Stochastic Modelling and Statistical Analysis of Spatial and Long-Range Dependent Data

Submitted by

Dangedara Egodage Ravindi Sachithra Nanayakkara

Bachelor of Science (Statistics with Computer Science) with Honours, 2016

University of Colombo

A thesis submitted in total fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics and Statistics

School of Engineering and Mathematical Sciences

College of Science, Health and Engineering

La Trobe University

Victoria

Australia

September 2021

Contents

List of Figures

List of Tables

Abstract

This thesis studies the stochastic modelling and statistical analysis of spatial and long-range dependent data. Temporal and spatial-dependent data are often encountered in many areas such as cosmology, geoscience, finance, genomics, embryology, hydrology, etc. It is important to study the long-range dependence and spatial dependence of such data for time series and spatial analysis.

First, the multifractality of spherical random fields is studied with cosmological applications. The motivation of this study is to investigate the cosmic microwave background radiation (CMB) data from the Planck mission. For random fields on the sphere, there are three models in the literature where the Rényi function is known explicitly. In this study, some new theoretical models and numerical multifractality studies are presented. Then, the methodology based on computing the Rényi function and the multifractal spectrum for different scenarios and actual CMB data is shown. The results suggest that a very minor multifractality of the CMB data may exist.

Next, the multifractionality of spherical random fields is studied with cosmological applications. The Hölder exponent is used to measure the roughness of random fields in a rigorous mathematical way. The pointwise Hölder exponent values are estimated for one- and two-dimensional sky regions using the HEALPix ring and nested orderings respectively. The analysis suggests that the CMB data are multifractional. The developed methodology is also used to detect anomalous regions in inpainted CMB maps.

Finally, simultaneous estimators of cyclic long-memory processes are studied. Spectral singularities at non-zero frequencies play an important role in investigating cyclic or seasonal time series. A wide class of semiparametric models with spectral singularities is studied. The generalized filtered method of moments simultaneous estimators of singularity location and long-memory parameters are considered. The results of the asymptotic normality of several statistics are obtained. The methodology includes wavelet transformations as a particular case.

Statement of Authorship

This thesis includes work by the author that has been published or accepted for publication as described in the text. Except where reference is made in the text of the thesis, this thesis contains no other material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Dangedara Egodage Ravindi Sachithra Nanayakkara

27/09/2021

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my principal supervisor, Associate Professor Dr. Andriy Olenko, for his excellent guidance, patience, and encouragement throughout my PhD candidature. Without his guidance, assistance, and dedicated involvement in every step throughout the process, this task would not have been accomplished. I would like to thank you very much for your immense support and friendship over the last few years. I would like to thank my co-supervisor, Professor Philip Broadbridge, for his valuable advice and constructive feedback given at the various stages of my PhD studies and for numerous discussions regarding cosmology and underlying physics models.

I would like to thank my co-authors, Professors Nikolai Leonenko, Antoine Ayache, and Myriam Fradon, for their ideas, contribution and for giving me the opportunity to learn from them. I would like to give special thanks to all the academic and administrative staff of the Department of Mathematics and Statistics for the friendship and knowledge they have imparted to me providing me with necessary technical support whenever needed.

Last but not the least, I would like to thank my loving parents for giving me the best education they could offer from the little they had and for encouraging me always, my loving husband, Isuru, for all the sacrifices he made behalf of me, and to my two little sisters for supporting me spiritually throughout writing this thesis and my life in general. I remain grateful to my best friends for their support and cooperation in my endeavours.

I am grateful to La Trobe University for the financial support offered to pursue my PhD studies. This work was supported by a La Trobe University Postgraduate Research Scholarship and a La Trobe University Full Fee Research Scholarship. This research was partially supported under the Australian Research Council's Discovery Projects funding scheme (project number DP160101366). This research includes computations using the Linux computational cluster Gadi of the National Computational Infrastructure (NCI), which is supported by the Australian Government and La Trobe University.

Chapter 1

Introduction

This chapter presents the historical background and motivation for the research problems considered in the thesis. This comprises the fundamental concepts, methods and results in the literature related to multifractality and the Rényi function, long-range dependence, multifractionality and the Hölder exponent, cyclic long-memory processes, wavelet-based estimation of parameters of stochastic processes and cosmological background. Finally, Section 1.8 provides a summary of the research problems, the main results and their novelty.

1.1 Stochastic processes and spherical random fields

Many real-world phenomena exhibit random behaviour and they often evolve over space and time in a random manner. Stochastic processes can be used as mathematical models to investigate the random behaviour of these real-life occurrences. Stochastic processes are widely used in many fields, for example, in cosmology, physics, meteorology, image processing, biology, neuroscience, signal processing, computer science, chemistry, ecology, cryptography, hydrology and many others, see Allen (2010), Bartlett (1955), Feller (1950), Gallager (2013), Laing and Lord (2010), Todorovic (2012), Van Kampen (1992), Paul and Baschnagel (1999) and the references therein. Stochastic models are essential for understanding the complex nature of a wide range of data. The probability distributions of the various outcomes can be estimated through the mechanism of stochastic modelling. Thus, these models can be distinguished as models for the progression of a system over time in which the random variable of interest undergoes some random changes according to probabilistic laws.

The theory of stochastic processes originated from the early 20th century. The foundation for the theory of probability was laid by Kolmogorov (1931) and Kolmogorov (1933). The notion of a random function was introduced by Kolmogorov (1933) and subsequently, the concept was further developed by Khintchine (1934). Other mathematicians who contributed significantly to the theory of stochastic processes are William Feller, Joseph Doob, Paul Lévy and Norbert Wiener. Joseph Doob pioneered the theory of stochastic processes for the continuous parameter case in his paper Doob (1937). Then, from 1940 to 1950, he progressed the martingale theory. From 1940 onwards, Kiyosi ltô made a significant contribution to the development of the stochastic calculus field, in particular, stochastic differential equations. He proposed the stochastic integrals in his publication Itô (1944). The results obtained by Paul Lévy were considered another important milestone in stochastic process theory. The monograph by Doob (1953) greatly influenced the theory of stochastic processes and focused on measure theory as the basis for probability theory. The classical monographs by Gikhman and Skorokhod (2004a), Gikhman and Skorokhod (2004b) and Gikhman and Skorokhod (2007) include a comprehensive discussion on the theory and applications of stochastic processes.

In statistics and probability theory, a stochastic process is a collection of random variables $X(t)$, $t \in T$, where for each t, $X(t)$ is a random variable and t varies in the parameter space T. The state space consists of possible values of $X(t)$. A stochastic process can be considered as a random function. Based on the parameter space and the state space, one-dimensional stochastic processes can be classified as stochastic processes with discrete parameter and discrete state space, continuous parameter and continuous state space, discrete parameter and continuous state space and continuous parameter and discrete state space (Cox and Miller (1977)). Stochastic processes are classified as discrete-time and continuous-time stochastic processes. For example, if the parameter space $T = \{0, 1, 2, ...\}$, the resulting stochastic process is a discrete-time stochastic process. In general, it is denoted by $X(n),\ n \in N$. If the parameter space is defined by $T = [0, \infty)$, the resulting process is a continuous-time stochastic process denoted by $X(t),\ t \geq 0$.

Traditionally, the observations are assumed to be independent in the theory of statistics. However, this assumption is too strong and not valid for real data. Thus, random processes with temporal and spatial-dependent observations became popular along with the availability of temporal and spatial data from diverse disciplines such as astronomy, finance, forestry, geology, telecommunications and many others. Accordingly, statistical methods were developed to accommodate such dependent data, see Appel and Pebesma (2020), Christakos (2017), Dehling and Philipp (2002), Emery and Porcu (2019), Emery et al. (2019), Jeong et al. (2017), Porcu et al. (2018) and the references therein for more details. Geostatistics is established on dependent random processes, see Cressie (1989). The covariance and spectral functions play a key role in describing the dependence properties of stochastic processes. Different types of stochastic processes were introduced to study the dependence structure. This thesis mainly deals with short-range and long-range dependent stochastic processes.

The short-range dependent processes are also known as weakly dependent or short-memory dependent processes. A stochastic process is said to be weakly dependent if its covariance function decays rapidly to zero. The spectral densities of such processes are bounded at the origin. Also, the autocovariance functions of short-range dependent processes are absolutely summable. In the literature, different types of weak dependence were introduced. Some examples are association, m-dependence, mixingales and near epoch dependence, quasi-association and strong mixing, see Dedecker et al. (2007), Nze and Doukhan (2004) and Spodarev (2013). Some typical examples for weakly dependent processes are autoregressive–moving-average models, Brownian motion, Markovian models, etc. The mixing conditions are a major tool in describing weak dependence, see

Bradley (2007), Doukhan (1994), Doukhan et al. (2009) and Rio (2017) for more details. The strong mixing coefficients were first introduced by Murray Rosenblatt in 1956 (Rosenblatt (1956)). For stationary stochastic processes, a modern notion of weak dependence was presented in Doukhan and Louhichi (1999). The presence of a strong mixing condition for a stochastic process suggests that it is short-range dependent (Rosenblatt (2015)).

The long-range dependent processes are also called strongly dependent or long-memory dependent processes. In the literature, multiple definitions for these processes exist, see Beran (1994), Beran et al. (2013), Doukhan et al. (2003), Ivanov and Leonenko (1989), Leonenko (1999), Samorodnitsky (2007) and Pipiras and Taqqu (2017). A stochastic process is said to be strongly dependent if it possesses a slowly decaying covariance function, see Section 1.3 for more details. The autocovariance functions of long-range dependent processes are not integrable or absolutely summable. With respect to the spectral domain, their spectral densities are unbounded at the origin or another point. Moreover, the limiting distributions of functionals of the strongly dependent processes are likely to be non-Gaussian. Chapter 5 deals with long-range dependent stochastic processes.

The one-dimensional stochastic processes are inadequate for modelling numerous types of spatial data. A random field is defined as a generalized form of a stochastic process in which the underlying parameter is a multidimensional vector. Thus, the outcome can be seen as a random multivariate function or as a random surface. In the one-dimensional case, the resulting random field is a stochastic process. In real-world applications, the values of random fields are dependent as the observations are spatially correlated with one another. An extensive discussion on the theory of random fields can be found in Hernàndez (1995), Ivanov and Leonenko (1989), Leonenko (1999) and Yadrenko (1983). Some other models of random fields in Physics, specifically for continuum mechanics were considered by Martin Ostoja-Starzewski and his co-authors , see Malyarenko and Ostoja-Starzewski (2019) and Malyarenko et al. (2020). In particular, the tensor-valued random fields underlying the stochastic continuum theories were studied.

