma;\prod_{i=1}^{4}\widehat{I}_{j_1^{(i)},j_1^{(i)}+K+m^{(i)},j_1^{(i)}+K+m^{(i)}})
\end{aligned}$$

For the first part it is easy to see that it is bounded by $16(r_1-s_1)^2(r_2-t_2)(t_2-s_2)$;

for the second part, for $j_1^{(i)}, j_2^{(i)}$ in each of their domain, we have,

$$\begin{aligned}
&\sum_{\gamma\in\Gamma_C}\rho(\gamma;\prod_{i=1}^{4}\widehat{I}_{j_1^{(i)}j_2^{(i)}j_2^{(i)}})\\
=\;&\sum_{\gamma\in\Gamma_{1C}}\rho(\gamma;\widehat{I}^2_{j_1j_2^{(1)}j_2^{(1)}}\widehat{I}^2_{j_1j_2^{(3)}j_2^{(3)}})\prod_{i=1}^{4}\delta^{j_1^{(i)}}_{j_1}\delta^{j_2^{(2)}j_2^{(4)}}_{j_2^{(1)}j_2^{(3)}}\\
&+\sum_{\gamma\in\Gamma_{2C}}\rho(\gamma;\widehat{I}^2_{j_1j_1'j_1'}\widehat{I}^2_{j_1'j_2^{(3)}j_2^{(3)}})\prod_{i=1}^{2}\delta^{j_1^{(i)}}_{j_2^{(i)}}\delta^{j_2^{(2)}j_2^{(4)}}_{j_2^{(1)}j_2^{(3)}},
\end{aligned}$$

where Γ_{1C} denotes the graphs with cliques (all nodes connected with all others), and Γ_{2C} refers to graphs with loops of order two; these two disjoint classes cover all possible connected graphs with four nodes. We have

$$\begin{aligned} & \frac{1}{L^4} \sum_{\substack{j_1^{(i)}=[Ls_1]+1, \\ i=1,..,4}}^{[Lr_1]} \sum_{\substack{m^{(i)}=[Ls_2]; \\ i=1,2}}^{[Lt_2]} \sum_{\substack{m^{(i)}=[Lt_2]; \\ i=3,4}}^{[Lr_2]} \sum_{\gamma \in \{fig2\}} \rho(\gamma; \prod_{i=1}^{4} \widehat{I}_{j_1^{(i)}, j_1^{(i)}+K+m^{(i)}, j_1^{(i)}+K+m^{(i)}}) \\ \leq & \frac{C}{L^4} \sum_{j_1, j_1'=[Ls_1]+1}^{[Lr_1]} \sum_{\substack{m^{(i)}=[Ls_2]; \\ i=1,2}}^{[Lt_2]} \sum_{\substack{m^{(i)}=[Lt_2]; \\ i=3,4}}^{[Lr_2]} 1 \leq C'(r_1-s_1)^2(r_2-t_2)(t_2-s_2) \end{aligned}$$

from which we obtain

$$\begin{aligned} & E[J_{1L}((s_1,r_1]\times(s_2,t_2])J_{1L}((s_1,r_1]\times(t_2,r_2])]^2 \\ \leq & C(r_1-s_1)^2(r_2-t_2)(t_2-s_2) \end{aligned}$$

Similarly, we can get the same result for

3) $s_1 \leq t_1 \leq r_1, s_2 \leq r_2$, that is

$$\begin{aligned} & E[J_{1L}((s_1,t_1]\times(s_2,r_2])J_{1L}((t_1,r_1]\times(s_2,r_2])]^2 \\ \leq & C(r_2-s_2)^2(r_1-t_1)(t_1-s_1) \end{aligned}$$

This concludes the proof of (3.17).

For (3.18), we start again from the convergence of the finite-dimensional distributions; for notational simplicity, we stick to the univariate case. It is obvious that $EJ_{2L}(r)=0$; on the other hand,

$$EJ_{2L}(r)J_{2L}(s)$$

$$
\begin{aligned}
&= \sum_{j_1=1}^{[Lr]} \sum_{j_1'=1}^{[Ls]} \left\{ \sum_{m_1,m_1'=0}^{N_1} \sum_{m_2=0}^{N(m_1)} \sum_{m_2'=0}^{N\left(m_1'\right)} \right\} \frac{E\widehat{I}_{j_1,j_1+K+m_1,j_1+2K+m_1+m_2} \widehat{I}_{j_1',j_1'+K+m_1',j_1'+2K+m_1'+m_2'}}{L\sum_{m_1=0}^{N_1}(N\left(m_1\right)+1)} \\
&= \frac{1}{L\sum_{m_1=0}^{N_1}(N\left(m_1\right)+1)} \sum_{j_1=1}^{[Lr]\wedge[Ls]} \sum_{m_1=0}^{N_1} \sum_{m_2=0}^{N(m_1)} E\widehat{I}^2_{j_1,j_1+K+m_1,j_1+2K+m_1+m_2} \\
&= \frac{1}{L}\left([Lr]\wedge[Ls]-1\right) \to [r\wedge s]\,, \text{ if } \min\{r,s\}>0.
\end{aligned}
$$

For Gaussianity, we analyze once again fourth-order cumulants, i.e. the connected components in the expansion of the fourth moment. As before, we need only focus on connected diagrams with four nodes, which can be partitioned into two classes: the cliques, where all nodes are connected with all three others, and diagrams with a loop of order 2. As before, these terms can be bounded by

$$
\rho(\gamma; \prod_{i=1}^{4} \widehat{I}_{j_1^{(i)} j_2^{(i)} j_3^{(i)}}) = O(B^{-\max\{j_1^{(i)}\}/2})\,,
$$

because for instance

$$
\begin{aligned}
&\frac{1}{B^{j_3^{(1)}+j_3^{(2)}+2j_3^{(3)}}} \sum_{k_3^{(1)},k_3^{(2)},k_3^{(3)},k_3^{(4)}} \frac{1}{(1+B^{j_1^{(1)}} d(k_1^{(1)},k_1^{(2)}))^M} \frac{1}{(1+B^{j_2^{(1)}} d(k_2^{(1)},k_1^{(4)}))^M} \\
&\times \frac{1}{(1+B^{j_2^{(2)}} d(k_2^{(2)},k_1^{(3)}))^M} \frac{1}{(1+B^{j_3^{(1)}} d(k_3^{(1)},k_2^{(3)}))^M} \frac{1}{(1+B^{j_3^{(2)}} d(k_3^{(2)},k_2^{(4)}))^M} \frac{1}{(1+B^{j_3^{(3)}} d(k_3^{(3)},k_3^{(4)}))^M} \\
&\leq \frac{\sum_{k_3^{(3)},k_3^{(4)}} \frac{1}{(1+B^{j_3^{(3)}} d(k_3^{(3)},k_3^{(4)}))^M}}{B^{j_3^{(1)}+j_3^{(2)}+2j_3^3}} \sum_{k_3^{(1)}} \frac{1}{(1+B^{j_3^{(1)}} d(k_3^{(1)},k_2^{(3)}))^M} \sum_{k_3^{(2)}} \frac{1}{(1+B^{j_3^{(2)}} d(k_3^{(2)},k_2^{(4)}))^M} \\
&\leq \frac{1}{B^{j_3^{(1)}+j_3^{(2)}+2j_3^3}} \sum_{k_3^{(3)},k_3^{(4)}} \frac{C_M}{(1+B^{j_3^{(3)}} d(k_3^{(3)},k_3^{(4)}))^M} \approx \frac{C_M B^{2j_3^3}}{B^{j_3^{(1)}+j_3^{(2)}+2j_3^3}} = O(B^{-(j_3^{(1)}+j_3^{(2)})})\,.
\end{aligned}
$$

To sum up

$$\begin{aligned}
&\prod_{i=1}^{4}\left\{\sum_{j_1^{(i)}=1}^{[Lr]}\sum_{m_1^{(i)}=0}^{N_1}\sum_{m_2^{(i)}=0}^{N\left(m_1^{(i)}\right)}\right\}\sum_{\gamma\in\Gamma_C}\rho\big(\gamma;\prod_{i=1}^{4}\widehat{I}_{j_1^{(i)},j_1^{(i)}+K+m_1^{(i)},j_1^{(i)}+2K+m_1^{(i)}+m_2^{(i)}}\big)\\
\leq\ & C[Lr]\left(\sum_{m_1=0}^{N_1}\left(N\left(m_1\right)+1\right)\right)^2\sum_{j_1=1}^{[Lr]}B^{-j_1/2}=O(L)\,.
\end{aligned}$$

It follows easily that

$$\begin{aligned}
EJ_L^4(r) &= \sum_{j_1=1}^{[Lr]}\sum_{m_1=0}^{N_1}\sum_{m_2=0}^{N(m_1)}\frac{3\left(E\hat{I}^2_{j_1,j_1+K+m_1,j_1+2K+m_1+m_2}\right)^2}{L^2\left(\sum_{m_1=0}^{N_1}(N\left(m_1\right)+1)\right)^2}\\
&\quad+\prod_{i=1}^{4}\left\{\sum_{j_1^{(i)}=1}^{[Lr]}\sum_{m_1^{(i)}=0}^{N_1}\sum_{m_2^{(i)}=0}^{N\left(m_1^{(i)}\right)}\right\}\frac{\sum_{\gamma\in\Gamma_C}\rho(\gamma;\prod_{i=1}^{4}\widehat{I}_{j_1^{(i)},j_1^{(i)}+K+m_1^{(i)},j_1^{(i)}+2K+m_1^{(i)}+m_2^{(i)}})}{L^2\left(\sum_{m_1=0}^{N_1}(N\left(m_1\right)+1)\right)^2}\\
&= 3|EJ_L^2(r)|^2+O(\frac{1}{L})\,,
\end{aligned}$$

which is enough to conclude the proof for the finite-dimensional distributions, in view of the standard argument from [69] that we used before.

To conclude the proof, we need only to consider tightness in $D([0,1])$. Note that $E[\hat{I}_{j_1,j_2,j_3}\hat{I}_{j_1',j_2',j_3'}]=\delta_{j_1}^{j_1'}\delta_{j_2}^{j_2'}\delta_{j_3}^{j_3'}$, the variance of J_{2L} in $[s,r]$ is provided by

$$\mathbb{E}|J_{2L}(r)-J_{2L}(s)|^2=\sum_{j_1=[Ls]+1}^{[Lr]}\sum_{m_1=0}^{N_1}\sum_{m_2=0}^{N(m_1)}\frac{\mathbb{E}\hat{I}^2_{j_1,j_1+K+m_1,j_1+2K+m_1+m_2}}{L\sum_{m_1=0}^{N_1}(N\left(m_1\right)+1)}\leq 2(r-s)$$

Now we establish our tightness criterion. For any $0 \leq s \leq t \leq r \leq 1$,

$$
\begin{aligned}
& \mathbb{E}|J_{2L}(r) - J_{2L}(t)|^2 |J_{2L}(t) - J_{2L}(s)|^2 \\
= & \sum_{j_1=[Lt]}^{[Lr]} \sum_{m_1=0}^{N_1} \sum_{m_2=0}^{N(m_1)} \frac{\mathbb{E}\widehat{I}^2_{j_1,j_1+K+m_1,j_1+2K+m_1+m_2}}{L\sum_{m_1=0}^{N_1}(N(m_1)+1)} \sum_{j_1'=[Ls]}^{[Lt]} \sum_{m_1'=0}^{N_1} \sum_{m_2'=0}^{N(m_1')} \frac{\mathbb{E}\widehat{I}^2_{j_1',j_1'+K+m_1',j_1'+2K+m_1'+m_2'}}{L\sum_{m_1=0}^{N_1}(N(m_1)+1)} \\
& + \sum_{\substack{j_1^{(i)}=[Lt],\, j_1^{(i)}=[Ls], \\ i=1,2 \qquad i=3,4}}^{[Lr] \qquad [Lt]} \left\{ \prod_{l=1}^{4} \sum_{m_1^{(l)}=0}^{N_1} \sum_{m_2^{(l)}=0}^{N(m_1^{(l)})} \right\} \sum_{\gamma \in \Gamma_C} \frac{\rho(\gamma; \prod_{l=1}^4 \widehat{I}_{j_1^{(i)},j_1^{(l)}+K+m_1^{(l)},j_1^{(l)}+2K+m_1^{(l)}+m_2^{(l)}})}{\left(L\sum_{m_1=0}^{N_1}(N(m_1)+1)\right)^2} \\
\leq & \frac{2}{L^2}([Lr]-[Lt])([Lt]-[Ls]) \leq 4(r-s)^2 \, .
\end{aligned}
$$

Thus we finished the proof of tightness. ■

3.3 Behaviour under non-Gaussianity

In this final Section, we shall provide some quick and informal discussion on the behaviour of our statistics under non-Gaussianity; see [75, 61] for other applications of the needlets to cosmological data analysis. There exist of course a huge variety of non-Gaussian models for spherical random fields, and we shall delay a much more detailed treatment to future work. Our purpose here is different, i.e. we want to provide some heuristic discussion on the expected behaviour of our procedures for physically motivated non-Gaussian models. This provides some guidance to practitioners for applications to CMB data. Much more refined simulations and applications to CMB done are provided in [76]. In that paper, in particular, we used the needlets bispectrum to constrain the non-linear parameter f_{NL}, which measures the amount of non-Gaussianity in CMB under the simplest models of so called inflation.