There are certain real-life applications that cannot be modelled by random fields defined on classical Euclidean geometries. In such cases, it is vital to deal with the random fields specified on manifolds. In recent years, enormous attention has been given to research on spherical random fields with the growing need for their applications in many fields such as astrophysics, climatology, cosmology, geology, geophysics, medical imaging, oceanography and many others, see Appel and Pebesma (2020), Christakos (2017), Cressie (1993), Fisher et al. (1993), Jeong et al. (2017), Marinucci and Peccati (2011) and the references therein. The monographs by Christakos (2017) and Fisher et al. (1993) demonstrated numerous statistical methods for spherical and spatiotemporal data. Over recent years, various mathematical and statistical software, particularly R packages such as GEOR (Ribeiro Jr et al. (2020)) and RCOSMO (Fryer and Olenko (2019) and Fryer et al. (2020)) and Python packages such as ASTROPY (Price-Whelan et al. (2018)) and HEALPY (Zonca et al. (2019)) were developed to analyze spatial and HEALPix data.

In the literature, random fields defined on the sphere are used as the standard stochastic model, describing various astrophysical, cosmological and environmental data. In this thesis, we mainly consider spherical random fields defined on the unit sphere with cosmological applications. The spectral methods are of great importance in studying and investigating second order random fields with homogeneous and isotropic properties, see Ivanov and Leonenko (1989), Leonenko (1999) and Yadrenko (1983) for more details. Initially, the monograph by Yadrenko (1983) extended the spectral theory of one-dimensional random processes to multidimensional spaces and spheres. It included a wide range of statistical problems that were addressed for spherical random fields such as linear forecasting, extrapolation and optimal linear estimation.

The monograph by Marinucci and Peccati (2011) discussed the recent developments in the theory and statistics of isotropic spherical random fields and stochastic modelling approaches with a focus on cosmological applications. It focused on harmonic analysis tools, studied the properties of angular power spectra and polyspectra, addressed their statistical estimation and investigated the asymptotic properties of spherical needlets. In accordance with the spectral representation of spherical random fields, they can be expanded in a series of spherical harmonics. These series were extensively studied in several publications. For example, Baldi and Marinucci (2007) provided elementary interpretations of the spherical harmonic coefficients of isotropic spherical random fields. They proved that the associated harmonic coefficients are uncorrelated and independent for isotropic Gaussian spherical random fields and the converse remains true. Marinucci (2005) presented Gaussianity tests for spherical random fields.

The angular power spectrum which is the Fourier transform of the covariance function (second-order cumulant) plays an important role in determining the Gaussianity of spherical random fields. Analogously, the higher-order spectra/polyspectra result in the angular bispectrum, trispectrum and so on. Their properties and statistical estimation have been extensively studied in Cammarota and Marinucci (2015), Durastanti et al. (2014), Marinucci (2006), Marinucci (2008), Marinucci and Peccati (2010b), Marinucci and Peccati (2010a) and Marinucci and Peccati (2011). Marinucci (2006) studied the asymptotic behaviour of the bispectrum using an analysis of Wigner coefficients and the obtained results were utilized to develop non-Gaussianity tests for spherical random fields. These results were improved in Marinucci (2008) where a multivariate central limit theorem was derived for the bispectrum using the methods of moments and showed the procedure for higher-order estimation.

In Marinucci and Peccati (2010b) and Marinucci and Peccati (2011), the corresponding results were developed for higher-order moments of isotropic spherical random fields. Marinucci and Peccati (2010b) and Marinucci and Peccati (2011) derived an expression for the angular polyspectra of isotropic spherical random fields using the convolution of Clebsch–Gordan and Wigner coefficients. Marinucci and Peccati (2010a) studied the high-frequency asymptotics of isotropic spherical random fields with a focus on the association between the conditions of ergodicity and asymptotic Gaussianity. Their results suggested that the two conditions are equivalent in many cases. Also, the obtained results are useful to understand the contribution of the cosmic variance problem in the analysis of CMB data. Durastanti et al. (2014) considered the single realization of an isotropic Gaussian spherical random field on the unit sphere and they probed its Gaussian semiparametric estimation, specifically the asymptotic behaviour of their spectral parameter estimators. Cammarota and Marinucci (2015) explored the stochastic behaviour of the isotropic Gaussian spherical random fields that are l^1-regularized and showed that under such circumstances, their properties such as isotropy and Gaussianity are not preserved.

Leonenko and Sakhno (2012) studied the spectral representation of vector- and tensor-valued spherical random fields with the motivation of probing the CMB polarization anisotropies. The spectral decomposition was obtained as a series of the generalized spherical functions for random fields that are weak-isotropic and mean-square continuous. Further, characterizations were obtained for the properties of the associated random spherical harmonic coefficients. Ma (2016) derived the stochastic series representation for an isotropic, mean-square continuous and a stationary vector-valued spherical random field. The obtained infinite series representation was in terms of ultraspherical/Gegenbauer's polynomials. Similar results were obtained by Ma (2017) for isotropic and mean-square continuous vector-valued spherical random fields that are stationary on a temporal-domain. Inspired by the cosmological applications, Leonenko and Ruiz-Medina (2018) considered the two scenarios of Gaussian and chi-squared spherical random fields and obtained central and non-central limit theorems for their first Minkowski functionals.

Various numerical approximation aspects of spherical random fields have been methodically studied by the research group of Ian Sloan, see Hamann et al. (2021), Hesse et al. (2017), Hesse et al. (2021), Le Gia et al. (2017), Le Gia et al. (2020), Sloan and Womersley (2000), Sloan and Wendland (2009), Wang and Sloan (2017), Wang et al. (2017) and Wang et al. (2018). Stochastic partial differential equations (SPDEs) are a major tool in the analysis of temporally evolving spatial data observed on the unit sphere or in the three-dimensional space, see Anh et al. (2008), Broadbridge et al. (2019), Broadbridge et al. (2020), Lang and Schwab (2015) and Ruiz-Medina et al. (2008). Lang and Schwab (2015) studied isotropic spherical random fields and their sample regularity. They also analyzed SPDEs on the unit sphere. They obtained the link between the angular power spectrum decay and other properties of spherical random fields and their approximation. These results formed a solid theoretical background and applied approximation methodology for the development of numeric methods and simulation techniques for spherical random fields.

1.2 Multifractality and the Rényi function

The fractal structures are common in nature, and they were illustrated comprehensively in Mandelbrot (1982). A fractal set has a fractal dimension and often has a pattern which is infinitely scaled and repeats itself. The fractal dimension measures the complexity of self-similar processes, i.e., their change in detail over the change in scale. Initially in 1967, Benoit Mandelbrot discussed the concepts, self-similarity and fractal dimension in his paper Mandelbrot (1967). Then, the results in Mandelbrot (1972) and Mandelbrot (1974) further emphasized the usefulness of scaling relations in the setting of turbulence modelling. For the first time in 1975, Benoit Mandelbrot defined the term fractal which was later developed in Mandelbrot (1977).

The theory of multifractality was introduced as the generalization of fractal sets. In 1982, Benoit Mandelbrot showed the importance of employing fractal theory to measures, see Mandelbrot (1982). Subsequently, the concept of multifractality was theoretically developed by various researchers, see Brown et al. (1992), Evertsz and Mandelbrot

(1992), Falconer (1990), Jaffard (1997a), Jaffard (1997b), Olsen (1995), Muzy et al. (1993) and Riedi (1995) for more details.

A multifractal pattern can be defined as a fractal pattern that scales with multiple scaling rules/fractal dimensions as opposed to a monofractal pattern that scales with one scaling rule/fractal dimension. In multifractal theory, the multifractal/singularity spectrum, also known as the curve of $f(\alpha(q))$ vs. $\alpha(q)$, assesses the strength of the nonlinearity of a fractal process, see Evertsz and Mandelbrot (1992) for more details. Here, $\alpha(q)$ denotes the singularity exponent and it determines the strength/order of singularity, where q is the moment order/resolution. Thus, $f(\alpha(q))$ indicates the local fractal dimensions at different resolutions q. For example, the multifractal spectra of non-fractal and monofractal processes depict an almost constant or flatter behaviour. For multifractal processes, their multifractal spectrum typically exhibits a concave down/parabolic behaviour and the curvature increases with the singularity exponent where it reaches a maximum for the most frequent fractal dimension.

With the initial applications to turbulence modelling (Benzi et al. (1984) and Chhabra et al. (1989)), the multifractal theory has been widely applied in other fields such as cosmology (Coleman and Pietronero (1992), Diego et al. (1999), Leonenko and Shieh (2013) and Martìnez (1990)), geophysics (Mandelbrot (1989) and Parisi and Frisch (1985)), meteorology (Gupta and Waymire (1993)), financial time series (Evertsz (1995), Mandelbrot (1997) and Muzy et al. (2001)), genomics (Arneodo et al. (1998) and Yu et al. (2001)), image processing (Turiel and Parga (2000) and Véhel and Mignot (1994)), internet traffic (Mannersalo et al. (2002), Riedi and Véhel (1997) and Véhel and Riedi (1997)), etc. Martìnez (1990) applied fractals and multifractals to interpret the distribution of large-scale galaxies and showed how the universe conforms to fractal geometry in small scales and collapses the scale invariance beyond a cut-off value. The results suggested that the galaxy distributions adjust to fractal behaviour in some of the scaling regions and the multifractal theory is more suitable to describe the structure of the universe.

When the probability distributions have singularities, the usual measures of location and spread are not suitable to describe conventional probability distributions. Therefore, in such cases, the multifractal measures play a key role. Of the various methods of constructing random multifractal measures, simple binomial cascades, multiplicative cascades and random branching processes were used, see Barral and Mandelbrot (2002), Gupta and Waymire (1993), Falconer (1997), Kahane (1985), Kahane (1987), Molchan

(1996) and Riedi (2002) for more details. Rhodes and Vargas (2010) introduced multi-dimensional multifractal random measures as a generalization of the multifractal random measures proposed by Bacry and Muzy (2003) for the one-dimensional case. Salat et al. (2017) compared the major multifractal methodologies that are used in practice. They mentioned that the moment method and the histogram method with a wide enough scaling range are the most suitable ones for spatial data.