We start from the expected value of the needlets bispectrum, which is provided by

$$\begin{aligned}
\mathbb{E}I_{j_1j_2j_3} &\approx \frac{1}{\sqrt{B^{2j_3}}}\sum_{k_1k_2k_3}\mathbb{E}\widehat{\beta}_{j_1k_1}\widehat{\beta}_{j_2k_2}\widehat{\beta}_{j_3k_3}\delta_{j_1j_2j_3}(k_1,k_2,k_3) \\
&\approx \frac{1}{\sigma_{j_1}\sigma_{j_2}\sigma_{j_3}}\frac{1}{\sqrt{B^{2j_3}}}\sum_{k_1k_2k_3}\sum_{\substack{l_1,l_2,l_3\\ m_1m_2m_3}} b(\frac{l_1}{B^{j_1}})b(\frac{l_2}{B^{j_2}})b(\frac{l_3}{B^{j_3}})\mathbb{E}\left(a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\right) \\
&\quad \times Y_{l_1m_1}(\xi_{j_1k_1})Y_{l_2m_2}(\xi_{j_2k_2})Y_{l_3m_3}(\xi_{j_3k_3})\delta_{j_1j_2j_3}(k_1,k_2,k_3)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sigma_{j_1}\sigma_{j_2}\sigma_{j_3}}B^{j_3}\sum_{l_1,l_2,l_3=B^{j-1}}^{B^{j+1}}\sum_{m_1m_2m_3} b(\frac{l_1}{B^{j_1}})b(\frac{l_2}{B^{j_2}})b(\frac{l_3}{B^{j_3}})b_{l_1l_2l_3} \\
&\quad \times \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}\sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \\
&\quad \times \left\{\frac{1}{B^{2j_3}}\sum_{k_1k_2k_3} Y_{l_1m_1}(\xi_{j_1k_1})Y_{l_2m_2}(\xi_{j_2k_2})Y_{l_3m_3}(\xi_{j_3k_3})\delta_{j_1j_2j_3}(k_1,k_2,k_3)\right\} .
\end{aligned}$$

Here, we recall that $b_{l_1l_2l_3}$ is the so-called reduced bispectrum (see for instance [52],[56]), which collects the non-Gaussian component in the third order moment $\mathbb{E}\left(a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\right)$, i.e. by definition

$$\mathbb{E}\left(a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\right) = \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}\sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}}b_{l_1l_2l_3} .$$

In the cosmological literature, a very popular model for $b_{l_1l_2l_3}$ is provided by the so-called Sachs-Wolfe bispectrum ([52], equation (21)), which yields

$$b_{l_1l_2l_3} = -6f_{NL}\left\{C_{l_1}C_{l_2} + C_{l_1}C_{l_3} + C_{l_2}C_{l_3}\right\} ,$$

where f_{NL} is a physical constant (see for instance [10, 88]). The Wigner coefficients on the right hand side ensure that the expected value $\mathbb{E}\left(a_{l_1m_1}a_{l_2m_2}a_{l_3m_3}\right)$ is rotation-

ally invariant under a change of coordinate, an obvious consequence of the isotropy of the random field. For our purposes below, it is sufficient to recall that

$$\left(\begin{array}{ccc} l & l & l \\ 0 & 0 & 0 \end{array} \right) \approx \frac{(-1)^{-3l/2}}{l} \, , \, \left(\begin{array}{ccc} l_0 & l & l+l_0 \\ 0 & 0 & 0 \end{array} \right) \approx \frac{(-1)^{-l_0+l}}{\sqrt{l}} \, . \tag{3.19}$$

Now exploiting again the cubature formula (2.18) we obtain as before

$$\begin{aligned} & \left\{ \frac{1}{B^{2j_3}} \sum_{k_1 k_2 k_3} Y_{l_1 m_1}(\xi_{j_1 k_1}) Y_{l_2 m_2}(\xi_{j_2 k_2}) Y_{l_3 m_3}(\xi_{j_3 k_3}) \delta_{j_1 j_2 j_3}(k_1, k_2, k_3) \right\} \\ \approx & \int_{\mathbb{S}^2} Y_{l_1 m_1}(\xi) Y_{l_2 m_2}(\xi) Y_{l_3 m_3}(\xi) d\xi \\ = & \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{array} \right) \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{array} \right) \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \, , \end{aligned}$$

whence

$$\begin{aligned} \mathbb{E} I_{j_1 j_2 j_3} = & \frac{1}{\sigma_{j_1} \sigma_{j_2} \sigma_{j_3}} B^{j_3} \sum_{l_1, l_2, l_3 = B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^{j_1}}) b(\frac{l_2}{B^{j_2}}) b(\frac{l_3}{B^{j_3}}) b_{l_1 l_2 l_3} \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{array} \right)^2 \\ & \times \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \sum_{m_1 m_2 m_3} \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{array} \right)^2 \end{aligned}$$

$$= \frac{B^{j_3}}{\sigma_{j_1} \sigma_{j_2} \sigma_{j_3}} \sum_{l_1, l_2, l_3 = B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^{j_1}}) b(\frac{l_2}{B^{j_2}}) b(\frac{l_3}{B^{j_3}}) b_{l_1 l_2 l_3} \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{array} \right)^2 \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \, ,$$

in view of the orthonormality properties of the Wigner's $3j$ coefficients. To keep the analogy with the cosmological literature, we shall focus on "equilateral" and "squeezed" configurations, see [4], [56]. In the equilateral case $j_1 = j_2 = j_3 = j$ we have

$$\mathbb{E} I_{jjj} = \frac{B^j}{\sigma_{j_1} \sigma_{j_2} \sigma_{j_3}} \sum_{l_1, l_2, l_3 = B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^j}) b(\frac{l_2}{B^j}) b(\frac{l_3}{B^j}) b_{l_1 l_2 l_3} \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{array} \right)^2 \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \, .$$

Now recall $l \approx B^j$, $\sigma_j^3 \approx C_{B^j}^{3/2} B^{3j}$, $b_{l_1 l_2 l_3} \approx f_{NL} l_1^{-\alpha} l_2^{-\alpha}$ so that, using also (3.10)

$$\begin{aligned} & B^j \sum_{l_1,l_2,l_3=B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^j}) b(\frac{l_2}{B^j}) b(\frac{l_3}{B^j}) \frac{b_{l_1 l_2 l_3}}{\sigma_j^3} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi} \\ \simeq \; & B^j \sum_{l_1,l_2,l_3=B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^j}) b(\frac{l_2}{B^j}) b(\frac{l_3}{B^j}) \frac{f_{NL} l_1^{-\alpha} l_2^{-\alpha}}{C_{B^j}^{3/2} B^{3j}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 B^{3j} \\ \simeq \; & B^j \sum_{l_1,l_2,l_3=B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^j}) b(\frac{l_2}{B^j}) b(\frac{l_3}{B^j}) \frac{f_{NL} B^{-2j\alpha}}{B^{-3j\alpha/2} B^{3j}} B^j \simeq f_{NL} B^{2j} B^{-j\alpha/2}. \end{aligned}$$

This suggests the expected value of the needlets bispectrum can either diverge or converge to zero, according to the asymptotic behaviour of the angular power spectrum; in particular, it does diverge for all $\alpha < 4$. On the other hand, for $j_1 << j_2 = j_3$ by an analogous argument we obtain

$$\begin{aligned} \mathbb{E} I_{j_1 j_2 j_2} \; & \simeq \; \frac{f_{NL} B^{j_2}}{B^{-j_1(\alpha/2-1)} B^{-j_2(\alpha-2)}} \sum_{l_1,l_2,l_3=B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^{j_1}}) b(\frac{l_2}{B^{j_2}}) b(\frac{l_3}{B^{j_2}}) l_1^{-\alpha} l_2^{-\alpha} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 B^{j_1} B^{2j_2} \\ & \simeq \; \frac{f_{NL} B^{j_2}}{B^{-j_1(\alpha/2-1)} B^{-j_2(\alpha-2)}} \sum_{l_1,l_2,l_3=B^{j-1}}^{B^{j+1}} b(\frac{l_1}{B^{j_1}}) b(\frac{l_2}{B^{j_2}}) b(\frac{l_3}{B^{j_2}}) B^{-j_1\alpha} B^{-j_2\alpha} B^{j_2} \\ & \simeq \; f_{NL} B^{-j_1\alpha/2} B^{2j_2}. \end{aligned}$$

As for the usual bispectrum, the previous computations suggest that the power is maximized by "squeezing" frequencies, i.e. maximizing the differences between the "side lengths" j_1 and j_2. This is the same sort of qualitative result which was found for the bispectrum in ([56]) and successfully applied to CMB data in [14]. Our heuristic calculations in this Section suggest very clearly that the needlets bispectrum may enjoy the same good power properties, at the same time healing the difficulties

that were met in [14] in the presence of missing observations. As mentioned before, these theoretical arguments have been confirmed by simulations in [76].

Chapter 4
On the Dependence Structure of Wavelet Coefficients for Spherical Random Fields

[2]As we have discussed in the previous chapters, random needlet coefficients enjoy a capital uncorrelation property: namely, for any fixed angular distance, random needlets coefficients are asymptotically uncorrelated as the frequency parameter grows larger and larger, see[7]. We stress again that the meaning of this uncorrelation property must be carefully understood, given the very specific setting of statistical inference in Cosmology. Indeed, CMB can be viewed as a single realization of an isotropic random field on a sphere of a finite radius ([24]). The asymptotic theory is then entertained in the high resolution sense, i.e. it is considered that observations at higher and higher frequencies (smaller and smaller scales) become available with the growing sophistication of CMB satellite experiments (see [56, 57]). Of course, uncorrelation entails independence in the Gaussian case: as a consequence, from the above-mentioned property it follows that an increasing array of asymptotically $i.i.d.$ coefficients can be derived out of a single realization of a spherical random field, making thus possible the introduction of a variety of consistent procedures for testing non-Gaussianity, estimating the angular power spectrum, testing for asymmetries, implementing bootstrap techniques, testing for cross-correlation among CMB

[2] This chapter is based on the paper [54], which has been tentatively accepted for publication on Stochastic Processes and their Applications.

and Large Scale Structure data, and many others, see for instance [7], [8], [33], [53], [61], [75].

Given such a widespread array of techniques which are made feasible by means of the uncorrelation property, it is natural to investigate to what extent this property should be considered unique for the construction in [66, 67], or else whether it is actually shared by other proposals. In particular, we shall focus here on the approach which has been very recently advocated by [34]. The latter approach (which we shall discuss in Section 4.1.2) can be labeled Mexican needlets, for reasons to be made clear later. Its analysis is made particularly interesting by the fact that, as we shall discuss below, the Mexican needlets can be considered asymptotically equivalent to the Spherical Mexican Hat Wavelet (SMHW), which is currently the most popular wavelet procedure in the Cosmological literature (compare Chapter 2.5.2, or see again [64, 18, 19]). As such, the investigation of their properties will fill a theoretical gap which is certainly of interest for CMB data analysis.

Our aim in this chapter is then to investigate the correlation properties of the Mexican needlets coefficients. We shall provide both a positive and a negative result: namely, we will provide necessary and sufficient conditions for the Mexican needlets coefficients to be correlated, depending on the behaviour of the angular power spectrum of the underlying (mean square continuous and isotropic) random fields. In particular, on the contrary of what happens for the needlets in [66, 67], we shall show that there is indeed correlation of the random coefficients when the angular power

spectrum is decaying faster than a certain limit. However, higher order versions (already considered in [34]) of the Mexican needlets can indeed provide uncorrelated coefficients, depending on a parameter which is related to the decay of the angular power spectrum. In some sense, a heuristic rationale under these results can be explained as follows: the correlation among coefficients is introduced basically by the presence in each of these terms of random elements which are fixed (with respect to growing frequencies) in a given realization of the random field, because they depend only on very large scale behaviour (this is known in the Physical literature as a Cosmic Variance effect). Because of the compact support in frequency in the needlets as developed by [66, 67], these low-frequency components are always dropped and uncorrelation is ensured. On the other hand, the same components can be dominant for Mexican needlets, in which cases it becomes necessary to introduce suitably modified versions which are better localized in the frequency domain (i.e., they allow less weight on very low frequency components). In view of the standard trade-off between localization properties in the frequency and real domains, an interesting consequence can be loosely suggested as follows: the better the localization in real domain, the worse the correlation properties. This is clearly a paradox, but it does shed some light within the class of needlets - in particular it clarifies that the uncorrelation property of wavelets coefficients does not follow at all by their localization properties in real domain: given the fixed-domain asymptotics we are considering, perfect local-

ization in real space does not ensure any form of uncorrelation (all random values at different locations in the sphere have in general a non-zero correlation).

The plan of this Chapter is as follows: in first Section we shall review the Mexican needlets construction. In the second Section we provide our main results on the dependence structure in real space, providing necessary and sufficient conditions for the uncorrelation properties to hold. In the third Section we focus on uncorrelation properties at different frequencies and in the fourth Section we review some statistical applications.

4.1 Spherical Mexican Needlets

The construction in [34] is in a sense similar to NPW needlets in [66, 67], insofar as a convolution of Legendre polynomials with a smooth function is proposed; the main difference is that for NPW needlets the kernel of the convolution is taken to be compactly supported, which allows on one hand for an exact reconstruction function (needlets make up a tight frame), at the same time granting exact localization in the frequency-domain. It should be added, however, that the approach by [34] enjoys some undeniable strong points: firstly, it covers general oriented manifolds and not simply the sphere; moreover it yields Gaussian localization properties in the real domain. A further nice benefit is that it can be formulated in terms of an explicit recipe in real space, a feature which is certainly valuable for practitioners. In particular, as we report below in the high-frequency limit the Mexican needlets are asymptotically

close to the Spherical Mexican Hat Wavelets, which have been exploited in several Cosmological papers but still lack a sound stochastic investigation.

More precisely, [34] propose to replace $b(l/B^j)$ in (2.20) by $f(l(l+1)/B^{2j})$, where $f(.)$ is some C^∞ function (not necessarily of bounded support) and the sequence $\{l(l+1)\}_{l=1,2,...}$ represents the eigenvalues of the Laplacian operator Δ_{S^2}. In particular, Mexican needlets can be obtained by taking $f(s) = s\exp(-s)$, so to obtain

$$\psi_{jk}^{M}(x) = \sqrt{\lambda_{jk}^{M}} \sum_{l\geq 1} \frac{l(l+1)}{B^{2j}} e^{-\frac{l(l+1)}{B^{2j}}} L_l\left(\left\langle x, \xi_{jk}\right\rangle\right) .$$

More generally, it is possible to consider higher order Mexican needlets by focussing on $f(s) = s^p \exp(-s)$, so to obtain

$$\psi_{jk;p}^{M}(x) := \sqrt{\lambda_{jk}^{M}} \sum_{l\geq 1} (\frac{l(l+1)}{B^j})^p e^{-l(l+1)/B^{2j}} L_l\left(\left\langle x, \xi_{jk}\right\rangle\right) .$$

Throughout this chaper, we shall only consider weight functions of the form $f(s) = s^p \exp(-s)$. It is certainly possible to consider more general constructions; however this specific shape lends itself to very neat results, allowing us to produce both upper and lower bounds for the coefficients' correlation. Also, it makes possible a clear interpretation of the final results, i.e. the effect of varying p on the structure of depedence is immediately understood; this is, we believe, a valuable asset for practitioners.

The resulting functions make up a frame which is not tight, but very close to, in a sense which is made rigorous in [34]. More precisely, we have explained in Section 2.6.4 that in order to make the needlets system to be a tight frame, the sequence

$b(l/B^j)$ must satisfy that for every l,

$$\sum_{j=-\infty}^{\infty} b^2\left(\frac{l}{B^j}\right) = 1.$$

and replacing $b(l/B^j)$ in the standard needlets construction with $\left(\frac{l(l+1)}{B^{2j}}\right)^p e^{-\frac{l(l+1)}{B^{2j}}}$, by simple manipulation we obtain that

$$\begin{aligned}
\sum_{j=-\infty}^{\infty} b^2\left(\frac{l}{B^j}\right) &= \sum_{j=-\infty}^{\infty}\left(\frac{l\,(l+1)}{B^{2j}}\right)^{2p} e^{-2l(l+1)/B^{2j}} \\
&= \sum_{j=-\infty}^{\infty}\left(\frac{l\,(l+1)}{B^{2j}}\right)^{2p} e^{-2l(l+1)/B^{2j}} \frac{l\,(l+1)\left(B^{-2j}-B^{-2(j+1)}\right)}{l\,(l+1)/B^{2j}\,(1-B^{-2})} \\
&\approx \frac{1}{(1-B^{-2})}\int_0^{+\infty} t^{2p-1}e^{-2t}dt \\
&= \frac{(2p-1)!}{(1-B^{-2})\,2^{2p}}
\end{aligned}$$

It can then be concluded that the mexican needlets frame is actually a nearly tight frame. The difference between the frame constants is of order $\left(\frac{A}{B}-1\right) \triangleq 10^{-4}$ for the first few values of p (see [20, 34]). On the other hand, exact cubature formulae cannot hold (in particular, $\left\{\lambda_{jk}^M\right\}$ are not exactly cubature weights in this case), because polynomials of infinitely large order are involved in the construction, but again this entails very minor approximations in practical terms. The random spherical Mexican-hat needlet coefficients are immediately seen to be given by

$$\begin{aligned}
\beta_{jk;p}^M &= \int_{S^2} T(x)\,\psi_{jk}(x)\,dx = \int_{S^2}\sum_{l\geq 0}\sum_{m=-l}^{l} a_{lm}Y_{lm}(x)\,\psi_{jk}(x)\,dx \\
&= \sqrt{\lambda_{jk}^M}\sum_{l\geq 1}\left(\frac{l(l+1)}{B^j}\right)^p e^{-l(l+1)/B^{2j}}\sum_{m=-l}^{l} a_{lm}Y_{lm}\left(\xi_{jk}\right)\ ,
\end{aligned}$$

whence their covariance is

$$E\beta^M_{jk;p}\beta^M_{jk';p} = \sqrt{\lambda^M_{jk}\lambda^M_{jk'}}\sum_{l\geq 1}(\frac{l(l+1)}{B^j})^{2p}e^{-2l(l+1)/B^{2j}}\frac{2l+1}{4\pi}C_l P_l\left(\left\langle \xi_{jk},\xi_{jk'}\right\rangle\right).$$

In the sequel, we shall drop the superscript M, without risk of confusion.

Remark As mentioned earlier, it can be suggested from results in ([34]) that Mexican needlets provide asymptotically a very good approximation to the widely popular Spherical Mexican Hat Wavelets (SMHW), which have been used in many physical papers; the asymptotic analysis of the stochastic properties of SMHW coefficients is still completely open for research. The discretized form of the SMHW can be written as

$$\Psi_{jk}(\theta;B^{-j}) = \frac{1}{(2\pi)^{\frac{1}{2}}\sqrt{2}B^{-j}(1+B^{-2j}+B^{-4j})^{\frac{1}{2}}}[1+(\frac{y}{2})^2]^2[2-\frac{y^2}{2t^2}]e^{-y^2/4B^{-j2}},$$

where the coordinates $y = 2\tan\frac{\theta}{2}$ follows from the stereographic projection on the tangent plane in each point of the sphere; here we take $\theta = \theta_{jk}(x) := d(x,\xi_{jk})$. Now write

$$\psi^M_{jk}\left(\theta_{jk}(x)\right) = \psi^M_{jk}(\theta)\ ;$$

by following the arguments in [34] and developing their bounds further, it is possible to show that

$$\left|\Psi_{jk}(\theta;B^{-j}) - K\psi^M_{jk}(\theta)\right| = B^{-j}O\left(\min\left\{\theta^4B^{4j},1\right\}\right)\ , \tag{4.1}$$

for some suitable normalization constant $K > 0$. Equation (4.1) suggests that our results below can be used as a guidance for the asymptotic theory of random SMHW coefficients.

4.2 Stochastic Properties of Mexican Needlet Coefficients

As mentioned in the Introduction, having established (2.24) opened the way to several developments for the statistical analysis of spherical random fields. It is therefore a very important question to establish under what circumstances these results can be extended to other constructions, such as Mexican needlets. In this Section, we provide a full characterization with positive and negative results. We start from the former, as follows:[3]

Theorem 9 *Assume Condition A holds with $\alpha < 4p+2$ and $M \geq 4p+2-\alpha$; then for some constant $C > 0$*

$$\left|Corr\left(\beta_{jk;p}, \beta_{jk';p}\right)\right| \leq \frac{C}{\left(1 + B^j d\left(\xi_{jk}, \xi_{jk'}\right)\right)^{(4p+2-\alpha)}}, \tag{4.2}$$

if $\alpha \in \mathbb{R}^+/\mathbb{N}$ (i.e. if it is non-integer), and

$$\left|Corr\left(\beta_{jk;p}, \beta_{jk';p}\right)\right| \leq \frac{C}{\left(1 + (B^{j-\log_B j}) d\left(\xi_{jk}, \xi_{jk'}\right)\right)^{(4p+2-\alpha)}}, \tag{4.3}$$

if $\alpha \in \mathbb{N}$.

[3] While finishing the paper [54], we learned by personal communication that working independently and at the same time as us, A.Mayeli has obtained a result similar to Theorem 9, see [63]. The statement and the assumptions in the two approaches are not equivalent and the methods of proofs are entirely different; we believe both are of independent interest.

Proof. We prove (4.2) following some ideas in [66]. We focus on

$$Corr\left(\beta_{jk;p},\beta_{jk';p}\right)=\frac{\sum\limits_{l\geq 1}(\frac{l(l+1)}{B^{2j}})^{2p}e^{-2l^2/B^{2j}}\frac{2l+1}{4\pi}C_lP_l\left(\left\langle\xi_{jk},\xi_{jk'}\right\rangle\right)}{\sum\limits_{l\geq 1}(\frac{l(l+1)}{B^{2j}})^{2p}e^{-2l^2/B^{2j}}\frac{2l+1}{4\pi}C_l}. \tag{4.4}$$

Now replace $C_l=l^{-\alpha}g_j\left(\frac{l}{B^j}\right)$ in the denominator of the above representation, for which we get

$$c_0^{-1}\frac{B^{-(4p+2-\alpha)j}}{4p+2-\alpha}\leq\sum_{l\geq 1}(\frac{l\left(l+1\right)}{B^{2j}})^{2p}e^{-2l^2/B^{2j}}\frac{2l+1}{4\pi}C_l\leq c_0\frac{B^{-(4p+2-\alpha)j}}{4p+2-\alpha}\ .$$

Denoting $\theta=\arccos\left\langle\xi_{j,k},\xi_{j,k'}\right\rangle$, the numerator can be written as

$$\sum_{l\geq 1}(\frac{l\left(l+1\right)}{B^{2j}})^{2p}e^{-2l^2/B^{2j}}\frac{2l+1}{4\pi}C_l\frac{1}{\pi}\int_\theta^\pi\frac{\sin\left(l+\frac{1}{2}\right)\varphi}{\left(\cos\theta-\cos\varphi\right)^{1/2}}d\varphi \tag{4.5}$$

where we have used the Dirichlet-Mehler integral representation for the Legendre polynomials [38].

The following steps and notations are very close to [66]. We write

$$\begin{aligned}C_{B,g_j}&=\sum_{l\geq 1}(\frac{l\left(l+1\right)}{B^{2j}})^{2p}e^{-2l^2/B^{2j}}\frac{2l+1}{4\pi}l^{-\alpha}g_j\left(\frac{l}{B^j}\right)\sin\left(l+\frac{1}{2}\right)\varphi \qquad (4.6)\\ &:=\frac{1}{2i}\sum_{l\geq 1}\left(h_{j+}(l)-h_{j-}\left(l\right)\right)\end{aligned}$$

where

$$h_{j\pm}(u)=(\frac{u\left(u+1\right)}{B^{2j}})^{2p}\frac{2u+1}{4\pi}u^{-\alpha}g_j\left(\frac{u}{B^j}\right)e^{-2(u/B^j)^2\pm i(u+\frac{1}{2})\varphi}.$$

By Poisson summation formula, we have

$$\sum_{l\geq 1}h_{j\pm}(l)=\frac{1}{2}\sum_{l\in\mathbb{Z}}h_{j\pm}(l)=\frac{1}{2}\sum_{\mu\in\mathbb{Z}}\widehat{h}_{j\pm}\left(2\pi\mu\right)\ .$$

Denote

$$G_{\alpha,j}\left(t\right):=t^{2p-\alpha}g_j\left(t\right)e^{-2t^2}I_{(B^{-j},\infty)}\ . \tag{4.7}$$

Let us now recall the following standard property of Fourier transforms:

$$(i\omega)^k \frac{d^m}{d\omega^m}\widehat{f}(\omega) = (\frac{d^k}{dx^k}\{x^m f(x)\})\widehat{}(\omega) .$$

Some simple computations yield

$$\begin{aligned}\widehat{h}_{\pm}(2\pi\mu) &= \frac{B^{(1-\alpha)j}}{4\pi}\sum_{m=0}^{2p}\binom{2p}{m}\left(2B^{-(2p-m)j}\frac{d^{m+1}}{d\omega^{m+1}} + B^{-(2p+1-m)j}\frac{d^m}{d\omega^m}\right) \\ &\quad \times 2\int_1^{\infty} G_{\alpha,j}\left(t/B^j\right) e^{\pm i(t+\frac{1}{2})\varphi - it\omega} dt\,|_{\omega=2\pi\mu} \\ &= \frac{B^{(2-\alpha)j}e^{\pm i\varphi}}{4\pi}\sum_{m=0}^{2p}\binom{2p}{m}\left(2B^{-(2p-m)j}\frac{d^{m+1}}{d\omega^{m+1}} + B^{-(2p+1-m)j}\frac{d^m}{d\omega^m}\right)\widehat{G}_{\alpha,j}(\omega)\,|_{\omega=B^j(2\pi\mu\mp\varphi)} ,\end{aligned}$$

where

$$\widehat{G}_{\alpha,j}(\omega) = \int_{\mathbb{R}} G_{\alpha,j}(t)\, e^{-it\omega} dt .$$

For all positive integers $k \leq M$, we can obtain

$$\left|\int_{B^{-j}}^{\infty}\frac{d^k}{dt^k}\{t^m G_{\alpha,j}(t)\}\,dt\right| \leq \begin{cases} \frac{k\Gamma(m+2p-\alpha)C_g}{L(p,m,\alpha,k)}B^{-j(2p+1+m-\alpha-k)}, \text{ for } 2p+1-\alpha+m \neq k \\ k\Gamma(m+2p-\alpha)\,C_g(j\log B) , \text{ for } 2p+1-\alpha+m = k \end{cases} \tag{4.8}$$

where $C_g = \max\{c_0, ..., c_M\}$ and

$$L(p,m,\alpha,k) := \begin{cases} \Gamma(2p+1+m-\alpha-k) \text{ when } (2p+1+m-\alpha-k) > 0 , \\ (\Gamma(k-2p-1-m+\alpha))^{-1} \text{ when } (2p+1+m-\alpha-k) < 0 . \end{cases} .$$

It should be noticed that our argument here differs from the one in [66], because we cannot assume the integrand function on the left-hand side to be in L^1 for all $k \leq M$.

Let us now focus on the case where $2p+1-\alpha+m \neq k$; the modifications needed for $2p+1-\alpha+m = k$ are obvious. We obtain

$$\begin{aligned} B^{-(2p+1-m)j}\left|\frac{d^m}{d\omega^m}\widehat{G}_{\alpha,j}(\omega)\right|\left|B^j(2\pi\mu-\varphi)\right|^k &\leq B^{-(2p+1-m)j}\left|\int_{B^{-j}}^{\infty}\frac{d^k}{dt^k}\{t^m G_j(t)\}\,dt\right| \\ &\leq \frac{k\Gamma(m+2p-\alpha)\,C_g}{L(p,m,\alpha,k)}B^{-j(4p+2-\alpha-k)}; \end{aligned}$$

therefore

$$B^{-(2p+1-m)j}\left|\frac{d^m}{d\omega^m}\widehat{G}_{\alpha,j}\left(\omega\right)\right| \leq \frac{C_{m,\alpha,k,g}B^{-j(4p+2-\alpha)}}{\left(2\pi\mu-\varphi\right)^k},$$

where

$$C_{m,\alpha,k,g} = \frac{k\left(m+2p-\alpha\right)!C_g}{L(p,m,\alpha,k)}.$$

Now let $C_{\alpha,k,g} = \max\left\{C_{m,g,\alpha,k},\ m=0,...,2p+1\right\}$; we have

$$\left|\widehat{h}_{\pm}\left(2\pi\mu\right)\right|$$

$$\begin{aligned}
&\leq \frac{B^{(2-\alpha)j}}{4\pi}\sum_{m=0}^{2p}\binom{2p}{m}\{2B^{-(2p-m)j}\left|\frac{d^{m+1}}{d\omega^{m+1}}\widehat{G}_{\alpha,j}\left(\omega\right)\right| \\
&\quad +B^{-(2p+1-m)j}\left|\frac{d^m}{d\omega^m}\widehat{G}_{\alpha,j}\left(\omega\right)\right|\}\,|_{\omega=B^j(2\pi\mu\mp\varphi)} \\
&\leq \frac{B^{(2-\alpha)j}}{4\pi}\sum_{m=0}^{2p}\binom{2p}{m}\frac{C_{\alpha,k,g}B^{-j(4p+2-\alpha)}}{\left(2\pi\mu-\varphi\right)^k} \\
&\leq \frac{2^{2p}C_{\alpha,k,g}B^{-4pj}}{\left(2\pi\mu\mp\varphi\right)^k},\ \mu=1,2,...
\end{aligned}$$

Therefore

$$\begin{aligned}
\left|C_{B,g_j}\right| &\leq 2^{2p}C_{\alpha,k,g}\left(\frac{1}{2\varphi^k}+\sum_{\mu\in N}\frac{1}{\left|2\pi\mu\pm\varphi\right|^k}\right)B^{-4pj} \\
&\leq 2^{2p}\left(\frac{1}{2\varphi^k}+\left(k-1\right)\pi^{1-k}\right)C_{\alpha,k,g}B^{-4pj}
\end{aligned}$$

Up to now, we can have for the numerator of the correlation

$$(4.5) \leq C_{\alpha,k,g}B^{-4pj}\int_{\theta}^{\pi}\frac{\left(\frac{1}{2\varphi^k}+\left(k-1\right)\pi^{1-k}\right)}{\left(\cos\theta-\cos\varphi\right)^{1/2}}d\varphi\ .$$

As in [66], we recall that $\cos\theta - \cos\varphi = 2\sin\frac{\theta+\varphi}{2}\sin\frac{\varphi-\theta}{2}$. For $0 \leq \theta \leq \pi/2$, using the fact that $\sin t/t$ is decreasing for $0 \leq t \leq \pi$, we obtain

$$\frac{\sin\frac{\theta+\varphi}{2}}{\frac{\theta+\varphi}{2}}\frac{\sin\frac{\varphi-\theta}{2}}{\frac{\varphi-\theta}{2}} \geq \frac{\sin\frac{3}{4}\pi}{\frac{3}{4}\pi}\frac{\sin\frac{1}{4}\pi}{\frac{1}{4}\pi} \geq 0.3 \times 0.9 \ ,$$

Therefore, we can get the following inequality

$$(4.5) \leq C_{\alpha,k,g}B^{-4pj}\int_{\theta}^{\pi}\frac{\pi^k}{0.27\varphi^k\left(\varphi^2-\theta^2\right)^{1/2}}d\varphi \leq C_1 B^{-4pj}\theta^{-k} \ .$$

If $\pi/2 \leq \theta \leq \pi$, letting $\widetilde{\theta} = \pi - \theta, \widetilde{\varphi} = \pi - \varphi$, we can obtain the same bound. Going back to 4.4, we obtain

$$Corr\left(\beta_{j,k},\beta_{j,k'}\right) \leq \frac{C_1\theta^{-k}B^{-4pj}}{C_2 B^{(2-\alpha)j}} = C\theta^{-k}B^{-(4p+2-\alpha)j} \to 0, \ as \ j \to \infty \ ,$$

where $C_1 = 2\sqrt{2}\pi^k\Gamma^2\left(1/2\right)C_{\alpha,k,g}$ and $C_2 = \left\{\pi\left(4p+2-\alpha\right)c_0\right\}^{-1}$.