Multifractal analysis deals with measuring the complexity of patterns using fractal dimensions and provides valuable information regarding the local and global complexities of many natural phenomena. Multifractal analysis has been extensively applied to one-dimensional time series data, but its applications in the multidimensional case are less developed. In multifractal analysis, the Rényi function played a vital role (Anh et al. (2008), Leonenko and Shieh (2013) and Mannersalo et al. (2002)). Further, wavelets were also used as one of the main techniques to investigate multifractality, see Audit et al. (2002), Gonçalvés and Riedi (2005), Jaffard (2004) and Riedi (2002) for more details.

The Rényi function ($T(q)$) is associated with the Rényi dimension which is also called the generalized dimension ($D(q)$) by $D(q) = T(q)/(q-1)$, see Evertsz and Mandelbrot (1992) and Harte (2001) for more details. $D(0)$, $D(1)$ and $D(2)$ are called fractal, information and correlation dimensions respectively (Jain et al. (1992) and Jizba and Arimitsu (2004)). The Rényi dimensions emerged from information theory (Rényi (1959), Rényi (1965) and Rényi (1970)). The concept of Rényi dimensions was developed from the Rényi entropy first proposed by Alfréd Rényi (Rényi (1961)). Jizba and Arimitsu (2004) demonstrated the relationship between the Rényi entropy and Rényi function. They also showed the connection between the Rényi entropy and the multifractal spectrum and used the Rényi entropy to study the statistical properties of multifractal processes. Further, Harte (2001) indicated that the fractal dimension estimates in empirical investigations are generally the estimates of Rényi dimensions.

In multifractal analysis, the Rényi function can be used to determine the multifractal behaviour of stochastic processes and random fields. It quantifies the variation of surface or trajectory characteristics along with the change in the box size of an image, resolution or scale. It is also known by the terms, deterministic partition function, spectrum of scaling exponents and moment-scaling function. Multifractal formalism was introduced by Parisi and Frisch (1985). In multifractal formalism, the Rényi function is related to the multifractal spectrum via a Legendre transformation (Evertsz and Mandelbrot

(1992), Harte (2001) and Jizba and Arimitsu (2004)). The Rényi functions of multifractal processes usually exhibit a non-linear/parabolic shape whereas those of non-fractal and monofractal processes exhibit an almost constant or a linear behaviour (Grahovac and Leonenko (2018)). Olsen (2002) considered self-similar multifractals and studied their Rényi functions. To distinguish the multifractal behaviour of the infinite products of homogeneous and isotropic random fields, Leonenko and Shieh (2013) determined the Rényi function for three types of multifractal spherical random fields. They derived the Rényi functions for the log-normal model, log-gamma model and log-negative-inverse-gamma model.

Multifractal approaches have been applied widely, especially in physics. In the one-dimensional time series case, Telesca et al. (2015) conducted a multifractal analysis of the gravity time series of Earth by employing multifractal detrended fluctuation analysis and power spectrum methods. Their results revealed a significant multifractality in the Earth's gravity time series which rely on long-range correlations. Caniego et al. (2005) investigated the spatial variability of soil by utilizing the multifractal spectrum and the Rényi spectrum. The results of both methods suggested monofractal and multifractal scaling behaviours in two short and long transects respectively along which the soil properties are measured. Ahmad et al. (2014) performed a multifractal analysis of nucleus collisions. The obtained multifractal spectrum depicted a concave down behaviour and suggested the multifractal presence in the multi-particle production in those collisions.

Multifractal stochastic processes were investigated by Calvet et al. (1997), Calvet and Fisher (2002), Fisher et al. (1997) and Mandelbrot et al. (1997), where they defined the scaling properties using process moments. Angulo and Esquivel (2015), Grahovac and Leonenko (2014) and Grahovac and Leonenko (2018) studied the multifractal behaviour of stochastic processes. Grahovac and Leonenko (2018) introduced some bounds for the support of the spectrum of multifractalities. They emphasized that self-similar processes could yield a multifractal spectrum because of the infinite order moments of a positive kind. They showed that the partition function detects the divergence of moments which is important in determining the multifractal spectrum. These results were used to introduce a robust version of the partition function. Jaffard (1999) showed the multifractal nature of Lévy processes apart from Brownian motion and Poisson processes which are considered to be monofractals. Moreover, he demonstrated that the singularity spectrum of Lévy processes exhibits a linear behaviour.

The multifractal products of stochastic processes were extensively studied in the publications by Anh et al. (2008), Kahane (1985), Kahane (1987) and Mannersalo et al. (2002). Mannersalo et al. (2002) studied the multifractal products of stochastic processes with the motivation of exploring new teletraffic models. They introduced random multifractal measure constructions established on a class of T-martingales and derived the Rényi functions based on them. They mainly investigated some exponential classes of distributions but their study was limited to one-dimensional processes. Then, Anh et al. (2008) studied the multifractal products of geometric Ornstein–Uhlenbeck type stochastic processes that are determined by the Lévy motion. For the background driving Lévy processes, they utilized five distributions namely the gamma, variance gamma, inverse Gaussian, normal inverse Gaussian and generalized-z (Meixner) that are infinitely divisible. They obtained the limiting processes for exponents of these distributions and examined their dependence structure. They derived Rényi functions and demonstrated the multifractal behaviour of the resulting log-gamma, log-variance gamma, log-inverse Gaussian, log-normal inverse Gaussian and log-generalized-z (log-Meixner) distributions.

Ruiz-Medina et al. (2008) established two classes of space-time random field models correspondingly in continuous and discrete space time that exhibit multifractal behaviour spatially. Further, they analyzed the spatial multifractal characteristics of the proposed models. Koenig and Chainais (2008) proposed a spherical multifractal analysis using the continuous spherical wavelet transforms to investigate the self-similar properties of spherical data. They conducted a numerical spherical multifractal analysis and used a multifractal spherical texture model to test their methodology.

1.3 Long-range dependence

The empirical phenomenon of long-range dependence was first observed by Hurst (1951) in his studies of the hydrological characteristics of the Nile River. Hurst computed the rescaled adjusted range (R/S) statistics of the Nile River data. He found that the empirical growth rate of the R/S statistics is approximately n^H where $H = 0.72$ and n is the number of observations, known as the Hurst effect. However, his results disagreed with short-range dependent stochastic processes based on iid Gaussian random variables as the R/S statistics of those processes have the asymptotic behaviour similar to constant $\times\, n^{1/2}$. The hydrological models contradicted the empirical verifications. Subsequently, in an attempt to find a suitable model for the Hurst effect, the notion of long-range dependence was developed by Benoit Mandelbrot and his co-authors in their publications, Mandelbrot (1965), Mandelbrot and Van Ness (1968) and Mandelbrot and Wallis (1968). They introduced the concept of self-similar processes, fractional Brownian motion (FBM) and fractional Gaussian noise (FGN). The self-similar processes exhibit similar statistical characteristics on all scales. For a self-similar stochastic process $X(t)$, it holds, $X(ct) \stackrel{d}{=} c^H X(t)$ where H, $0 < H < 1$, is the self-similarity/Hurst parameter. The FBM is a self-similar, continuous-time stochastic process and it is a generalized version of the ordinary Brownian motion. Moreover, they showed that the FGN (the discrete increment process of the FBM) with the Hurst parameter, $H \in (1/2, 1)$ can be used to model the Hurst effect (Beran (1994)). These significant results were the foremost ones that modelled the Hurst effect by using a stationary Gaussian stochastic process. For more details regarding the chronological record of the phenomenon of long-range dependence, see Beran (1994), Graves et al. (2017) and Samorodnitsky (2007).

In many natural phenomena, the independence assumption of observations is often an approximation and the dependence between observations is inevitable due to the existence of serial correlations that are slowly decaying. Thus, long-range dependent processes received increasing attention in the last two decades. The long-range dependence is also known as long-memory, long-term persistence and the correlation structure of the singularities. The correlation function of a process with long-range dependence typically demonstrates a power-like decay and it is slower than an exponential decay. The Hurst parameter, H is used to quantify the long-range dependence. For long-range dependent processes $H \in (1/2, 1)$. Traditionally, two parametric models were introduced to model long-range dependent stochastic processes, namely, the FBM model and fractional autoregressive integrated moving average (fractional ARIMA/FARIMA) or ARFIMA model suggested by Granger and Joyeux (1980) and Hosking (1981).

The phenomenon of long-range dependence was found in both time series and spatial data and in numerous fields such as astronomy (Jeffreys (1939), Pearson (1902)), agronomy (Smith (1938)), climatology (Franzke (2010)), cosmology (Anh et al. (2018)), economics (Carlin and Dempster (1989)), geosciences (Montillet and Yu (2015)), hydrology (Lawrance and Kottegoda (1977)), internet traffic modelling (Karagiannis et al. (2004)) and many others, see Beran et al. (2013), Doukhan et al. (2003) and the references therein. Robinson (2020) discussed statistical inference methods and modelling approaches for spatial data that exhibit long-range dependence. The long-range behaviour of processes and fields has been found to be associated with their fractal behaviour and the scaling properties. The concept of long-memory concerns spectral singularities of stochastic processes and random fields, see Leonenko (1999).

There are several inequivalent definitions of long-range dependent processes, see Beran (1994), Doukhan et al. (2003), Pipiras and Taqqu (2017) and Samorodnitsky (2007). The majority of these definitions are established on the second-order characteristics of stochastic processes and random fields, such as the asymptotic behaviour of the covariance function at the infinity and spectral density at the origin or at another point or variance of partial sums. In the literature, the most frequent definition of long-memory is the existence of a hyperbolic-type decaying non-integrable covariance function. The relationship between the properties of long-range dependent processes in temporal and spectral domains is given by the well-known Tauberian and Abelian theorems. They link the asymptotic behaviour of the covariance function at the infinity and the singularity characteristics of the spectral density at zero, see Klykavka et al. (2012), Olenko (2006) and Olenko (2013).