Taking $k = 4p + 2 - \alpha$, we thus get inequality (4.2) and (4.3). ■

The previous result shows that Mexican needlets can enjoy the same uncorrelation properties as standard needlets, in the circumstances where the angular power spectrum is decaying "slowly enough". The extra log term in (4.3) where α is integer is a consequence of a standard technical difficulty when dealing with a boundary case in the integral in (4.8); however, from the point of view of Physical applications this does not seem to entail any practical drawback, since the case of an integer α is a measure zero set of the parameter space.

The following Theorem complete the analysis, establishing indeed that the random Mexican needlets coefficients are necessarily correlated at some angular dis-

tance in the presence of faster memory decay. This is clearly different from needlets, which are always uncorrelated. The heuristic rationale behind this duality can be explained as follows: it should be stressed that we are focussing on high-resolution asymptotics, i.e. the asymptotic behaviour of random coefficients at smaller and smaller scales in the same random realization. For such asymptotics, a crucial role can be played by terms which remain constant across different scales. In the case of usual needlets, which have bounded support over the multipoles, terms like these are simply dropped by construction. This is not so for Mexican needlets, which in any case include components at the lowest scales. These components are dominant when the angular power spectrum decays fast, and as such they prevent the possibility of asymptotic uncorrelation. In particular, we have correlation when the angular power spectrum is such that $\alpha > 4p + 2$.

Theorem 10 *Under condition A, for* $\alpha > 4p + 2$, $\forall\, \varepsilon \in (0, 1)$, *there exists positive* $\delta \leq \varepsilon \left(1 + c_0^2\right)^{-1/(\alpha - 4p - 2)}$, $s.t.$ *if* $d(\xi_{jk}, \xi_{jk'}) \leq \delta$, *then*

$$\lim_{j \to \infty} Corr\left(\beta_{jk;p}, \beta_{jk';p}\right) > 1 - \varepsilon\,.$$

Proof. For notational simplicity, we perform the proof for $p = 1$; the general case can be established under the same lines.

We first divide the variance of the coefficients into three parts, as follows

$$\left(\sum_{1 \leq l < \epsilon_1 B^j} + \sum_{\epsilon_1 B^j \leq l < \epsilon_2 B^j} + \sum_{l \geq \epsilon_2 B^j}\right) \frac{l^2 \left(l + 1\right)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l + 1}{4\pi} C_l\,.$$

For the first part we obtain easily

$$\sum_{1\leq l<\epsilon_1 B^j} \frac{l^2\left(l+1\right)^2}{B^{4j}} e^{-2l(l-1)/B^{2j}} \frac{2l+1}{4\pi} C_l \leq 2\sum_{1\leq l<\epsilon_1 B^j} \frac{l^4}{B^{4j}} \frac{l}{\pi} l^{-\alpha} g_j(\frac{l}{B^j})$$
$$\leq \quad 2\frac{c_0}{\pi} B^{(2-\alpha)j} \int_{B^{-j}}^{\epsilon_1} x^{5-\alpha} dx = 2\frac{c_0\left(B^{(\alpha-6)j} - \epsilon_1^{6-\alpha}\right)}{\pi\left(\alpha-6\right)} B^{(2-\alpha)j},$$

and

$$\sum_{1\leq l<\epsilon_1 B^j} \frac{l^2\left(l+1\right)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l \geq \frac{\left(B^{(\alpha-6)j} - \epsilon_1^{6-\alpha}\right)}{2\pi c_0\left(\alpha-6\right)} e^{-2\epsilon_1^2} B^{(2-\alpha)j},$$

Similarly, for the second part

$$\frac{\left(\epsilon_2^{6-\alpha} - \epsilon_1^{6-\alpha}\right)}{2\pi\left(6-\alpha\right)c_0} e^{-2\epsilon_2^2} B^{(2-\alpha)j} \quad \leq \quad \sum_{\epsilon_1 B^j \leq l<\epsilon_2 B^j} \frac{l^2\left(l+1\right)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l$$
$$\leq \quad 2\frac{\left(\epsilon_1^{6-\alpha} - \epsilon_2^{6-\alpha}\right)}{\pi\left(\alpha-6\right)} c_0 e^{-2\epsilon_1^2} B^{(2-\alpha)j};$$

for the third part

$$\sum_{l\geq\epsilon_2 B^j} \frac{l^2\left(l+1\right)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l \leq \frac{3c_0}{4\pi} B^{(\alpha+2)j} \int_{\epsilon_2}^{\infty} x^{5-\alpha} e^{-x^2} dx$$
$$\leq \quad \frac{3c_0}{4\pi} \epsilon_2^{4-\alpha} e^{-\epsilon_2^2} B^{(2-\alpha)j}, \ \left(\epsilon_2 > 1\right) .$$

The last but one inequality follows from the asymptotic formula

$$\frac{1}{\sqrt{2\pi}} \int_y^{\infty} e^{-x^2/2} dx \sim \frac{1}{\sqrt{2\pi} y} e^{-y^2/2}, \ \ y \to \infty .$$

If we choose $\epsilon_1 = NB^{-j}$, where N is sufficiently large that $N^{6-\alpha}\left(1+c_0^2\right) < 1$, then

$$\frac{\left(B^{(\alpha-6)j} - \epsilon_1^{6-\alpha}\right)}{c_0} > \left(\epsilon_1^{6-\alpha} - \epsilon_2^{6-\alpha}\right) c_0, \text{ as } j \to \infty .$$

Therefore

$$\left\{ \sum_{1\leq l<\epsilon_1 B^j} \frac{l^2\,(l+1)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l \right\}^{-1}$$
$$\times \left\{ \left(\sum_{\epsilon_1 B^j \leq l<\epsilon_2 B^j} + \sum_{l\geq \epsilon_2 B^j} \right) \frac{l^2\,(l+1)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l \right\}$$
$$= O(B^{j(6-\alpha)}) = o(1)\ ,\ as\ j \to \infty\ .$$

It follows that the order of the whole sum is the same as for the summation restricted to the elements in the range $1 \leq l < \epsilon_1 B^j$. We recall also that

$$\sup_{\theta\in[0,\pi]} P_l(\cos\theta) = P_l(\cos 0) = 1\ ,$$

and

$$\sup_{\theta\in[0,\pi]} |\frac{d}{d\theta} P_l(\cos\theta)| \leq 3l$$

(See Appendix C for detailed proof of the last inequality). As a consequence, for any positive $\varepsilon < 1$, there exists a $\delta > 0$, s.t. if $0 < \theta \leq \delta \leq \varepsilon/N$, then,

$$|P_l(\cos\theta) - P_l(\cos 0)| \leq 3l\theta \leq \varepsilon$$

for any $l > N$, where the above inequalities follow from

$$0 \leq \cos(\theta_0+\theta) - \cos\theta_0 = 2\sin^2\frac{\theta}{2} \leq \theta\ ,\ \forall\ \theta_0 \in [0, 2\pi)\ .$$

Therefore, for any $\xi_{jk}, \xi_{jk'} \in \mathbb{S}^2$ s.t. $\arccos\left\langle \xi_{jk}, \xi_{jk'} \right\rangle \leq \delta$, we have

$$Corr\left(\beta_{jk}, \beta_{jk'}\right) = \frac{\sum_{l\geq 1} \frac{l^2(l+1)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l P_l\left(\left\langle \xi_{jk}, \xi_{jk'} \right\rangle\right)}{\sum_{l\geq 1} \frac{l^2(l+1)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l}$$

$$\geq \frac{\sum\limits_{1\leq l\leq N} \frac{l^2(l+1)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l \times (1-\varepsilon)}{\sum\limits_{l\geq 1} \frac{l^2(l+1)^2}{B^{4j}} e^{-2l(l+1)/B^{2j}} \frac{2l+1}{4\pi} C_l} + O(B^{j(6-\alpha)}) = (1-\varepsilon) + o(1)\,,\ as\ j\to\infty\,.$$

Thus, the proof of the theorem is completed. ■

The results in the previous two Theorems illustrate an interesting trade-off between the localization and correlation properties of spherical needlets. In particular, we can always achieve uncorrelation by choosing $p > (\alpha - 2)/4$; of course α is generally unknown and must be estimated from the data (in this sense standard needlets have better robustness properties). Introducing higher order terms implies lowering the weight of the lowest multipoles, i.e. improving the localization properties in frequency space; on the other hand, this may entail some worsening of the localization in pixel space, in view of the usual trade-off between real and harmonic space localization properties (i.e., the indetermination principle).

Remark In Theorem 10, we decided to keep the assumptions as close as possible to Theorem 9, in order to ease comparisons and highlight the symmetry between the two results. However, it is simple to show that the correlation result holds in much greater generality, for angular power spectra that have a decay which is faster than polynomial. In particular, assume that

$$C_l = H(l)\exp(-l^p)\,,\ l = 1, 2, ...$$

where $H(l)$ is any kind of polynomial such that $H(l) > c > 0$ and $p > 0$. Then it is simple to establish the same result as in Theorem 10, by means of a simplified version of the same argument. The underlying rationale should be easy to get: for exponentially decaying power spectra the dominating components are at the lowest frequencies, and they introduce correlations among all random coefficients which cannot be neglected.

4.3 Correlation Across Different Frequencies

The asymptotic uncorrelation we established in Theorem 9 is indeed sufficient to obtain central limit results for finite-dimensional statistics based on nonlinear transformations of the Mexican needlets coefficients, see below. However, for many applications it is useful to consider different scales $\{j\}$ at the same time. Because of this, it is also important to focus on the correlation of Mexican needlet coefficients at different j, j'. We stress that no such analysis was needed for standard needlets; given the compactly supported kernel $b(.)$, the frequency support of the various coefficients is automatically disjoint when $|j - j'| \geq 2$. We start by writing the expression for the correlation coefficients, which is given by

$$Corr\left(\beta_{j_1k_1}, \beta_{j_2k_2}\right)$$

$$= \frac{\sum_{l\geq 1}\left(\frac{l(l+1)}{B^{j_1}}\right)^p\left(\frac{l(l+1)}{B^{j_2}}\right)^p e^{-l(l+1)(B^{2j_1}+B^{2j_2})^{-1}}(2l+1)C_l P_l^{\left(\frac{1}{2}\right)}\left(\left\langle \xi_{j_1k_1}, \xi_{j_2k_2}\right\rangle\right)}{\left\{\sum_{l\geq 1}\left(\frac{l(l+1)}{B^{j_1}}\right)^{4p} e^{-2l(l+1)/B^{2j_1}}(2l+1)C_l\right\}^{1/2}\left\{\sum_{l\geq 1}\left(\frac{l(l+1)}{B^{j_2}}\right)^{4p} e^{-2l(l+1)/B^{2j_2}}(2l+1)C_l\right\}^{1/2}}.$$

We have the following result.

Theorem 11 *Assume Condition A holds with* $\alpha < 4p+2$ *and* $M \geq 4p+2-\alpha$; *then*

$$\left|Corr\left(\beta_{j_1k_1;p}, \beta_{j_2k_2;p}\right)\right| \leq \frac{C_M}{\left(1+B^{(j_1+j_2)/2}d\left(\xi_{j_1k_1}, \xi_{j_2k_2}\right)\right)^{(4p+2-\alpha)}}, \tag{4.9}$$

if $\alpha \in \mathbb{R}^+/\mathbb{N}$ *(i.e. if it is non-integer), and*

$$\left|Corr\left(\beta_{j_1k_1;p}, \beta_{j_2k_2;p}\right)\right| \leq \frac{C_M}{\left(1+(B^{(j_1+j_2)/2-\log_B(j_1+j_2)/2}d\left(\xi_{j_1k_1}, \xi_{j_2k_2}\right)\right)^{(4p+2-\alpha)}}, \tag{4.10}$$

if $\alpha \in \mathbb{N}$, *for some constant* $C_M > 0$. *On the other hand, for* $\alpha > 4p+2$, $\forall\, \varepsilon \in (0,1)$, *there exists positive* $\delta \leq \varepsilon\left(1+c_0^2\right)^{-1/(\alpha-4p-2)}$, $s.t.$ *if* $d(\xi_{j_1k_1}, \xi_{j_2k_2}) \leq \delta$, *then*

$$\lim_{j\to\infty} Corr\left(\beta_{j_1k_1;p}, \beta_{j_2k_2;p}\right) > 1-\varepsilon\,. \tag{4.11}$$

Proof. As the proof is very similar to the arguments for Theorems 9 and 10, we just provide the main ideas and we omit many details. In the sequel, for any two sequences a_l, b_l, we write $a_l \approx b_l$ if and only if $a_l = O(b_l)$ and $b_l = O(a_l)$.