Various statistical methods and results have been established to address the problems arising within the phenomenon of long-range dependence, see Beran (1992), Beran (1994), Giraitis et al. (2012), Pipiras and Taqqu (2017) and Samorodnitsky (2007). Some of these are: distinguishing the long-memory behaviour of data, the estimation of the Hurst parameter, the statistical inference of long-range dependence, the study of limit theorems of long-range dependence, spectral estimation and simulating processes with long-range dependence. Beran (1992) reviewed and addressed diverse statistical methods for long-range dependent data. The monograph by Beran (1994) studied various methodologies for the statistical inference of long-memory processes with real-life applications. He demonstrated limit theorems for the statistical inference of long-memory processes, specifically those for simple sums, quadratic forms and Fourier transforms. He also proposed three estimation methods for the Hurst parameter: heuristic approaches, time domain and frequency domain maximum likelihood estimation (MLE) techniques. The suggested heuristic approaches comprise the rescale range (R/S) statistics initially introduced by Hurst (1951), variance plots, log-log correlogram, semivariogram and log-log plots. Moreover, the least squares regression methods were proposed for the spectral domain as another important heuristic approach. These heuristic approaches are best considered as exploratory tools and are less efficient for statistical inference compared to MLE methods. Also, deriving simple confidence intervals for these methods is a difficult task. Other important problems such as the robust estimation of location and scale parameters in long-memory, forecasting, regression modelling and goodness of fit tests were also discussed in Beran (1994).

The estimation of the long-range dependence parameter, H, in the discrete setting was of growing interest in the recent past. It was extensively studied by many authors employing three major methodologies, namely the Gaussian parametric estimation/Gauss-Whittle contrast function (Fox and Taqqu (1986), Gao and Anh (1999), Gao et al. (2001), Heyde and Gay (1989), Heyde and Gay (1993) and Yajima (1985)), log-periodogram linear regression (Geweke and Porter-Hudak (1983), Hurvich and Beltrao (1993), Robinson (1994c), Robinson (1994b) and Robinson (1995b)) and Gaussian semiparametric/local Whittle estimation (Robinson (1995a)). The parametric estimation approach has drawbacks as the misspecification of the parametric model can result in inconsistent estimators of the location and memory parameters. Therefore, the latter two methods which are semiparametric approaches are better in many circumstances though they have slower convergence rates than the parametric ones. Yajima (1985) studied the parameter estimation of long-memory time series models introduced by Granger and Joyeux (1980) and Hosking (1981) by utilizing MLE and least squares estimation approaches.

Robinson (1995a) suggested a semiparametric Gaussian estimate for H. The proposed estimator was consistent and asymptotically normal. However, all these techniques were limited to the discrete case. Anh et al. (1999) introduced the fractional Riesz-Bessel motion (FRBM) which is an extension and a stationary counterpart of the classical FBM to model fractional random fields with long-range dependence. The FRBM models exhibited both long-range dependence and second-order intermittency properties simultaneously. The parameter estimation of FRBM models was studied in Gao et al. (2001). They proposed a novel estimation method for the parameter estimation of long-range dependence and second-order intermittency. For this, they utilized a continuous form of the Gauss-Whittle contrast function. The suggested estimators were asymptotically normal and consistent.

In the statistical inference of long-range dependent processes, the minimum contrast estimator (MCE) technique played an important role. Anh et al. (2004) studied statistical inference based on the higher-order spectra of random fields with probable long-range dependence in the frequency domain. They proposed a quasi-likelihood or, in other terms, an MCE method that was established from the information associated with the spectral densities of higher order. They included data tapering into the estimation method to eliminate possible bias caused by the edge effects and showed the consistency and the asymptotic normality of the consequent estimators. Similar results were obtained by Anh et al. (2007a) and Anh et al. (2007b). Anh et al. (2007a) examined the consistency and asymptotic normality of the proposed group of MCEs for short- and long-range dependent stochastic processes that were based on cumulant spectra of second- and third-orders. Espejo et al. (2015) introduced a class of spatial long-memory models and showed the consistency and asymptotic normality of the suggested MCEs for the long-memory parameters of Gegenbauer random fields. Further information regarding the present status of the theory of the MCE approach and its application for long-range dependent random processes can be found in Alomari et al. (2017) and the references therein. The asymptotic properties of long-range dependent random fields and their functionals have been extensively studied by Alodat and Olenko (2018), Alodat and Olenko (2020) and Alodat et al. (2020).

Anh et al. (2019b) considered long-range dependent spherical random fields with increasing radii and investigated the asymptotic behaviour of regression parameter estimators involved in the least squares estimation method of linear regression. For the least squares estimator, they derived the limiting distribution and its convergence rate and verified the results obtained through simulation studies. Li et al. (2020b) extended the framework of functional data analysis to processes with long-range dependence. For the temporal sum of observations of functional long-range dependent time-series, they developed the central limit theorem. They also estimated the Hurst parameter by employing a semiparametric R/S technique. Due to the slower convergence rates and lower efficiency of the previous estimation method, Li et al. (2020a) extended the results by utilizing a semiparametric local Whittle estimation procedure. They considered long-range dependent functional time series that are stationary and studied the estimation of the associated memory parameter. They demonstrated the consistency and the asymptotic normality of the resultant estimator.

1.4 Multifractionality and the Hölder exponent

The notion of multifractionality originated from fractionality. Initially, Kolmogorov (1940) presented the concept of the fractional Brownian motion (FBM) as a generalization of the classical Brownian motion. Subsequently, the FBM was analytically defined by Mandelbrot and Van Ness (1968) and emphasized its importance in real-world problems. The FBM is characterized by the Hurst parameter H where $H \in (0, 1)$. The pointwise Hölder exponent was used to determine the regularity of the paths of a stochastic process. The Hölder regularity and correspondingly, the pointwise Hölder exponent of the FBM can be specified by its Hurst parameter H (Ayache and Véhel (2004)). For the FBM, it is a constant so that the regularity does not change along the paths. The FBM was considered a mathematical model in applications related to fields such as hydrology, finance, image processing, internet traffic modelling, physics, turbulence, to name a few. The monograph by Doukhan et al. (2003) methodically investigated the FBM and its various applications. Biagini et al. (2008) discussed the stochastic calculus theory for the FBM with applications. Cohen and Istas (2013) provided an impressive deliberation on fractional Brownian fields.

However, the FBM processes had limited practical applications in many situations since they didn't conform to real data. In particular, the FBM was not suitable to model natural phenomena that had variable Hölder exponents. Thus, the multifractional Brownian motion (MBM) was proposed as an alternative flexible stochastic model to overcome such limitations. The MBM is an extension of the FBM where the Hurst parameter H changes over time. The pointwise Hölder exponent of the MBM can be specified by its Hölder function $H(t)$ (Ayache and Véhel (2004)). The MBM was defined in two different ways. Firstly, Peltier and Véhel (1995) presented the MBM by incorporating an integral representation of the FBM in the time-domain whereas secondly, Benassi et al. (1998a) illustrated the MBM by utilizing an integral representation of the FBM in the frequency-domain. Stoev and Taqqu (2006) considered a family of MBM processes which included these two representations and studied the connection between them. Multifractionality concerns the fractal properties of the process: self-similarity properties and the roughness of the trajectories. The multifractional processes found ample applications in modelling many complex phenomena, see Bianchi and Pianese (2007) and Sheng et al. (2011). The classical monograph by Ayache (2018) provided a comprehensive discussion on multifractional stochastic fields and their wavelet decompositions. It also provided excellent studies of the Hölder exponents that determine the roughness of paths in pointwise, local and global settings.

The Hölder exponent is a numerical measure of roughness. It is used to distinguish the Hölder regularity of erratic functions locally. It was defined as H, $H \in (0, 1)$, through a Hölder condition by Mallat (1998) and Muzy et al. (1994). As the value of H gets closer to zero, the regularity of the function decreases. Conversely, the smoothness of the process increases along with the increasing Hölder exponent. For continuous functions, their family of pointwise Hölder exponents were thoroughly studied by Andersson (1997), Daoudi et al. (1998) and Jaffard (1995). They demonstrated that this family is the same as the lower limits of series of functions which are continuous and non-negative. Traditionally, the quadratic variations approach was used frequently for the estimation of the pointwise Hölder exponents, see Benassi et al. (1998a), Benassi et al. (1998b), Benassi et al. (2000), Coeurjolly (2001), Istas and Lang (1997a), and Kent and Wood (1997).

Many research studies focused on the estimation of multifractional processes to comprehend and utilize them better. Benassi et al. (1998a) studied the estimation of the MBM processes and obtained an estimator for the Hurst functions which are continuously differentiable. They considered a local version of the global approach used for the estimation of the FBM processes and proved the consistency of the proposed estimator. These results were extended by Coeurjolly (2005) in which they studied the Hurst functions that are Hölderian. They proposed a method of local quadratic variations for the estimation of the MBM, in particular a moments approach and showed the asymptotic normality and the consistency of the estimates. Benassi et al. (2000) and Ayache et al. (2005) investigated the estimation of multifractional processes that possess discontinuous pointwise Hölder exponents and used a generalized quadratic variations approach in each of these studies. Benassi et al. (2000) studied the estimation of the step fractional Brownian motion. Ayache et al. (2005) proposed a family of multifractional processes, generalized multifractional Gaussian processes, and explored the estimation of the erratic multifractional function.

Ayache and Taqqu (2005) examined another development of the MBM, multifractional processes with a random exponent, where a stochastic process was substituted for the Hurst function. Loboda et al. (2021) studied the regularity properties of such multifractional processes with a random exponent using a probabilistic approach. The obtained results were valid for a wide class of moving average processes, in particular for multifractional Matérn processes. Ayache and Jaffard (2010) considered discontinuous functions and studied their family of pointwise Hölder exponents. They showed that these Hölder exponents can also be stated in terms of the lim infs of series of continuous functions. Bardet and Surgailis (2013) proposed a nonparametric local estimation method for multifractional Gaussian processes. They used the increment ratio statistic to estimate the local Hurst function and developed the multidimensional asymptotic theory for the proposed estimator. Another global estimation methodology established on a ratio statistic was presented by Lebovits and Podolskij (2017) for the global regularity of the MBM.

The MBM has a major drawback as it relies on the mandatory requirement of the continuous Hölder regularity in their sample paths. In practice, the pointwise Hölder exponent changes extensively in many applications, for example in finance, medical imaging, turbulence and in telecommunications. Therefore, when the Hölder function is irregular over time, the generalized multifractional Brownian motion (GMBM) was employed as a good candidate to model such processes, extending the FBM and MBM. The GMBM was introduced by Ayache and Véhel (1999) and Ayache and Véhel (2000). It was characterized by using a functional parameter $H(t)$ which belonged to a wider class of Hölder functions. It was shown in Ayache and Véhel (1999) and Ayache and Véhel (2000) that under some mild conditions, the pointwise Hölder exponent of the GMBM coincided with its functional Hurst parameter. One of the major advantages of the GMBM is that it can be used to model processes which exhibit both long-range dependence and erratic behaviour.

The generalized multifractional field was introduced by Ayache (2002) to study the GMBM. It was represented by using a wavelet series expansion. The paper derived some important properties that enabled the GMBM to be examined, in particular to identify its pointwise Hölder exponent. Ayache and Véhel (2004) considered the GMBM and studied the estimation of its pointwise Hölder exponent. They utilized a generalized quadratic variations approach and proposed two estimators for the pointwise Hölder exponent. Both the estimators were strongly consistent. Chapter 4 of this thesis develops a suggested estimation methodology for multifractional spherical random fields. Another methodology which generalized the MBM to obtain a very irregular Hölder function was suggested by Herbin (2006). He introduced two classes of multiparameter developments of the MBM: multifractional Brownian field and multifractional Brownian sheet. It was shown that the derived multifractional Brownian field coincided with the results of Benassi et al. (1997).

Over recent years, fractional and multifractional paradigms were developed in multidimensional settings. Richard (2018) considered Hölder random fields that include fractional Brownian fields and examined the directional properties of the Hölder regularity of their trajectories. D'Ovidio et al. (2016) studied the fractional spherical stochastic fields: spherical fields that are dependent on space-time locations and driven by various stochastic differential equations. Anh et al. (2018) investigated the fractional SPDEs defined on the unit sphere with cosmological applications. They examined the fractionality in the derivatives. The utilized fractional SPDE models showed long-range dependent behaviour. The conceptual multifractional space-time models were examined by Calcagni et al. (2016) and Calcagni (2019) in a detailed manner. They indicated that the expansion of the universe corresponds to multifractional behaviour. Calcagni et al. (2016) considered the multi-fractional models that possess q-derivatives and provided numerical studies using CMB data from the Far Infrared Absolute Spectrophotometer and Planck mission. Calcagni (2021) presented some theoretical perspectives of multifractional models and discussed their applications in the fields of cosmology and quantum gravity.

1.5 Cyclic long-memory processes

In reality, numerous time series often exhibit fluctuations with periodic behaviour/seasonal variation. The phenomenon of long-range dependence was also observed in such cyclic and seasonal processes, see Hassler and Wolters (1995), Montanari et al. (2000), Porter-Hudak (1990) and Ray (1993). Hence, research on cyclic long-memory processes received enormous attention over the past years, see Arteche and Robinson (2000), Alomari et al. (2017), del Barrio Castro and Rachinger (2021), Hidalgo and Soulier (2004), Whitcher (2004) and the references therein. Arteche and Robinson (1999) presented an excellent discussion on the application, estimation and statistical inference of seasonal and cyclic long-memory processes. In the literature, long-memory processes with spectral singularities at the origin is a comprehensively studied research area, see Dahlhaus (1989), Fox and Taqqu (1986), Giraitis and Surgailis (1990), Heyde and Gay (1993) and Yajima (1985). However, spectral singularities at unknown poles that characterize cyclic and seasonal long-memory processes are less studied. Traditionally, cyclic processes are modelled by the two major models: an ARMA model that possesses a non-zero spectral peak and a non-random trend with a stationary random noise. In the first case, the cyclical behaviour fades with time whereas in the second case, the cyclical feature remains fixed with time. Cyclic long-memory processes are an intermediate class that has a pole in their spectral densities, see Arteche and Robinson (1999).

Various statistical approaches have been historically developed for cyclic long-memory processes. Several parametric models for cyclic long-memory processes have been proposed by Anděl (1986), Hosking (1984) and Gray et al. (1989). Gray et al. (1989) proposed a stationary class of the long-memory processes which is a generalized version of the FARIMA models. These models were referred to as the Gegenbauer autoregressive moving-average (GARMA) models as they utilized the properties of the Gegenbauer polynomials' generating function in their construction. Porter-Hudak (1990) considered seasonal auto-regressive fractionally integrated moving-average (SARFIMA) models, also known as seasonal fractional ARIMA models, to study USA monetary aggregate data that exhibited seasonal long-memory behaviour. Hassler (1994) proposed flexible ARFISMA models. Chung (1996b) provided an in-depth analysis of the GARMA processes' long-memory characteristics. Woodward et al. (1998) suggested a k-factor GARMA model for long-memory processes that have multiple spectral poles. For cyclic long-range dependent random fields, the asymptotic properties of their functionals and the associated limit theorems were extensively studied by Espejo et al. (2014), Espejo et al. (2015), Ivanov et al. (2013), Klykavka et al. (2012) and Olenko (2013).

In practice, the parameters of seasonal/cyclic long-memory processes were often estimated in a two-step procedure. First, the spectral singularities were estimated and subsequently, a classical parametric method was employed to estimate the memory parameter. For Gegenbauer processes, Yajima (1996) proposed the initial estimation of the singularity parameter by the periodogram maximization. The statistical inference for the simultaneous estimation of the two parameters of cyclic long-range dependent time series is the main interest of this thesis; see Chapter 5.

Over recent years, diverse strategies were developed for the parameter estimation of cyclic-long memory processes with one non-zero spectral singularity. Chung (1996a) studied the parametric estimation of Gegenbauer processes with two parameters by employing a conditional sum of squares estimation strategy in the time-domain. Giraitis et al. (2001) introduced a parametric joint estimation method for the unknown spectral singularity and memory parameters of cyclic long-memory processes in the frequency-domain. It utilized a maximization of Whittle approximation of the Gaussian likelihood. They showed that the asymptotic behaviour of the Whittle estimates remains the same, irrespective of the knowledge of singularity location. Though they proved the consistency of the estimates of singularity parameter, its asymptotic normality and the optimal rates of convergence remained undetermined. In comparison, Hidalgo and Soulier (2004) studied the semiparametric estimation of the hidden location and memory parameters in a two-phase strategy and showed that the convergence rates of the estimators are optimal. They utilized the periodogram maximization technique of Yajima (1996) for the estimation of the location parameter and a modification of the GPH estimator of Geweke and Porter-Hudak (1983) for the long-memory parameter. They also showed that the statistical properties of the estimator of long-memory parameter are unaffected by the inadequate knowledge of the location parameter.

Hidalgo (2005) considered a consistent nonparametric estimator of the location of the spectral pole and its limiting behaviour was derived. He suggested a two-phase semiparametric estimator for the long-memory parameter and studied its asymptotic behaviour. The results suggested that the convergence rate and the limiting distribution of the estimate of the memory parameter were not affected by the singularity parameter estimation. Artiach and Arteche (2011) proposed an iterative algorithm procedure established on the periodogram maximizer method for the parameter estimation of the unknown non-zero frequency of cyclic long-memory processes. It outperformed the traditional techniques limited to Fourier frequencies, in particular to that of Hidalgo and Soulier (2004). The obtained theoretical findings were demonstrated with applications to the famous sunspot data and determination of the business cycle of the US unemployment increments. Dissanayake et al. (2016) introduced a quasi-likelihood approach employing the Kalman filter for the parameter estimation of Gegenbauer long-memory processes.

Although the parameter estimation of cyclic long-memory processes has resulted in a proliferation of research, the simultaneous estimation of both long-memory and singularity parameters is a challenging problem. Notably, less is known regarding the inferential statistics of the proposed estimators. Alomari et al. (2020) considered Gegenbauer-type cyclic long-memory processes. They proposed a semiparametric estimation method in which novel simultaneous estimators were introduced for the long-memory and singularity location parameters. For the estimation procedure, they utilized the generalized filtered method of moments technique which was established on general filter transforms. Chapter 5 of this thesis extends the results obtained by Alomari et al. (2020) and studies the asymptotic properties of the introduced simultaneous estimators for cyclic long-memory processes. Beaumont and Smallwood (2019) examined the asymptotic properties of the time- and frequency-domain likelihood-based estimators of the Gegenbauer processes, correspondingly conditional sum of squares and Whittle estimators. By utilizing a Monte Carlo analysis, the substantial efficacy of the two types of estimators was demonstrated and supported the distribution theory suggested by Chung (1996a) and Chung (1996b) for the cycle length ruling parameter.

Several researchers investigated cyclic long-memory processes with multiple spectral singularities and developed parametric and semiparametric estimation techniques for them, see Arteche and Robinson (2000), Arteche (2020), del Barrio Castro and Rachinger (2021), Chan and Tsai (2012), Ferrara and Guégan (2001) Giraitis and Leipus (1995), Hunt et al. (2021), Robinson (1994a) and Whitcher (2004). Arteche and Robinson (2000) introduced seasonal or cyclic processes that consist of asymmetric long-memory characteristics. They considered a semiparametric inference approach for cyclic long-memory processes with one or more non-zero frequencies by extending the general semiparametric techniques such as log-periodogram and Gaussian/Whittle estimation methods. Three semiparametric tests of spectral symmetry were provided and empirical studies with application to UK inflation data were presented. Arteche and Velasco (2005) extended the results of Arteche and Robinson (2000) by considering the tapering and trimming of the proposed semiparametric estimates. They showed that data tapering diminishes the bias made by asymmetrical spectral densities around cyclical frequencies.