First, we consider the variance of the random coefficients, which can be represented by:

$$\sum(\frac{l(l+1)}{B^{j_1}})^{4p}e^{-2l(l+1)/B^{2j_1}}(2l+1)C_l \approx B^{(2-\alpha)j_1}\left\{\left(\int_{B^{-j_1}}^1 + \int_1^\infty\right)t^{4p+1-\alpha}e^{-2t(t+1)}g_{j_1}(t)\,dt\right\}$$

$$= B^{(2-\alpha)j_1}\left\{\frac{c_0^{\pm 1}}{2p+2-\alpha}B^{-j_1(4p+2-\alpha)} + O(1)\right\}\,;$$

therefore

$$\begin{aligned}
& \left\{\sum_{l\geq 1}(\frac{l(l+1)}{B^{j_1}})^{4p}e^{-2l(l+1)/B^{2j_1}}(2l+1)C_l\right\}^{1/2}\left\{\sum_{l\geq 1}(\frac{l(l+1)}{B^{j_2}})^{4p}e^{-2l(l+1)/B^{2j_2}}(2l+1)C_l\right\}^{1/2} \\
= & B^{(1-\alpha/2)(j_1+j_2)}\left\{\frac{c_0^{\pm 1}}{4p+2-\alpha}B^{-j_1(4p+2-\alpha)}+O(1)\right\}^{1/2}\left\{\frac{c_0^{\pm 1}}{4p+2-\alpha}B^{j_2(4p+2-\alpha)}+O(1)\right\}^{1/2} \\
= & \begin{cases} O(1)B^{(1-\alpha/2)(j_1+j_2)}, & 4p+2-\alpha>0; \\ O(1)B^{-2p(j_1+j_2)}, & 4p+2-\alpha<0. \end{cases}
\end{aligned}$$

Now let us focus on the case where $4p+2-\alpha>0$; without loss of generality, we can always assume $j_1<j_2$. To establish (4.9), we can implement the same argument as in Theorem 9, provided we replace C_{B,g_j} in (4.6) by

$$\begin{aligned}
C_{B,g_{j_1},j_2} &= \sum_{l\geq 1}(\frac{l(l+1)}{B^{j_1+j_2}})^{2p}e^{-l(l+1)(B^{-2j_1}+B^{-2j_2})}\frac{2l+1}{4\pi}l^{-\alpha}g_{j_1}\left(\frac{l}{B^{j_1}}\right)\sin\left(l+\frac{1}{2}\right)\varphi \\
&: = \frac{1}{2i}\sum_{l\geq 1}\left(h_{j_1j_2+}(l)-h_{j_1j_2-}(l)\right)
\end{aligned}$$

where

$$h_{j_1j_2\pm}(u)=(\frac{u(u+1)}{B^{j_1+j_2}})^{2p}\frac{2u+1}{4\pi}u^{-\alpha}g_{j_1}\left(\frac{u}{B^{j_1}}\right)e^{-u(u+1)(B^{-2j_1}+B^{-2j_2})\pm i(u+\frac{1}{2})\varphi}.$$

As before, by Poisson summation formula, we have

$$\sum_{l\geq 1}h_{j_1j_2\pm}(l)=\frac{1}{2}\sum_{l\in\mathbb{Z}}h_{j_1j_2\pm}(l)=\frac{1}{2}\sum_{\mu\in\mathbb{Z}}\widehat{h}_{j_1j_2\pm}(2\pi\mu)\ .$$

Denote

$$G_{\alpha,j_1j_2}(t):=B^{2p(j_1-j_2)}t^{2p-\alpha}g_{j_1}(t)\,e^{-t\left(t+B^{-j_1}\right)\left(1+B^{2(j_1-j_2)}\right)}I_{(B^{-j},\infty)}\ ;$$

by the same argument and notation as in Theorem 9, we have

$$\begin{aligned}\left|\widehat{h}_{j_1 j_2 \pm}\left(2\pi\mu\right)\right| &= \left|\frac{B^{(1-\alpha)j_1}}{4\pi}\sum_{m=0}^{2p}\binom{2p}{m}\left(2B^{-(2p-m)j_1}\frac{d^{m+1}}{d\omega^{m+1}}+B^{-(2p+1-m)j_1}\frac{d^m}{d\omega^m}\right)\right. \\ &\quad \left.\times 2\int_1^\infty G_{\alpha,j_1 j_2}\left(t/B^{j_1}\right)e^{\pm i(t+\frac{1}{2})\varphi - it\omega}dt\,|_{\omega=2\pi\mu}\right| \\ &\leq \frac{B^{(2-\alpha)j_1}}{4\pi}\sum_{m=0}^{2p}\binom{2p}{m}\left(2B^{-(2p-m)j_1}\left|\frac{d^{m+1}}{d\omega^{m+1}}\widehat{G}_{\alpha,j_1 j_2}\left(\omega\right)\right|+B^{-(2p+1-m)j_1}\left|\frac{d^m}{d\omega^m}\widehat{G}_{\alpha,j_1 j_2}\left(\omega\right)\right|\right)|_{\omega=B^j(2\pi\mu} \\ &\leq \frac{B^{(2-\alpha)j_1}}{4\pi}\sum_{m=0}^{2p}\binom{2p}{m}\frac{C_{\alpha,k,g}B^{-j_1(2-\alpha)-2p(j_1+j_2)}}{\left(2\pi\mu-\varphi\right)^k} \\ &\leq \frac{2^{2p}C_{\alpha,k,g}B^{-2p(j_1+j_2)}}{\left(2\pi\mu\mp\varphi\right)^k},\ \mu=1,2,...\end{aligned}$$

Therefore

$$\begin{aligned}\left|C_{B,g_{j_1},j_2}\right| &\leq 2^{2p}C_{\alpha,k,g}\left(\frac{1}{2\varphi^k}+\sum_{\mu\in N}\frac{1}{\left|2\pi\mu\pm\varphi\right|^k}\right)B^{-2p(j_1+j_2)} \\ &\leq 2^{2p}\left(\frac{1}{2\varphi^k}+\left(k-1\right)\pi^{1-k}\right)C_{\alpha,k,g}B^{-2p(j_1+j_2)}.\end{aligned}$$

It is then straightforward to conclude as in the case where $j_1 = j_2$, to obtain

$$Corr\left(\beta_{j_1,k},\beta_{j_2,k'}\right)\leq\frac{C_2\theta^{-k}B^{-2p(j_1+j_2)}}{C_1B^{(1-\alpha/2)(j_1+j_2)}}=C\theta^{-k}B^{-(2p+1-\alpha/2)(j_1+j_2)}\to 0,\ \text{as } j_2\to\infty\ .$$

Thus (4.9) is established.

Let us now focus on the case where $4p+2-\alpha<0$. We can again fix two integers $1\leq N_1\leq N_2$ and evaluate

$$\begin{aligned}
&\sum_{N_1\leq l\leq N_2}(\frac{l(l+1)}{B^{j_1}})^p(\frac{l(l+1)}{B^{j_2}})^p e^{-l(l+1)(B^{2j_1}+B^{2j_2})^{-1}}(2l+1)C_l\\
=&\sum_{N_1\leq l\leq N_2}\frac{l^{2p}(l+1)^{2p}}{B^{2(j_1+j_2)p}}e^{-l(l+1)(B^{2j_1}+B^{2j_2})^{-1}}(2l+1)C_l\\
\approx& e^{-N_i^{4p}(B^{2j_1}+B^{2j_2})^{-1}}c_0^{\pm 1}B^{(j_1+j_2)(1-\alpha/2)}\int_{N_1/B^{(j_1+j_2)/2}}^{N_2/B^{(j_1+j_2)/2}}t^{4p+1-\alpha}dt\\
\approx& B^{(j_1+j_2)(1-\alpha/2)}\left\{\left(N_1^{4p+2-\alpha}-N_2^{4p+2-\alpha}\right)\frac{B^{-(j_1+j_2)(2p+1-\alpha/2)}}{\alpha-4p-2}\right\}\\
\approx& \left(N_1^{4p+2-\alpha}-N_2^{4p+2-\alpha}\right)B^{-(j_1+j_2)2p}/\left(\alpha-4p-2\right)\ .
\end{aligned}$$

The proof of (4.11) can then be concluded along the same lines as in Theorem 10. ■

4.4 Statistical Applications

The previous results lend themselves to several applications for the statistical analysis of spherical random fields, in particular with a view to CMB data analysis. Similarly to [7], let us consider polynomials functions of the normalized Mexican needlets coefficients, as follows

$$h_{u,N_j}:=\frac{1}{\sqrt{N_j}}\sum_{k=1}^{N_j}\sum_{q=1}^{Q}w_{uq}H_q(\widehat{\beta}_{jk;p})\ ,\ \widehat{\beta}_{jk;p}:=\frac{\beta_{jk;p}}{\sqrt{E\beta_{jk;p}^2}}\ ,\ u=1,...,U\ ,$$

where $H_q(.)$ denotes the q-th order Hermite polynomials (see [80]), N_j is the cardinality of coefficients corresponding to frequency j (where we take $\{\xi_{jk}\}$ to form a B^{-j}-mesh, see [8], so that $N_j\approx B^{2j}$), and $\{w_{uq}\}$ is a set of deterministic weights that must ensure these statistics are asymptotically non-degenerate, i.e.

Condition B There exist j_0 such that for all $j > j_0$

$$rank(\Omega_j) = U \ , \Omega_j := Eh_{N_j}h'_{N_j} \ , h_{N_j} := (h_{1,N_j}, ..., h_{U,N_j})' \ .$$

Condition B is a standard invertibility assumption which will ensure our statistics are asymptotically non-degenerate (for instance, it rules out multicollinearity). Several examples of relevant polynomials are given in [7]; for instance, given a theoretical model for the angular power spectrum $\{C_l\}$, a goodness-of-fit statistic might be based upon

$$\begin{aligned} \frac{1}{\sqrt{N_j}} \sum_{k=1}^{N_j} H_2(\widehat{\beta}_{jk;p}) &= \frac{1}{N_j} \sum_{k=1}^{N_j} (\widehat{\beta}^2_{jk;p} - 1) \\ &= \frac{1}{\sqrt{N_j}} \sum_{k=1}^{N_j} (\frac{\beta^2_{jk;p}}{\lambda_{jk} \sum\limits_{l \geq 1} b^2(\frac{l}{B^j}) \frac{2l+1}{4\pi} C_l} - 1) \ ; \end{aligned}$$

likewise, tests of Gaussianity could be implemented by focussing on the skewness and kurtosis of the wavelets coefficients (see for instance [13],[18, 19]), i.e. by focussing on

$$\begin{aligned} \frac{1}{\sqrt{N_j}} \sum_{k=1}^{N_j} \left\{ H_3(\widehat{\beta}_{jk;p}) + H_1(\widehat{\beta}_{jk;p}) \right\} &= \frac{1}{\sqrt{N_j}} \sum_{k=1}^{N_j} \widehat{\beta}^3_{jk;p} \text{ and} \\ \frac{1}{\sqrt{N_j}} \sum_{k=1}^{N_j} \left\{ H_4(\widehat{\beta}_{jk;p}) + 6H_2(\widehat{\beta}_{jk;p}) \right\} &= \frac{1}{\sqrt{N_j}} \sum_{k=1}^{N_j} \left\{ \widehat{\beta}^4_{jk;p} - 3 \right\} \ . \end{aligned}$$

The joint distribution for these statistics is provided by the following results:

Theorem 12 *Assume T is a Gaussian mean square continuous and isotropic random field; assume also that Conditions A and B are satisfied and choose $p > (\alpha + \delta)/4$, some $\delta > 0$. Then as $N_j \to \infty$*

$$\Omega_j^{-1/2} h_{N_j} \to_d N(0, I_U) \, .$$

Proof. The asymptotic behaviour of our polynomial statistics can be established by means of the method of moments. In particular, it is possible to exploit the diagram formula for higher order moments of Hermite polynomial, as explained for instance in [69],[80]. The details are the same as in [7], and thus they are omitted for brevity's sake. We only note that, in order to be able to use Lemma 6 in that reference, we need to ensure that

$$\left| Corr \left(\beta_{jk;p}, \beta_{jk';p} \right) \right| \leq \frac{C}{\left(1 + B^j d \left(\xi_{j,k}, \xi_{j,k'} \right) \right)^{2+\delta}} \text{ , some } C > 0 \, .$$

In view of (4.2), this motivates the tighter limit we need to impose on the value of p.

■

It may be noted that the covariance matrix Ω_j can itself be consistently estimated from the data at any level j, for instance by means of the bootstrap/subsampling techniques that are detailed in [8]. Again, the arguments of that paper lend themselves to straightforward extensions to the present circumstances, as they rely uniquely upon the covariance inequalities for the random wavelet coefficients. A much more challenging issue relate to the relaxation of the Gaussianity assumptions, which is still under investigation.

References

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

[2] **Antoine, J.-P. and Vandergheynst, P. (2007)** Wavelets on the Sphere and Other Conic Sections, *Journal of Fourier Analysis and its Applications, 13, 369-386*

[3] **Antoine, J-P., Murenzi, R., Vandergheynst, P., and Twareque Ali S. (2004)** Two-Dimensional Wavelets and their Relatives, *Cambridge University Press*

[4] **Babich, D., Creminelli, P., Zaldarriaga, M. (2004)** The Shape of non-Gaussianities, *Journal of Cosmology and Astroparticle Physics 8, 009*

[5] **Baldi, P., Marinucci, D.(2007)**, Some Characterizations of the Spherical Harmonics Coefficients for Isotropic Random Fields, *Statistics and Probability Letters 77, 490–496*

[6] **Baldi, P., Kerkyacharian, G., Marinucci, D. and Picard, D. (2008)** High Frequency Asymptotics for Wavelet-Based Tests for Gaussianity and Isotropy on the Torus, *Journal of Multivariate Analysis, Vol. 99, pp. 606-636*

[7] **Baldi, P., Kerkyacharian, G., Marinucci, D. and Picard, D. (2006)** Asymptotics for Spherical Needlets, *Annals of Statistics, in press,* arxiv:math/0606599

[8] **Baldi, P., Kerkyacharian, G., Marinucci, D. and Picard, D. (2007)** Subsampling Needlet Coefficients on the Sphere, *Bernoulli, in press,* arXiv: 0706.4169

[9] **Baldi, P., Kerkyacharian, G., Marinucci, D. and Picard, D. (2008)** Adaptive Density Estimation for Directional Data Using Needlets, *Annals of Statistics, in press,* arXiv:0807.5059

[10] **Bartolo, N., Komatsu, E., Matarrese, S., Riotto, A. (2004)** Non-Gaussianity from Inflation: Theory and Observations, *Phys.Rept. 402, 103-266*

[11] **Billingsley, P.(1968)** *Convergence of Probability Measures.* Wiley, New York.