Ferrara and Guégan (2001) studied the parameter estimation of k-factor Gegenbauer processes. They presented two types of estimation techniques, namely a semiparametric approach and a quasi-likelihood approach for the long-memory parameter of cyclic long-memory processes. These techniques were established on the log-periodogram and the Whittle-likelihood, respectively. The obtained results were demonstrated by applications to Nikkei spot index data. Palma and Chan (2005) proved the asymptotic normality and consistency of the exact MLE method for Gaussian seasonal long-memory processes and demonstrated applications to internet traffic data that displayed multiple spectral singularities. Chan and Tsai (2012) studied the asymptotic properties of an aggregated SARFIMA process and the MLE method of its limiting spectral density. The obtained results were demonstrated with application to internet traffic data that consisted of multiple spectral peaks in its periodogram. Arteche (2020) introduced a generalized form of the exact local Whittle estimator suggested by Shimotsu and Phillips (2005) for the joint estimation of the memory parameters in seasonal and cyclic long-memory processes with multiple spectral poles. Hunt et al. (2021) suggested a novel estimation method for the parameter estimation of k-factor Gegenbauer processes which was motivated by Whittle's methodology. It was established on a nonlinear least-squares regression method in the frequency-domain.

1.6 Wavelet-based estimation of parameters of stochastic processes

Wavelets have been applied in the identification of abrupt changes and in the study of the local behaviour of time series. The theory of wavelets emerged in the 1980s. Farge (1992), Meyer (1992) and Meyer (1993) provide an extensive discussion on the history of wavelets, wavelet transforms and their applications in cosmology, image processing, numerical analysis, signal processing, turbulence, to name a few. Fourier analysis is ideally suited to stationary time series analysis. In contrast, wavelets are efficient, simpler and capable of handling the non-stationarities in signals due to the fact that they are localized in time/space. Wavelets have enabled time series analysis both in scale and time. There are two main classes of wavelet transforms used in the literature. They are continuous wavelet transform (CWT) and discrete wavelet transform (DWT). A wavelet sequence is constructed by means of dilations and translations of the mother wavelet. A plethora of mother wavelet forms has been introduced such as Haar, Mexican hat, Meyer, Shannon and Morlet. Numerous R packages have been developed to accommodate such wavelets, for instance, MassSpecWavelet (Du et al. (2006)), mwaved (Wishart (2019)), Rwave (Carmona and Torresani (2021)), waveslim (Whitcher (2020)) and wavethresh (Nason (2016)).

Wavelet analysis was well developed in recent years. It plays a key role in the spectral analysis of time series and time-dependent data, see Beran and Shumeyko (2012), Cazelles et al. (2007), Chesneau et al. (2019), Chiann and Morettin (1998), Cornish et al. (2006), Nason and Sachs (1999), Percival and Walden (2000), Priestley (1996), Shirazi and Doosti (2015) and the references therein. The monograph by Ogden (1997) provided the basic theory of wavelets and addressed its statistical applications. For stationary discrete processes, Chiann and Morettin (1998) developed the wavelet-based spectral analysis. They introduced the wavelet spectrum and derived the large sample properties of the DWT of sample values. The wavelet periodogram was utilized as the wavelet spectrum estimator. The classical monograph by Percival and Walden (2000) introduced the wavelet approach for the spectral analysis of time series. It was mainly focused on the DWT, the maximal overlap DWT and the discrete wavelet packet transform with applications to real data. It also discussed various facets of wavelet theory, in particular, wavelet variance, the study of long-memory processes using wavelets and wavelet-based estimation of signals. Serroukh et al. (2000) studied the time-scale characteristics of time series using a wavelet variance estimator. Cazelles et al. (2007) discussed the wavelet-based approach for the spectral analysis of time-dependent time series from epidemiological studies. Guerrier et al. (2021) proposed a robust wavelet-based approach for the statistical inference of latent time series models. They utilized the generalized method of wavelet moments technique to derive the asymptotic properties of the robust estimators.

In recent years, wavelet-based estimation techniques have been instrumental in the parameter estimation of long-memory processes. This is due to the fact that the DWT of long-memory time series results in wavelet coefficients that have almost no correlation within and between each of the levels. Moulines et al. (2007) studied the estimation of the memory parameter of long-memory time series models by employing a wavelet-based semiparametric methodology. They used the log-regression estimation method proposed by Abry and Veitch (1998). They considered observations in discrete-time and their wavelet coefficients, derived their spectral density, and studied the asymptotic approximations. Another wavelet-based approach was considered by Moulines et al. (2008) for the long-memory parameter estimation by utilizing a semiparametric local Whittle estimation method. They derived the asymptotic normality of the estimator in the case of a Gaussian process and its consistency and rate optimality under the condition that it's a linear process. For the trend-function estimation of long-range dependent time series models, Beran and Shumeyko (2012) proposed a data-adaptive wavelet-based technique.

Whitcher (2004) considered stationary seasonal long-memory time series models with one or more spectral singularities and investigated how wavelet transforms can be used for their estimation. He studied the discrete wavelet packet transform and its maximal overlap version for the analysis of seasonal long-memory time series. He proposed an approximate MLE approach which was established using the wavelets. The results suggested that the performance of the wavelet-based method was similar to the Whittle likelihood approach under the condition of a parametric-type spectrum for the likelihood construction. Similarly, an alternative wavelet-based approximate MLE method was presented by Boubaker (2015) for the parameter estimation of stationary k-factor Gegenbauer processes. According to the obtained results, the new estimator achieved an improved performance in many settings compared to Whitcher (2004) and classical Whittle estimators. Then, for non-stationary seasonal long-memory processes, Lu and Guegan (2011) proposed a novel class of stochastic models, a k-factor Gegenbauer process that is locally stationary. They implemented a wavelet approach for the parameter estimation of time-varying memory parameters and proved the consistency of the parameter estimates. The proposed wavelet algorithm was demonstrated with applications to Nikkei stock average data. Analogously, Boubaker and Sghaier (2015) introduced a novel semiparametric family of seasonal long-memory (SEMIGARMA) models extending the classical GARMA models. The corresponding class of models permitted the simultaneous presence of a nonlinear stochastic trend and seasonal time-dependence. They incorporated the wavelet-based estimation method suggested by Whitcher (2004) for the parameter estimation of the novel k-factor GARMA-FIAPARCH model with applications to time series data from MENA stock markets.

Wavelet-based estimation techniques have been popular among parameter estimation strategies for different types of stochastic processes. For example, Abry and Didier (2018) studied the wavelet analysis of the vector-valued counterpart of the FBM process: operator FBM. They suggested a wavelet-based approach (wavelet spectrum eigenstructure) for the parameter estimation of the matrix Hurst exponent of the bivariate operator FBM process. In addition, Abry et al. (2019) examined the semiparametric two-phase estimation of the Gaussian mixed fractional processes. They employed a wavelet-based methodology for the estimation of memory parameters and the demixing matrix. In addition, the asymptotic normality of the proposed estimators was derived in discrete-time as well as in continuous-time. Boniece et al. (2021) considered a wavelet-based estimation technique for the parameter estimation of the tempered FBM. They used wavelets to develop a computationally powerful hypothesis test to compare the FBM model with the alternative tempered FBM model. They also determined the asymptotic normality and the consistency of the wavelet-based estimator.

Furthermore, wavelet methods played an important role in the study and series representation of fractional and multifractional processes, see Ayache and Taqqu (2005), Ayache et al. (2007), Ayache (2018), Ayache and Esmili (2020) and Ayache et al. (2020). Ayache and Taqqu (2005) considered wavelet series expansions of multifractional processes with a random exponent and investigated their self-similarity properties and Hölder regularity. Ayache et al. (2007) studied the wavelet-based representation of the generalized multifractional processes. The monograph by Ayache (2018) discussed the development of the wavelet series expansion of the MBM. It also addressed the application of the wavelet methodology in the path regularity identification of the multifractional stochastic fields. Ayache and Esmili (2020) studied the generalized Rosenblatt process and its series representation using wavelets. Analogously, Ayache et al. (2020) established a wavelet approach for the expansion of harmonizable fractional stable sheets.

1.7 Cosmological background

The universe originated about 14 billion years ago. The cosmic microwave background radiation (CMB) eventuated about 380,000 years after the Big Bang. After its discovery it was thought to be the remnant radiation from the Big Bang that fills the universe as the background. Its unforeseen discovery was made by Arno Penzias and Robert Wilson who were American radio astronomers. It was done in 1964 and they won the Nobel prize in Physics in 1978, see Penzias and Wilson (1965). The CMB cannot be detected by usual telescopes. Its temperature is approximately 2.73° above absolute zero (2.73 kelvin). It consists of a temperature and an electromagnetic spectrum and is best observable in the microwave part of the electromagnetic spectrum, see NASA (2021). The CMB is detectable only through a far-infrared or a radio telescope. It is the primordial source of information left to understand the origin of the early universe.

Earlier, the universe was in an extremely hot and dense state. During this time, the universe comprised an opaque composition of a plasma/ionised gas. Therefore, the atoms presented at that time couldn't reconcile and were broken down into protons and electrons. After the Big Bang, the universe underwent accelerated inflation and expansion and still continues to expand at present. Along with the expansion, the universe started to cool down. Thus, the recombination of atoms was possible about 380,000 years after the Big Bang. The first hydrogen atoms were formed once the free electrons, scattered through the universe, became trapped in the orbits of the atomic nuclei, making them stable. As the electrons in high-energy levels started falling down on lower energy levels, photons were emitted and the first light was radiated, see Castelvecchi (2019). This ancient light in the cosmic history is termed as the CMB and this occurrence is known as the recombination/ epoch of re-ionization. After the recombination era, other atoms such as oxygen, carbon and iron formed within the hearts of stars.

For the very first time, in the late 1940s, George Gamow, Ralph Alpher, and Robert Herman suggested the presence of CMB in the midst of their studies investigating the nucleosynthesis of light elements such as hydrogen, helium and lithium in the ancient universe (Alpher et al. (1948a) and Alpher et al. (1948b)). At this time, they understood that the universe must be hot for the combination of the nuclei of those light elements (Naselsky et al. (2006)). Also, they perceived that the CMB would be able to be seen in the present day with the spread of the residue radiation from the Big Bang throughout the universe. The most historic measurements of the CMB were made by Arno Penzias and Robert Wilson in 1964. At that time, although they verified the presence of the CMB, they weren't able to identify the CMB in a detailed manner.