[12] **Bishop, R. L. and Goldberg, S. I. (1980)** *Tensor Analysis on Manifolds.* Dover Publications, Inc. New York.

[13] **Cabella, P., Hansen, F. K., Marinucci, D., Pagano, D. and Vittorio, N. (2004)** Search for non-Gaussianity in Pixel, Harmonic, and Wavelet Space: Compared and Combined. *Physical Review D **69** 063007.*

[14] **Cabella, P., Hansen, F. K., Liguori, M., Marinucci, D., Matarrese, S., Moscardini, L. and Vittorio, N. (2006)** The Integrated Bispectrum as a Test of CMB non-Gaussianity: Detection Power and Limits on f_{NL} with WMAP Data, *Mon.Not.Roy.Astron.Soc., Vol.369, 819-824*

[15] **Cabella, P. and Marinucci, D. (2008)** Statistical Challenges in the Analysis of Cosmic Microwave Background Radiation. *Annals of Applied Statistics, in press.*

[16] **Cayon, L., Sanz, J. L., Martinez-Gonzalez, E., Banday, A. J., Argueso, F., Gallegos, J.E., Gorski, K. M., Hinshaw, G. (2001)** Spherical Mexican Hat wavelet: an application to detect non-Gaussianity in the COBE-DMR maps, *Mon.Not.Roy.Astron.Soc., Vol.326, 1243*

[17] **Copi, C. J., Hunterer, D. and Starkman, G. D. (2004)** Multipole Vectors: A New Representation of the CMB Sky and Evidence for Statistical Anisotropy or non-Gaussianity at $2 \leq l \leq 8$. *Phys. Rev. D., 70(4) arxiv:astro-ph/0310511*

[18] **Cruz, M., Cayon, L., Martinez-Gonzalez, E., Vielva, P., Jin, J. (2007)** The non-Gaussian Cold Spot in the 3-year WMAP data, *Astrophysical Journal 655 , 11-20*

[19] **Cruz, M., Cayon, L., Martinez-Gonzalez, E., Vielva, P. (2006)** The non-Gaussian Cold Spot in WMAP: significance, morphology and foreground contribution, *Monthly Notices of the Royal Astronomical Society 369, 57-67*

[20] **Daubechies, I. (1992)** *Ten lectures on Wavelets,* Capital City Press, Montpelier, Vermont

[21] **de Gasperis, G., Balbi, A., Cabella, P., Natoli, P., Vittorio, N.(2005)** ROMA: A map-making algorithm for polarised CMB data sets, *Astronomy and Astrophysics, Volume 436, Issue 3, pp.1159-1165*

[22] **DeJong, P. (1990)** A Central Limit Theorem for Generalized Multilinear Forms. *Journal of Multivariate Analysis 34, 275-289.*

[23] **Delabrouille, J., Cardoso, J.-F,. Le Jeune, M., Betoule, M., Fay, G., Guilloux, F. (2008)** A Full Sky, Low Foreground, High Resolution CMB Map from WMAP, arXiv:0807.0773

[24] **Dodelson, S. (2003)** *Modern Cosmology*, Academic Press

[25] **Doukhan, P. (1988)** Formes de Toeplitz Associées à une Analyse Multiéchelle. (French) [Toeplitz Forms Associated to a Multiscale Analysis] *C. R. Acad. Sci. Paris Sér. I Math. 306, no. 15, 663–666*

[26] **Doukhan, P. and Leon, J. R. (1990)** Formes Quadratique d'estimateurs de Densité par Projections Orthogonales. (French) [Quadratic Deviation of Projection Density Estimates] *C. R. Acad. Sci. Paris Sér. I Math. 310, no. 6, 425–430*

[27] **Doroshkevich, A. G., Naselsky, P. D., Verkhodanov, O. V. , Novikov, D. I., Turchaninov, V. I., Novikov, I. D. , Christensen, P. R. , Chiang, L.-Y. (2005)** Gauss-Legendre Sky Pixelization (GLESP) for CMB Maps, *International Journal of Modern Physics D 14, 275*

[28] **Eriksen, H. K., Hansen, F. K., Banday, A. J., Gorski, K. M., Lilje, P. B. (2004)** Asymmetries in the CMB Anisotropy Field, *Astrophys.J. 605 14-20; Erratum-ibid. 609 1198*

[29] **Faraut, J. (2008)** *Analysis of Lie Groups,* Cambridge University Press, Cambridge

[30] **Faÿ, G., Guilloux, F. (2008)** Consistency of a Needlet Spectral Estimator on the Sphere, arXiv:0807.2162

[31] **Faÿ, G., Guilloux, F., Betoule, M., Cardoso, J.-F., Delabrouille, J., Le Jeune, M. (2008)** CMB Power Spectrum Estimation Using Wavelets, *Phys.Rev.D78:083013*

[32] **Flandrin, P. (1998)** Time-Frequency/Time-Scale Analysis, *San Diego, CA: Academic Press*

[33] **Guilloux, F., Faÿ, G., Cardoso, J.-F. (2008)** Practical Wavelet Design on the Sphere, *Applied and computational harmonic analysis, in press,* arxiv 0706.2598

[34] **Geller, D. and Mayeli, A. (2007)** Nearly Tight Frames and Space-Frequency Analysis on Compact Manifolds, arxiv:0706.3642v2

[35] **Geller, D. and Mayeli, A. (2007)** Besov Spaces and Frames on Compact Manifolds, arXiv:0709.2452

[36] **Genovese, C., Miller, C. J., Nichol, R. C., Arjunwadkar, M. and Wasserman, L. (2004)** Nonparametric Inference for the Cosmic Microwave Background, *Statistical Science, Vo.19, No.2, 308-321*

[37] **Genovese, C., Perone-Pacifico, M., Verdinelli, I. and Wasserman, L. (2008)** On the Path Density of a Gradient Field, *The Annals of Statistics, to appear,* arXiv:0805.4141

[38] **Gradshteyn, I. S. and Ryzhik, I. M. (1980)** *Table of Integrals, Series, and Products*, Academic Press

[39] **Gorski, K. M. , Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K. , Reinecke, M., Bartelman, M., (2005)** HEALPix – a Framework for High Resolution Discretization, and Fast Analysis of Data Distributed on the Sphere, *Astrophysical Journal 622, 759-771*

[40] **Hannan, E. J. (1970)** Multiple Time Series, *Wiley. New York*

[41] **Hansen, F. K., Cabella, P., Marinucci, D., Vittorio, N., (2004)** Asymmetries in the Local Curvature of the WMAP Data, *Astrophys.J. 607 (2004) L67-L70*

[42] **Hardle, W., Kerkyacharian, G., Picard, D., Tsybakov, A. (1997)** Wavelets, Approximation, and Statistical Applications, *John Wiley & Sons*

[43] **Hernandez, E. and Weiss G. L. (1996)** *A First Course on Wavelets*, Boca Raton, CRC Press

[44] **Holschneider, M. and Iglewska-Nowak, I. (2007)** Poisson Wavelets on the Sphere, *Journal of Fourier Analysis and its Applications, 13, 405-420*

[45] **Hu, W. (2001)** The angular Trispectrum of the CMB, *Physical Review D, Vol.64, Issue 8, id.3005*

[46] **Jones, M. N. (1985)** Spherical Harmonics and Tensors for Classical Field Theory, *Research Studies Press LTD. and John Wiley & Sons Inc.*.

[47] **Kagan, A.M., Linnik, Y.V., and Rao, C.R., (1973)** *Characterization problems in mathematical statistics, John Wiley & Sons*

[48] **Keihanen, E., Kurki-Suonio, H., Poutanen, T. (2005)** MADAM- a Map-making Method for CMB Experiments, *Monthly Notices of the Royal Astronomical Society, Volume 360, Issue 1, pp. 390-400*

[49] **Kerkyacharian, G., Petrushev, P., Picard, D., Willer, T. (2007)** Needlet Algorithms for Estimation in Inverse Problems, *Electronic Journal of Statistics*, *1, 30-76*

[50] **Klauder, J. R. and Skagerstam,B. (1985)** *Coherent States,* World Scientific Publishing Co. Pte. Ltd., Singapore

[51] **Kogo, N. and Komatsu, E. (2006)** Angular Trispectrum of CMB Temperature Anisotropy from Primodial Non-Gaussianity with the Full Radiation Transfer Function, *Physical Review D 73:8*

[52] **Komatsu, E. and Spergel, D.N. (2001)** Acoustic Signatures in the Primary Microwave Background Bispectrum. *Physycal Review D **63** 063002*

[53] **Lan, X. and Marinucci, D. (2008)** The Needlets Bispectrum, *Electronic Journal of Statistics, Vol. 2, pp.289-319*

[54] **Lan, X. and Marinucci, D. (2008)** On the Dependence Structure of Wavelet Coefficients for Spherical Random Fields, under revision for *Stochastic Processes and their Applications,* arXiv:0805.4154

[55] **Leonov,V.P. and Shiryaev, A.N. (1959)** On a Method of Calculations of Semi-invariants, *Theory Probab.Appl. 4,319-329.*

[56] **Marinucci, D. (2006)** High-Resolution Asymptotics for the Angular Bispectrum of Spherical Random Fields, *The Annals of Statistics 34, 1–41*

[57] **Marinucci, D. (2007)** A Central Limit Theorem and Higher Order Results for the Angular Bispectrum, *Probability Theory and Related Fields, no. 3-4, 389–409*

[58] **Marinucci, D. and Peccati, G. (2007)** Group Representations and High-Resolution Central Limit Theorems for Subordinated Spherical Random Fields, under revision for *Bernoulli,* arXiv:0706.2851

[59] **Marinucci, D. and Peccati, G. (2008)** Representations of $SO(3)$ and Angular Polyspectra, under revision for *Journal of Multivariate Analysis, arXiv:0807.0687*

[60] **Marinucci, D. and Piccioni, M. (2008)** The Empirical Process on Gaussian Spherical Harmonics, *Annals of Statistics 2004, Vol. 32, No. 3, 1261-1288*

[61] **Marinucci, D., Pietrobon, D., Balbi, A., Baldi, P., Cabella, P., Kerkyacharian, G., Natoli, P., Picard, D., Vittorio, N. (2008)** Spherical Needlets for CMB Data Analysis, *Monthly Notices of the Royal Astronomical Society, Vol.383, 539-545*

[62] **Martinez-Gonzalez,E., Gallegos, J.E., Argueso, F., Cayon, L., Sanz, J.L. (2002)** The Performance of Spherical Wavelets to Detect non-Gaussianity in the CMB Sky, *Monthly Notices of the Royal Astronomical Society, Vol. 336, no.1 pp.22-32*

[63] **Mayeli, A. (2008)** Asymptotic Uncorrelation for Mexican Needlets, arXiv:0806.3009

[64] **McEwen, J. D., Vielva, P., Wiaux, Y., Barreiro, R. B., Cayon, L., Hobson, M. P., Lasenby, A. N., Martinez-Gonzalez, E., Sanz, J. (2007)** Cosmological Applications of a Wavelet Analysis on the Sphere, *Journal of Fourier Analysis and its Applications, 13, 495-510*

[65] **Mhaskar, H.N, Narcowich, F.J and Ward, J.D. (2000)** Spherical Marcinkiewicz-Zygmund Inequalities and Positive Quadrature, *Math. Comp. 70 (2001), no. 235, 1113–1130. 41A55 (42C10 65D30)*

[66] **Narcowich, F. J., Petrushev, P. and Ward, J. D. (2006a)** Localized Tight Frames on Spheres, *SIAM Journal of Mathematical Analysis 38, 2, 574–594*

[67] **Narcowich, F. J., Petrushev, P. and Ward, J. D. (2006b)** Decomposition of Besov and Triebel-Lizorkin Spaces on the Sphere, *Journal of Functional Analysis 238, 2, 530–564*

[68] **Nourdin, I. and Peccati, G. (2007)** Non-central Convergence of Multiple Integrals, *Annals of Probability, to appear,* arXiv:0709.3903

[69] **Nualart, D. and Peccati, G. (2005)** Central Limit Theorems For Sequences of Multiple Stochastic Integrals, *Annals of Probability 33, 1, 177-193.*

[70] **Park, C. (2004)** Non-Gaussian Signatures in the Temperature Fluctuation Observed by the Wilkinson Microwave Anisotropy Probe, *Mon.Not.Roy.Astron.Soc. 349 313-320*

[71] **Peccati, G. and Pycke, J. R. (2005)** Decompositions of Stochastic Processes Based on Irreductible Group Representations, arXiv:math/0509569

[72] **Peccati, G. and Tudor, C. A. (2005)** Gaussian Limits for Vector-valued Multiple Stochastic Integrals. *Séinaire de Probabilités XXXVIII, 247–262, Lecture Notes in Math., 1857, Springer, Berlin*

[73] **Peccati, G., and Taqqu, M S. (2008)** Moments, Cumulants and Diagram Formulae for Non-linear Functionals of Random Measures, arXiv:0811.1726

[74] **Pietrobon, D., Amblard, A., Balbi, A., Cabella, P., Cooray, A., Marinucci, D. (2008)** Needlet Detection of Features in WMAP CMB Sky and the Impact on Anisotropies and Hemispherical Asymmetries, *Phys.Rev.D78:103504*

[75] **Pietrobon, D., Balbi, A., Marinucci, D. (2006)** 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, 74, 043524*

[76] **Rudjord, O., Hansen, F., Lan, X., Liguori,M., Marinucci, D., Matarrese,S., (2008)** An Estimate of the Primordial Non-Gaussianity Parameter f_NL Using the Needlet Bispectrum from WMAP, arXiv:0901.3154

[77] **Smoot, G.F.; Bennett, C.L.; Kogut, A.; Wright, E.L.; Aymon, J; Boggess, N.W.; Cheng, E.S.; de Amici, G.; Gulkis, S.; Hauser, M.G.; Hinshaw, G.; Jackson, P.D.; Janssen, M.; Kaita, E.; Kelsall, T.; Keegstra, P.; Lineweaver, C.; Loewenstein, K.; Lubin, P.; Mather, J.; Meyer, S.S.; Moseley, S.H.; Murdock, T.; Rokke, L.; Silverberg, R.F.; Tenorio, L.; Weiss, R.; Wilkinson, D.T. (1992)** Structure in the COBE Differential Microwave Radiometer First-Year Maps, *Astrophysical Journal, Part 2-Letters, Vol. 396, no.1, pp. L1-L5*

[78] **Spivak, M. (1979)** *A Comprehensive Introduction to Differential Geometry Vol. I, Berkeley*

[79] **Straf, M.L. (1972)** Weak Convergence of Stochastic Processes with Several Parameters, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, II, 187–221.*