Until now, the space missions, Cosmic Background Explorer (COBE), Wilkinson Microwave Anisotropy Probe (WMAP) and Planck have been voyaged to investigate the CMB. In 1989, NASA launched its first mission on CMB, namely the COBE (Efstathiou et al. (1992)). One of the major discoveries of the COBE mission was to verify that the CMB spectrum conforms to a blackbody spectrum with 2.73 kelvin temperature radiation (Smoot (2007)). Further, the COBE mission discovered that the CMB consists of slight temperature variations across the sky. They are also known as CMB anisotropies. In recognition of the significant explorations achieved by the COBE mission, John Mather and George Smoot were awarded the Nobel prize in Physics in 2006. Then, in 2001, NASA launched its second space mission on CMB, namely the WMAP (Spergel et al. (2003)). The sensitivity of the CMB measurements taken by the WMAP mission is 45 times and the angular resolution is 33 times that of the COBE mission (Smoot (2007)) which allowed the small temperature fluctuations to be measured more precisely. The WMAP mission resulted in the development of the standard cosmological model (Komatsu et al. (2009)).

The standard cosmological model is known as the ΛCDM (Lambda Cold Dark Matter) model (Robson (2019)). According to this model, the universe comprises two main characteristics, homogeneity and isotropy on a large scale. The property of homogeneity suggests that each part of the universe has approximately equivalent attributes and the isotropy feature implies that the properties of the universe are independent of the spatial directions. The standard cosmological model can be explained with a few parameters such as dark matter and dark energy, the ordinary matter's density, the universe geometry, the Hubble constant, etc, see Bjorken (2003) and Robson (2019) for more details. The universe consists of 68% dark energy, 27% dark matter and 5% ordinary matter (Ade et al. (2014) and NASA Science (2021)).

The Planck mission was launched in 2009 by the European Space Agency (Adam et al. (2016b) and Planck Science Team (2021)) with the main objective of finding evidence to validate and observe deviations from the established standard cosmological model.

In 2013, they released the highest precision CMB map from Planck CMB data (Ade et al. (2014)). The Planck mission captured the widest frequency range considering the CMB anisotropies of far-infrared and microwave regions with higher resolution than the previous space missions. The Planck CMB data have been observed at 5 arc minutes resolution on the CMB sky (Smoot (2007)). Figure 1.1 shows fractional variations of the CMB temperature that were observed by the Planck mission in 2009. It was obtained using the R package RCOSMO (Fryer et al. (2020) and Fryer et al. (2019)).

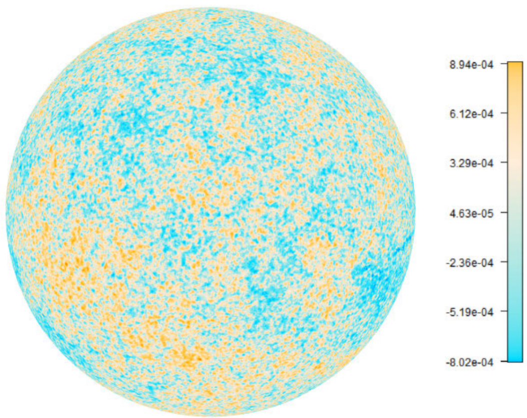


Figure 1.1: The CMB observed by the Planck mission

Planck enables astronomers to estimate the parameters that describe the early universe by investigating the CMB and developing various models of its changes over billions of years. Figure 1.2 depicts the CMB angular power spectrum distinguished by the Planck mission (European Space Agency and the Planck Collaboration (2021)). It comprises temperature fluctuations observed at different angular resolutions in decreasing order of magnitude. The sample values based on the Planck CMB observations are shown by the red dots. The optimum fit with the standard cosmological model is shown by the green curve and the shaded region depicts the cosmic variance. It is clear from Figure 1.2 that the data points conform to the standard model of cosmology at very small angular scales whereas they deviate from this in the large angular scales (say greater than 60 degrees). Also, a distinct outlier can be observed near the angular scales greater than 6 degrees which is an anomalous CMB observation. Thus, these findings question the validity of the standard cosmological model and in particular, demonstrate the importance of investigating the non-Gaussianity of the CMB data.

Several studies uncovered the new models underlying the CMB data, checked for the

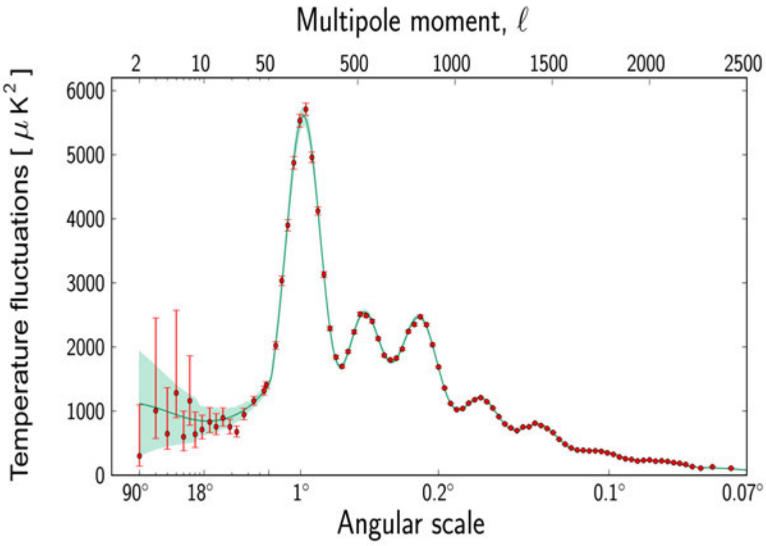


Figure 1.2: The CMB angular power spectrum observed by the Planck mission (Image credits: European Space Agency and the Planck Collaboration (2021))

Gaussianity of the CMB data and explored the outer solar system based on the CMB data. Hawking and Ellis (1968) investigated the presence of singularities in the universe, regarding the radiation of cosmic black-bodies and discussed the phenomenon of space-time singularities. They showed that under some specific assumptions, the existence of a singularity in the universe is manifested by the black-body radiation of 3 kelvin. Leclercq et al. (2014) showed the constraints on primitive non-Gaussianity with substantial examples from major analysis approaches applied to Planck 2013 data.

Traditionally, CMB data are modelled by the isotropic, Gaussian spherical random fields under the assumptions of the standard cosmological model. However, there have been issues and concerns regarding the statistical distribution of the CMB data, see Bartolo et al. (2010) and Marinucci (2004) for a detailed discussion. Various methodologies were used to investigate the non-Gaussianity of the CMB data and to identify departures from the standard cosmological model. For example, they considered the angular power spectrum (Durrer (2015)), bispectrum statistic (Ferreira et al. (1998), Hill (2018), Yadav et al. (2007)), trispectrum statistic (Munshi et al. (2011), Smidt et al. (2010)), wavelets (Barreiro et al. (2000), Cayón et al. (2001), Hobson et al. (1999), McEwen et al. (2005), McEwen et al. (2006), Pando et al. (1998), Starck et al. (2004)), Minkowski functionals (Buchert et al. (2017), Hikage et al. (2006), Hikage et al. (2008), Hikage et al. (2009), Novikov et al. (2000)), fractal analysis (Coleman and Pietronero (1992), Diego et al.

(1999), Leonenko and Shieh (2013)), hot and cold spot statistics (Chingangbam et al. (2012), Cruz et al. (2005), Larson and Wandelt (2004)), principal component analysis (Bromley and Tegmark (1999), Regan and Munshi (2015)) correlation functions (Kogut et al. (1996)), entropy methods (Minkov et al. (2019)), etc.

Of these, Durrer (2015) considered the CMB angular power spectrum to investigate the non-Gaussianity of the CMB. He explored the evolution of cosmological perturbations and their impacts, such as polarization and temperature fluctuations on the CMB. He mentioned that the oscillatory behaviour in the CMB angular power spectrum including the acoustic peaks demonstrate the primordial density fluctuations that originated after the Big Bang. These density fluctuations have been subjected to gravitational instability and resulted in the formation of large-scale structures, such as galaxies and galaxy clusters (de Bernardis et al. (2000)).

Bromley and Tegmark (1999) investigated the Gaussianity of the CMB data from the COBE mission using an eigenmode analysis approach. The series of tests that were conducted agreed with the Gaussianity of the CMB data. Further, they assessed the significance of the non-Gaussianity detected using the bispectrum and fourth-order wavelet statistics by Pando et al. (1998) and Ferreira et al. (1998). They stated that Gaussianity cannot be completely rejected from COBE data, and the non-Gaussian fields should be studied in a quantitative manner. Regan and Munshi (2015) performed a principal component analysis (PCA) in the framework of a skew-spectrum statistic to investigate the primary and secondary non-Gaussianities in the CMB data. They applied the techniques to a Planck-like data set and highlighted the importance of PCA analysis in estimating and distinguishing the different sources of non-Gaussianity.

Hobson et al. (1999) used wavelet transforms to detect the non-Gaussianity in CMB maps. They considered non-Gaussian CMB maps that resulted from cosmic strings together with the superimposed Gaussian signals. Their results surpassed the ones obtained by other techniques such as Minkowski functionals and methods based on quantifying temperature distribution moments. Subsequently, Barreiro et al. (2000) used spherical Haar wavelets and Cayón et al. (2001) utilized spherical Mexican Hat wavelets to investigate the non-Gaussianity of the CMB data from the COBE mission. Both of their studies found no strong evidence for the non-Gaussianity in the four-year COBE-DMR data.

Starck et al. (2004) employed multi-scale methods such as wavelet, curvelet and ridgelet transforms to determine non-Gaussian indications in the CMB. They performed a multi-scale analysis of three anisotropic CMB maps generated by the Gaussian, cosmic strings and the kinetic Sunyaev-Zel'dovich (SZ) effects where the latter two are non-Gaussian consequences. The bi-wavelet transform was able to more strongly distinguish the non-Gaussian signs of the CMB maps than the other two methods. Thus, the utilized techniques are useful to detect the non-Gaussian signatures of CMB maps due to a mixture of such effects.

McEwen et al. (2005) and McEwen et al. (2006) used directional spherical wavelets to examine non-Gaussianity in the CMB data from the WMAP mission. McEwen et al. (2005) detected significant departures from Gaussianity from the observed skewness and kurtosis statistics in the coefficients of the real Morlet wavelet and spherical elliptical Mexican hat wavelets. As the Bianchi correction in the WMAP maps removed many detected anomalies from the previous studies, McEwen et al. (2006) conducted a further analysis to confirm the results. The results confirmed the deviations from Gaussianity observed via the skewness statistics of spherical wavelet coefficients while they rejected the ones observed from the kurtosis statistic.