[80] **Surgailis, D. (2003)** CLTs for Polynomials of Linear Sequences: Diagram Formula with Illustrations,(English summary) Theory and applications of long-range dependence, *111–127, Birkhäuser Boston, Boston, MA, 2003. 60F05 (60G15)*

[81] **Surgailis, D. (2000)** Long-range Dependence and Appell Rank, *Ann. Probability, Vol.28, No.1, 478-497.*

[82] **Sweldens, W.(1996)** The Lifting Scheme: a Custom-design Construction of Biorthogonal Wavelets. *Appl. Comput. Harmon. Anal. 3 (1996), no. 2, 186–200*

[83] **Tenorio, L., Jaffe, A.H., Hanany, S., Lineweaver, C.H.(1999)** Applications of Wavelets to the Analysis of Cosmic Microwave Background Maps, *Monthly Notice of the Royal Astronomical Society, Vol. 310, 823*

[84] **Varshalovich, D.A., Moskalev, A.N. and Khersonskii, V.K. (1988)** *Quantum Theory of Angular Momentum.* World Scientific, Singapore

[85] **Vielva, P., Martinez-Gonzalez, E., Gallegos,J. E., Toffolatti,L. and Sanz,J. L. (2003)** Point Source Detection using the Spherical Mexican Hat Wavelet on Simulated All-sky Planck Maps, *Monthly Notice of the Royal Astronomical Society, Vol. 344, no.1, pp. 89-104*

[86] **Vielva, P., Martinez-Gonzalez,E.,Cayon, L., Diego, J.M., Sanz, J.L., Toffolatti, L. (2001)** Predicted Planck Extragalactic Point Source Catalogue, *Mon.Not.Roy.Astron.Soc. Vol.326, 181*

[87] **Vilenkin, N. Ja. and Klimyk, A. U. (1993)** *Representation of Lie Groups and Special functions Vol. 2: Class I Representations, Special Functions, and Integral Transforms,* KluwerAcademic Press, Dordrecht

[88] **Yadav, A. P. S. and Wandelt, B. D. (2008)** Detection of Primordial non-Gaussianity (f_{NL}) in the WMAP 3-Year Data at above 99.5% Confidence, *Phys.Rev.Lett.100:181301*

[89] **Wiaux, Y., McEwen, J.D., Vielva, P., (2007)** Complex Data Processing: Fast Wavelet Analysis on the Sphere, *Journal of Fourier Analysis and its Applications, 13, 477-494*

Appendix A
The 3-j Symbols

Here we list a few properties of the Wigner's 3-j symbols which are relevant for the analysis of the bispectrum. A much more detailed discussion can be found in [84], chapter 8. An explicit formula for Wigner's 3-j is as follows

$$\begin{aligned}
&\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \qquad (A.1)\\
= \;& (-1)^{j_1-j_2-m_3} \left[(j_1+j_2-j_3)!\,(j_1-j_2+j_3)!\,(-j_1+j_2+j_3)!\,(j_1+m_1)!\,(j_1-m_1)!\right.\\
& \left.\times (j_2+m_2)!\,(j_2-m_2)!\,(j_3+m_3)!\,(j_3-m_3)!\right]^{1/2} / \left[(j_1+j_2+j_3+1)!\right]^{1/2}\\
& \times \sum_z (-1)^z / z!\,(j_1+j_2-j_3-z)!\,(j_1-m_1-z)!\,(j_2-m_2-z)!\\
& \times (j_3-j_2+m_1+z)!\,(j_3-j_1-m_2+z)!\,.
\end{aligned}$$

The 3-j symbol is zero when either of the following conditions is not all satisfied:

$$\left.\begin{array}{c} m_1+m_2+m_3=0; \\ j_1+j_2-j_3\geq 0 \\ j_1-j_2+j_3\geq 0 \\ -j_1+j_2+j_3\geq 0 \end{array}\right\} Triangular\ Conditions.$$

(a) *Symmetry properties*

$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ is invariant under even permutation of columns and is multiplied by the factor $(-1)^{j_1+j_2+j_3}$ for odd permutations.Also

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}.$$

There are altogether 72 known symmetries of the 3-j symbol;

(b) (symmetry) for any triple l_1, l_2, l_3

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} l_2 & l_3 & l_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = \begin{pmatrix} l_3 & l_1 & l_2 \\ m_3 & m_1 & m_2 \end{pmatrix}$$

$$= (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_3 & l_2 & l_1 \\ m_3 & m_2 & m_1 \end{pmatrix} = (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_1 & l_3 & l_2 \\ m_1 & m_3 & m_2 \end{pmatrix}$$
$$= (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_2 & l_1 & l_3 \\ m_2 & m_1 & m_3 \end{pmatrix};$$

(c) *Orthogonality properties*

$$\sum_{m_1,m_2,m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 = 1,$$

$$\sum_{j_3 m_3} (2j_3+1) \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1' & m_2' & m_3' \end{pmatrix} = \delta_{m_1 m_1'} \delta_{m_2 m_2'},$$

$$\sum_{m_1 m_2} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3' \\ m_1 & m_2 & m_3' \end{pmatrix} = \frac{\delta_{j_3 j_3'} \delta_{m_3 m_3'}}{(2j_3+1)};$$

(d) *Useful special cases*

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ 0 & 0 & 0 \end{pmatrix} = (-1)^{1/2J} \left[\frac{(J-2j_1)!\,(J-2j_2)!\,(J-2j_3)!}{(J+1)!} \right]^{1/2} \times \frac{(1/2J)!}{\left(\frac{1}{2}J-j_1\right)!\left(\frac{1}{2}J-j_2\right)!\left(\frac{1}{2}J-j_3\right)!}$$

(non-zero for $J = j_1 + j_2 + j_3$ even);

(e) *Upper bound*

for any j_1, j_2, j_3

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = O\left(\left[\max\{j_1, j_2, j_3\} \right]^{-1/2} \right).$$

Appendix B
Construction of $b(\cdot)$

There are of course many explicit recipes for the construction of functions $b(\cdot)$ which satisfy the properties we required for needlets; one simple example is the following (see also [61]):

STEP 1: Construct the function

$$f(t) = \begin{cases} \exp(-\frac{1}{1-t^2}), & -1 \leq t \leq 1 \\ 0, & otherwise \end{cases} .$$

It is immediate to check that the function $f(\cdot)$ is C^∞ and compactly supported in the interval $(-1, 1)$;

STEP 2: Construct the function

$$\phi(u) = \frac{\int_{-1}^{u} f(t)dt}{\int_{-1}^{1} f(t)dt}.$$

The function $\phi(\cdot)$ is again C^∞; it is moreover non-decreasing and normalized so that $\phi(-1) = 0, \phi(1) = 1$;

STEP 3: Construct the function

$$\varphi(t) = \begin{cases} 1 & if \quad 0 \leq \quad t \quad \leq \frac{1}{B} \\ \phi(1 - \frac{2B}{B-1}(t - \frac{1}{B})) & if \quad \frac{1}{B} < \quad t \quad \leq 1 \\ 0 & if \qquad\quad t \quad > 1 \end{cases} .$$

Here we are simply implementing a change of variable so that the resulting function $\phi(\cdot)$ is constant on $(0, B^{-1})$ and monotonically decreasing to zero in the interval $(B^{-1}, 1)$. Indeed it can be checked that

$$1 - \frac{2B}{B-1}(t - \frac{1}{B}) = \begin{cases} 1 & for \quad t = \frac{1}{B} \\ -1 & for \quad t = 1 \end{cases}$$

and

$$\phi(\frac{1}{B}) = \phi(1) = 1,$$

$$\phi(1) = \phi(-1) = 0;$$

STEP 4: Construct

$$b^2(\xi) = \phi(\frac{\xi}{B}) - \phi(\xi)$$

and for $b(\xi)$ we take the positive root. In view of all the above four steps, we see that $b(.) \in C^\infty$ in this construction method.

For the needlets

$$\varphi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_l b(\frac{l}{B^j}) \sum_{m=-l}^{l} Y_{lm}(x) Y_{lm}^*(\xi_{jk}),$$

as mentioned before, the main result of Narcowich,Petrushev and Ward [66] is the followings localization property of φ_{jk}.

For any M there exists a constant c_M such that, for every $\eta \in S^2$:

$$\left|\varphi_{jk}(\eta)\right| \leq \frac{c_M B^j}{\left(1 + B^j \arccos\left\langle \eta, \xi_{jk} \right\rangle\right)^k}.$$

Likewise, the main result of [7] lemma 5 is as follows :

$$\left|Cor(\beta_{jk}, \beta_{jk'})\right| \leq \frac{C_M}{\left(1 + B^j d\left(\xi_{jk}, \xi_{jk'}\right)\right)^M}.$$

Our aim here is to provide more explicit results for the bound C_M. More precisely, in [66] it indicates that if a function $b(t) \in C^M(\mathbb{R})$ satisfies

$$\left|b^{(r)}(t)\right| \leq C_b (1 + |t|)^{r-\alpha}, \text{ for all } t \in \mathbb{R}, r = 0, ..., M,$$

where $\alpha > 2 + M$, as well as

$$\left\| \frac{d^r}{dt^r} \left\{ t^j b \right\} \right\|_{L^1} \leq B_{M,b}, \text{ for } r = 0, ..., M, \ j = 0, 1,$$

then we can get the bound

$$c_M = \max \left\{ 6C_b/\omega_n, \frac{7\omega_{n-1} B_{M,b}}{2\sqrt{\pi}} \right\},$$

and

$$C_M = c_0 c_M = c_0 \max \left\{ 6C_b/\left(4\pi^2\right), \frac{14\pi B_{M,b}}{2\sqrt{\pi}} \right\},$$

in view of $n = 2$. Here c_0 is the same constant as in Condition A.

Without loss of generality, we take $M = 3$. Then it is easy to see that we can get the bound C_M through calculating the upper bounds of the derivatives of $b(t)$ up to the third order. To exploit $b(t)$ a little bit more, we notice that

$$\begin{aligned} b^2(t) &= \frac{1}{c} \int_{-1}^{x_1} f(s) ds \Big|_{x_1 = 1 - \frac{2}{B-1}(t-1)} - \frac{1}{c} \int_{-1}^{x_2} f(s) ds \Big|_{x_2 = 1 - \frac{2B}{B-1}(t - \frac{1}{B})} \\ &= \begin{cases} \frac{1}{c} \int_{-1}^{x_1} \exp(-\frac{1}{1-s^2}) ds \Big|_{x_1 = 1 - \frac{2}{B-1}(t-1)}, & 1 < t \leq B; \\ 1 - \frac{1}{c} \int_{-1}^{x_2} \exp(-\frac{1}{1-s^2}) ds \Big|_{x_2 = 1 - \frac{2B}{B-1}(t - \frac{1}{B})}, & \frac{1}{B} \leq t \leq 1; \\ 0, & otherwise. \end{cases} \end{aligned}$$

Therefore, we can divide $b(t)$ into 3 cases, which are $t \in \left[\frac{1}{B}, 1\right]$, $[1, B]$ and $\mathbb{R} / \left[\frac{1}{B}, B\right]$.

(i) $1 \leq t \leq B$; here we take $x_1 = 1 - \frac{2}{B-1}(t-1)$,

$$\begin{aligned} b'(t) &= b(t)^{-1} \left(\frac{-1}{c(B-1)} \exp(-\frac{1}{1-x_1^2}) \right. \\ &= \left(\int_{-1}^{x_1} \exp(-\frac{1}{1-s^2}) ds \right)^{-1/2} \frac{-1}{\sqrt{c}(B-1)} \exp(-\frac{1}{1-x_1^2}) \end{aligned}$$

whence

$$\max |b'(t)| \leq 1.85/(B-1), t_{\max} = 0.96175B + 0.03825 \ (< B) \ .$$

Moreover

$$\begin{aligned} b''(t) &= b(t)^{-1}\left(\frac{-4}{c(B-1)^2}\frac{\exp(-\frac{1}{1-x_1^2})x_1}{(1-x_1^2)^2}-[b'(t)]^2\right) \\ &= \frac{-\exp(-\frac{1}{1-x_1^2})}{\sqrt{c}(B-1)^2}\left(\int_{-1}^{x_1}\exp(-\frac{1}{1-s^2})ds\right)^{-3/2}\left(\frac{4x_1\int_{-1}^{x_1}\exp(-\frac{1}{1-s^2})ds}{(1-x_1^2)^2}+\exp(-\frac{1}{1-x_1^2})\right) \end{aligned}$$

which leads to

$$\max|b''(t)| \leq 32.33(B-1)^{-2}, t_{\max} = 0.8476B + 0.1524.$$

Also

$$\begin{aligned} b^{(3)}(t) &= b(t)^{-1}\left(\frac{-4}{c(B-1)^2}\frac{\exp(-\frac{1}{1-x_1^2})}{(1-x_1^2)^4}\left(1-3x_1^4\right)-3b'(t)b''(t)\right) \\ &= \frac{-1}{\sqrt{c}(B-1)^3}\left(\int_{-1}^{x_1}\exp(-\frac{1}{1-s^2})ds\right)^{-5/2}\exp(-\frac{1}{1-x_1^2}) \\ &\quad\times\left\{\left(4\int_{-1}^{x_1}\exp(-\frac{1}{1-s^2})ds\right)^2\frac{(1-3x_1^4)}{(1-x_1^2)^4}\right. \\ &\quad\left.+3\exp(-\frac{1}{1-x_1^2})\left(\frac{4x_1\int_{-1}^{x_1}\exp(-\frac{1}{1-s^2})ds}{(1-x_1^2)^2}+\exp(-\frac{1}{1-x_1^2})\right)\right\} \end{aligned}$$

which gives

$$\max\left|b^{(3)}(t)\right| \leq 24602(B-1)^{-3}, t_{\max} = 0.9753B + 0.0247.$$

(ii) $\frac{1}{B} \leq t \leq 1$; here we take $x_2 = 1 - \frac{2B}{B-1}(t - 1/B)$,

$$\begin{aligned} b'(t) &= b(t)^{-1}\frac{B}{c(B-1)}f(x_2) \\ &= \left(1-\frac{1}{c}\int_{-1}^{x_2}\exp(-\frac{1}{1-s^2})ds\right)^{-1/2}\frac{B\exp(-\frac{1}{1-x_2^2})}{c(B-1)} \end{aligned}$$

giving

$$\max|b'(t)| \leq 1.85B/(B-1), t_{\max} = \frac{0.8476}{B} + 0.1524\,(> 1/B).$$

Also

$$\begin{aligned} b''(t) &= b(t)^{-1}\left(\frac{4B^2}{c(B-1)^2}\frac{\exp(-\frac{1}{1-x_1^2})x_1}{(1-x_1^2)^2} - [b'(t)]^2\right) \\ &= \left(1-\frac{1}{c}\int_{-1}^{x_2}\exp(-\frac{1}{1-s^2})ds\right)^{-3/2}\frac{B^2\exp(-\frac{1}{1-x_1^2})}{c(B-1)^2} \\ &\quad\times\left(\frac{4x_1\left(1-\frac{1}{c}\int_{-1}^{x_2}\exp(-\frac{1}{1-s^2})ds\right)}{(1-x_1^2)^2} - \exp(-\frac{1}{1-x_1^2})\right) \end{aligned}$$

so that we obtain

$$\max|b''(t)| \leq 64B^2(B-1)^{-2}, t_{\max} = \frac{0.95755}{B} + 0.04245\ ,$$

and finally

$$\begin{aligned} b^{(3)}(t) &= b(t)^{-1}\left(\frac{-4B^2}{c(B-1)^2}\frac{\exp(-\frac{1}{1-x_1^2})}{(1-x_1^2)^4}\left(1-3x_1^4\right) - 3b'(t)b''(t)\right) \\ &= \frac{-B^3}{c(B-1)^3}\left(1-\frac{1}{c}\int_{-1}^{x_2}\exp(-\frac{1}{1-s^2})ds\right)^{-5/2}\exp(-\frac{1}{1-x_1^2}) \\ &\quad\times\{4\left(1-\frac{1}{c}\int_{-1}^{x_2}\exp(-\frac{1}{1-s^2})ds\right)^2\frac{(1-3x_1^4)}{(1-x_1^2)^4} \\ &\quad+3\exp(-\frac{1}{1-x_1^2})\left(\frac{4x_1\left(1-\frac{1}{c}\int_{-1}^{x_2}\exp(-\frac{1}{1-s^2})ds\right)}{(1-x_1^2)^2} - \exp(-\frac{1}{1-x_1^2})\right)\} \end{aligned}$$

which gives

$$\max\left|b^{(3)}(t)\right| \leq 3360B^3(B-1)^{-3}, t_{\max} = \frac{0.9591}{B} + 0.0409\ .$$

Combining the all the bounds of derivatives (up to the third order) of $b(.)$ above, we obtain

$$C_b, B_{3,b} \leq 24603(B-1)^{-3}\ .$$

Therefore $C_M = \max\left\{6C_b/(4\pi^2), \frac{14\pi B_{3,b}}{2\sqrt{\pi}}\right\} c_0 = 305254(B-1)^{-3}c_0$.

The constant seems very large, due to form of $b(\cdot)$. Here we give another method of construction of $b(\cdot)$. The steps for constructing $b(x) \in C^{\alpha}$ $(\alpha < k + 1/2)$ are as follows:

1) define polynomials (of Benstein)

$$B_i^{(n)}(t) = \left(\begin{array}{c} n \\ i \end{array} \right) t^i (1-t)^{n-1}.$$

Example: for $n = 1$

$$B_0^{(1)}(t) = (1-t) \, , B_1^{(1)}(t) = t \, ,$$

for $n = 2$

$$B_0^{(2)}(t) = (1-t)^2 \, , B_1^{(2)}(t) = 2t(1-t) \, , B_2^{(2)}(t) = t^2 \, ,$$

for $n = 3$

$$B_0^{(3)}(t) = (1-t)^3 \, , B_1^{(3)}(t) = 3t(1-t)^2 \, , B_2^{(3)}(t) = 3t^2(1-t) \, , B_3^{(3)}(t) = t^3 \, .$$

2) define polynomials

$$p_{2k+1}(t) = \sum_{i=0}^{k} c_i B_i^{(2k+1)}(t) \, ,$$

where

$$c_i = 1 \text{ , for } i = 1, ..., k, \text{ and } c_i = 0 \text{ otherwise.}$$

Example: for $k = 1$

$$p_3(t) = \sum_{i=0}^{1} B_i^{(3)}(t) = (1-t)^3 + 3t(1-t)^2 \, ,$$

$$p_5(t) = \sum_{i=0}^{2} B_i^{(5)}(t) = (1-t)^5 + 5t(1-t)^4 + 10t^2(1-t)^3 \, .$$

Note that

$$p_3(0) = p_5(0) = 1\ ,$$

and

$$\begin{aligned} p_3'(1) &= p_3'(0) = 0\ , \\ p_5'(t) &= 30t^2(1-t)^2, \end{aligned}$$

Therefore the derivatives of p_5 have value zero up to order 2 at points 0 and 1. In general

$$p_{2k+1}^{(r)}(1) = p_{2k+1}^{(r)}(0) = 0 \text{ for } r = 1, ..., k\ .$$

3) define

$$t = \frac{x - 1/B}{1 - 1/B}$$

where $B > 1$.

4) define the function

$$\varphi(x) := \begin{cases} 1, \text{ if } x \in [0, \frac{1}{B}] \\ p_{2k+1}(\frac{x-1/B}{1-1/B}), \text{ if } x \in [\frac{1}{B}, 1] \\ 0, \text{ if } x > 1 \end{cases}\ .$$

5) Finally, we have

$$b(x) = \begin{cases} \sqrt{\varphi(\frac{x}{B}) - \varphi(x)}\ , \frac{1}{B} \le x \le B \\ 0\ , \text{otherwise} \end{cases}\ .$$

In view of the examples given in this construction, we can also easily see that the constant C_M is in the order of $\frac{M^M}{(B-1/B)^M}$, which is rather small compared with the one we obtained from the first recipe. For the sake of brevity and also due to its similarity to the procedure we have done for the former one, we omit the computation of C_M derived from this latter method.

Appendix C
A proof for the bound of $\frac{dP_l(\cos\theta)}{d\theta}$

In this final appendix, we shall provide some bounds on Legendre polynomials which were used throughout the thesis. These results are classical, and can be obtained by many techniques - we provide one more derivation for completeness.

Let us first recall a few properties of Legendre polynomials, that we shall use extensively soon (see [84] for details); we have

$$\left(1-x^2\right)\frac{dP_l^m(x)}{dx} = \left(xP_l^m(x) - P_{l-1}(x)\right)l\ .$$

If we replace $x = \cos\theta$, we obtain

$$\frac{dP_l(\cos\theta)}{d\theta} = \frac{1}{\sin\theta}\left(\cos\theta P_l(\cos\theta) - P_{l-1}(\cos\theta)\right)l\ . \tag{C.1}$$

Differentiating the component in the bracket on the right side of the above differential equality, we have

$$\begin{aligned}
& \frac{d}{d\theta}\left(\cos\theta P_l(\cos\theta) - P_{l-1}(\cos\theta)\right) \\
= & -\sin\theta P_l(\cos\theta) + \frac{\cos\theta}{\sin\theta}\left(\cos\theta P_l(\cos\theta) - P_{l-1}(\cos\theta)\right)l \\
& -\frac{1}{\sin\theta}\left(\cos\theta P_{l-1}(\cos\theta) - P_{l-2}(\cos\theta)\right)(l-1) \\
= & -(l+1)\sin\theta P_l(\cos\theta) \\
& +\frac{1}{\sin\theta}\left(lP_l(\cos\theta) - (2l-1)\cos\theta P_{l-1}(\cos\theta) + (l-1)P_{l-2}(\cos\theta)\right) \\
= & -(l+1)\sin\theta P_l(\cos\theta)\ .
\end{aligned}$$

The last equality comes from the fact that the Legendre polynomials satisfy

$$lP_l(\cos\theta) = (2l-1)\cos\theta P_{l-1}(\cos\theta) - (l-1)P_{l-2}(\cos\theta) \ .$$

We know that the Dirichlet-Mehler Integral representation for the Gegenbauer polynomials is

$$P_l^{(\lambda)}(\cos\theta) = \frac{2^\lambda \Gamma\left(\lambda+\frac{1}{2}\right)\Gamma(2\lambda+l)}{\sqrt{\pi}l!\Gamma(\lambda)\Gamma(2\lambda)(\sin\theta)^{2\lambda-1}} \int_\theta^\pi \frac{\cos\left((l+\lambda)\varphi - \lambda\pi\right)}{(\cos\theta-\cos\varphi)^{1-\lambda}} d\varphi,$$

and in our case $P_l(\cos\theta) = P_l^{\left(\frac{1}{2}\right)}(\cos\theta)$. Thus

$$\begin{aligned}
\frac{dP_l(\cos\theta)}{d\theta} &= \frac{-l(l+1)}{\sin\theta}\int_0^\theta \sin\psi \frac{\sqrt{2}\Gamma(1)}{\sqrt{\pi}\Gamma(1/2)}\int_\psi^\pi \frac{\cos\left(\left(l+\frac{1}{2}\right)\varphi - \frac{1}{2}\pi\right)}{(\cos\psi-\cos\varphi)^{1/2}} d\varphi d\psi \\
&= \frac{-l(l+1)}{\sin\theta}\left(\int_0^\theta d\varphi\int_0^\varphi + \int_\theta^\pi d\varphi\int_0^\theta\right)\sin\left(l+\frac{1}{2}\right)\varphi \frac{\sin\psi d\psi}{(\cos\psi-\cos\varphi)^{1/2}} \\
&= \frac{-l(l+1)}{\sin\theta}\Big[\int_0^\pi 2\sin\left(l+\frac{1}{2}\right)\varphi(1-\cos\varphi)^{1/2} d\varphi \\
&\quad - \int_\theta^\pi 2\sin\left(l+\frac{1}{2}\right)\varphi(\cos\theta-\cos\varphi)^{1/2} d\varphi\Big]. \qquad \text{(C.2)}
\end{aligned}$$

The first part in the bracket is equal to

$$\int_0^\pi 2\sin\left(l+\frac{1}{2}\right)\varphi(1-\cos\varphi)^{1/2} d\varphi = \int_0^\pi 2\sqrt{2}\sin\left(l+\frac{1}{2}\right)\varphi\sin\frac{\varphi}{2} d\varphi = 0, \qquad \text{(C.3)}$$

while for the second part , we can see that it is absolutely integrable, so as $m, j \to \infty$, we can obtain

$$
\begin{aligned}
& \int_{\theta}^{\pi} 2\sin\left(l+\frac{1}{2}\right)\varphi\left(\cos\theta-\cos\varphi\right)^{1/2}d\varphi \\
= & \lim_{m\to\infty}\frac{2\pi}{(2l+1)\,m}\sum_{k=[(l+1/2)\theta/\pi]}^{l}\sum_{h=0}^{m}\{\sin\left(2k\pi+\frac{\pi h}{2m}\right)\left(\cos\theta-\cos\frac{2k+h/m}{2l+1}\pi\right)^{1/2} \\
& -\sin\left(2k\pi+\frac{\pi h}{2m}\right)\left(\cos\theta-\cos\frac{2k+2+h/m}{2l+1}\pi\right)^{1/2} \\
& +\sin\left(2k\pi+\frac{\pi}{2}+\frac{\pi h}{2m}\right)\left(\cos\theta-\cos\frac{2k+1+h/m}{2l+1}\pi\right)^{1/2} \\
& -\sin\left(2k\pi+\frac{\pi}{2}+\frac{\pi h}{2m}\right)\left(\cos\theta-\cos\frac{2k+3+h/m}{2l+1}\pi\right)^{1/2}\} \\
= & \lim_{m\to\infty}\frac{\pi}{(2l+1)\,m}\sum_{h=0}^{m}\sin\frac{\pi h}{2m}\sum_{k=[(l+1/2)\theta/\pi]}^{l}\frac{d}{dx}\left(\cos\theta-\cos\left(\frac{2k+\pi h/m}{2l+1}+x\right)\right)^{1/2}|_{x=0}\,\frac{2\pi}{2l+1} \\
& +\frac{\pi}{(2l+1)\,m}\sum_{h=0}^{m}\cos\frac{\pi h}{2m}\sum_{k=[(l+1/2)\theta/\pi]}^{l}\frac{d}{dx}\left(\cos\theta-\cos\left(\frac{2k+1+h/m}{2l+1}+x\right)\right)^{1/2}|_{x=0}\,\frac{2\pi}{2l+1} \\
= & \lim_{m\to\infty}\frac{\pi}{(2l+1)\,m}\sum_{h=0}^{m}\left(\sin\frac{\pi h}{2m}+\cos\frac{\pi h}{2m}\right)\left(\cos\theta-\cos\left(\frac{\pi h}{(2l+1)\,m}+\frac{2l\pi}{(2l+1)}\right)\right)^{1/2} \\
\leq & \frac{\pi\sqrt{2}}{2l+1}\left(\cos\theta-\cos\pi\right)^{1/2}\ .
\end{aligned}
$$

We are now in the position to establish the upper bound of the first derivative of Legendre polynomial of degree l. First consider the case that $\theta \in [\pi/2, \pi]$, and recall one of the trigonometric equalities $1-\cos\alpha = 2\sin^2\alpha/2$, we can then bound $(C.2)$

with

$$\begin{aligned} & \left| \frac{1}{\sin\theta} \int_{\theta}^{\pi} 2\sin\left(l+\frac{1}{2}\right)\varphi \left(\cos\theta-\cos\varphi\right)^{1/2} d\varphi \right| \\ \leq & \frac{\pi \sin\left(\pi-\theta\right)/2}{\left(l+1/2\right)\sin\left(\pi-\theta\right)} \\ = & \frac{\pi}{\left(2l+1\right)\cos\left(\pi-\theta\right)/2}. \end{aligned} \tag{C.4}$$

Therefore

$$\left| \frac{dP_l\left(\cos\theta\right)}{d\theta} \right| \leq \frac{\pi l\left(l+1\right)}{\left(2l+1\right)\cos\left(\pi-\theta\right)/2} \leq 3l,$$

in view of $(C.3)$, $(C.4)$ and $(C.2)$.

Similarly, in the case of $\theta \in [0, \pi/2]$, we do the transform $\widetilde{\theta} = \pi - \theta, \widetilde{\varphi} = \pi - \varphi$ into the integral representation of $P_l\left(\cos\theta\right)$,

$$P_l^{(\lambda)}\left(\cos\theta\right) = \frac{\sqrt{2}\Gamma\left(1\right)}{\sqrt{\pi}\Gamma\left(1/2\right)} \int_0^{\theta} \frac{\left(-1\right)^l \cos\left(l+1/2\right)\varphi}{\left(\cos\varphi-\cos\theta\right)^{1/2}} d\varphi,$$

The follwoing steps are the same as before. Thus we can obtain the same bound, and make the conclusion that the following inequality exists for any $\theta \in [0, \pi]$, which is

$$\left| \frac{dP_l\left(\cos\theta\right)}{d\theta} \right| \leq 3l.$$