Amidst the employed multifractal methods, Diego et al. (1999) conducted a multifractal analysis using a partition-function-based approach to investigate the fractal nature and non-Gaussianity of the CMB data from the COBE mission. The results showed the absence of a fractal nature with no evidence of non-Gaussianity. Further, Leonenko and Shieh (2013) investigated the Rényi function which is an important technique in mutifractal analysis. They derived the Rényi functions for three multifractal models, namely the log-normal model, log-gamma model and log-negative-inverse-gamma model which can be used to study the non-Gaussianity of the CMB data.

Novikov et al. (2000) utilized Minkowski functionals and peak statistics to investigate the non-Gaussianity of CMB data from the four-year COBE mission. They computed the Minkowski functionals and peak statistics for a non-Gaussian model which is χ^2 distributed, a Gaussian model and for COBE data and compared the results. The results have been remarkably different with no significant evidence from COBE data that confirm either the Gaussian or non-Gaussian models. Similarly, Buchert et al. (2017) used Minkowski functionals to study the non-Gaussianity in Planck 2015 CMB maps. They performed comparisons using the Hermite and perturbative expansions for the Planck CMB data and simulated maps with ΛCDM model. They also disclosed weak evidence for non-Gaussianity in the Planck 2015 CMB data.

Chingangbam et al. (2012) used hot and cold spot statistics to discriminate the different types of non-Gaussian models. They considered simulated non-Gaussian models; one with primordial non-Gaussianity and another with temperature fluctuations according to a probability density function. They suggested that in addition to Minkowski functionals, hot and cold spot statistics are probes of non-Gaussianity in distinguishing the different types of non-Gaussianity in the CMB data.

Further, cosmic anomalies have been investigated as an indication of the non-Gaussian CMB and ample discussions and research have been conducted regarding them, see Adhikari et al. (2016), Copi et al. (2010), European Space Agency (2021c), Hamann et al. (2021) and Muir et al. (2018). Cosmic anomalies are vague features observed at some angular scales of the CMB sky sphere. It is difficult to explain the underlying reasons for such anomalies from the standard cosmological model. It is a known fact that the interference from the Milky Way obstructs the CMB near the galactic plane as the radio signals emitted from Milky Way are noisier than the CMB. Nevertheless, as these radio wave emissions have a predictable smooth spectrum, they can be deducted from the observed CMB spectrum to create a precise CMB map, see Castelvecchi (2019).

Adhikari et al. (2016), Copi et al. (2010) and Muir et al. (2018) investigated the large-scale anomalies in the CMB data. Of these, Adhikari et al. (2016) studied the expected CMB power asymmetry under the non-Gaussianity of the primordial temperature fluctuations. They considered the local non-Gaussianity and their numerical studies resulted in weak confirmation for the large-scale non-Gaussianity in the CMB data from WMAP and Planck missions. Hamann et al. (2021) proposed a direction-dependent probe (AC discrepancy approach) to check the Gaussianity of the CMB data with applications to Planck 2015 and 2018 data. They suggested that these probe coefficients should be independent Gaussian random variables for a given direction. They compared their results with the simulated isotropic Gaussian CMB maps and observed significant departures from Gaussianity for the inpainted Planck maps, particularly near the masked regions.

According to the latest findings of a study (Planck Collaboration et al. (2020)) which analyzed CMB polarization data, the Planck mission reveals no advanced verification for the underlying causes of the cosmic anomalies and still they remain an open problem. Though these cosmic anomalies might be due to outliers in data or inappropriate statistical methods, they may also be due to new physics. Therefore, astronomers hope to further unveil the mystery behind these cosmic anomalies through the next generation missions.

The next generation CMB missions, such as CMB-S4 (Abazajian et al. (2019)), COrE (Collaboration et al. (2011)), Euclid (European Space Agency (2021a) and Racca et al. (2016)), and LiteBIRD (Matsumura et al. (2014)) have been designed for more refined measurements of the detailed structure of CMB and the universe. CMB-S4 is a ground-based CMB mission which will observe the CMB with 21 telescopes physically located at Atacama desert, Chile and South Pole. COrE is a full-sky satellite which will detect microwave-band frequencies and will mainly focus on the CMB polarization. Euclid is a European Space Agency based mission which will use a Korsch-type telescope to measure the visible to near-infrared wavelengths. LiteBIRD is a satellite which will explore the B-mode polarization patterns and cosmic inflation. They will investigate the CMB deeply at narrower angles that will result in high resolution CMB maps. It is believed that the CMB data from these missions will interpret the slight variations of the CMB temperature and polarization with an unprecedented precision. Thus, giving hope to further unveil the non-Gaussianity of the CMB data, these missions will hopefully discover the in-depth details of dark matter, dark energy and the development of the cosmic structures.

1.8 Overview of the results

The main objective of this thesis is to study and develop stochastic models and statistical methods for spatial and long-range dependent data. These data are modelled as realizations of spherical random fields and functional time series. Also, it concerns the investigation of non-Gaussianity in the CMB data from the Planck mission and develops a methodology which is useful for investigating non-Gaussianity in future CMB missions.

First, the thesis studies the multifractal and multifractional behaviour of spherical random fields. The contents of Chapters 3 and 4 focus on this context. Secondly, the thesis studies the statistical inference of long-range dependent data with specific consideration to cyclic long-memory processes. The content of Chapter 5 focuses on this study. The results obtained in this thesis are novel and were derived considering rather general assumptions.

The concepts of multifractality and multifractionality play an important role in studying the properties of the underlying random fields. Although multifractal and multifractional approaches have been extensively used in the one-dimensional case, their applications in the multi-dimensional case or on manifolds are less developed. This thesis extends the applications of multifractal and multifractional theory to the case of spherical random fields. Next, developing inferential statistics for functional time series with spectral singularities at non-zero frequencies make a significant contribution to the statistics of random processes with long-range dependence. Numerical studies including Meyer, Shannon and Mexican wavelets and extensive simulation studies were conducted to confirm the theoretical findings.

Chapter 2 presents the fundamental concepts, notations, definitions and auxiliary results related to the theory of stochastic processes, random fields and spherical random fields. Further, it introduces the HEALPix ordering schemes employed in the CMB data analysis, spherical harmonics and Hermite polynomials. Some of the definitions in this chapter may reappear in Chapters 3, 4 and 5 as these chapters are the corresponding published/submitted articles.

Chapter 3 is based on the paper published by Leonenko, N., Nanayakkara, R. and Olenko, A. Analysis of spherical monofractal and multifractal random fields, *Stochastic Environmental Research and Risk Assessment*, 35(3):681–701, (2021) (Leonenko et al. (2021)).

Chapter 4 is based on the paper submitted by Broadbridge, P., Nanayakkara, R. and Olenko, A. On multifractionality of spherical random fields with cosmological applications (Broadbridge et al. (2021)).

Chapter 5 is based on the paper submitted by Ayache, A., Fradon, M., Nanayakkara, R. and Olenko, A. Asymptotic normality of simultaneous estimators of cyclic long-memory processes, which will appear in *Electronic Journal of Statistics* (Ayache et al. (2021)).

In Chapter 3, the Rényi function approach was used to examine the multifractality of spherical random fields whereas in Chapter 4, the Hölder exponent approach was used to probe the multifractionality of spherical random fields. Chapter 5 of this thesis proves the asymptotic normality behaviour of the simultaneous estimators of parameters of cyclic long-memory processes. Initially, we planned to develop the methodology for stochastic processes and then to extend the results in Chapter 5 to the case of random fields. Due to a lack of time and as the thesis was prolonged, the third research project was undertaken only for the one-dimensional case. The results for random fields similar to Chapter 5 will be finalized in future publications.

The computing techniques implemented for modelling purposes, to obtain results and to uphold the theoretical findings in Chapters 3, 4 and 5 are correspondingly included in Appendices A, B and C as source codes. As the serial computational methodology to obtain many results, in particular, Figure 5.4, which is included in Appendix C is rather time consuming, its parallelized version is included in Appendix D. Appendix D provides a detailed analysis of the performance of the parallelized code over the serial version. All the numerical simulation studies were carried out using the high performance computer Gadi of the NCI. The obtained results code and methodology can be applied to other spherical, geoscience, environmental, medical imaging, embryology and functional time series data.

The main content of this thesis is based on three published or submitted articles. The author made equal contributions in the process of developing the theoretical section, proving results, providing examples, composing the articles and preparing the final editions of the articles for publication. The author made the main contribution in developing the software code. In addition, the author conducted all the statistical simulation studies specifically applying high performance computing techniques through the Linux computational cluster Gadi of the NCI.

The obtained research findings and results were presented at the Young Statisticians Showcase 2019 of the Victorian branch of the Statistical Society of Australia, Young Statisticians Conference 2019 Canberra, SEMS and SHE College Three Minute Thesis (3MT) competition at La Trobe University, Virtual Poster Pitch Competition of the Statistical Society of Australia, Fourth Victorian Research Students' Meeting in Probability and Statistics 2020, AustMS 2020 Conference, Spatial and Temporal Statistics Symposium 2021 at the University of Wollongong and Early-Career & Student Statisticians Conference 2021.

Chapter 2

Definitions, notations and auxiliary results

This chapter presents main concepts and notations relevant to the study. This includes the mathematical notations related to d-dimensional Euclidean space, basic material related to HEALPix ordering schemes, spherical harmonics, Hermite polynomials, probability theory, stochastic processes, random fields and spherical random fields that will compose the background of the study. Most of the materials included in this chapter are based on Abramowitz and Stegun (1964), Bingham et al. (1989), Broadbridge et al. (2020), Florescu (2014), Fryer and Olenko (2019), Fryer et al. (2020), Gorski et al. (2005), Hernàndez (1995), Hivon (2021), Ivanov and Leonenko (1989), Lang and Schwab (2015), Leonenko (1999), Leonenko and Olenko (2013), Leonenko and Shieh (2013), Leonenko and Olenko (2014), Marinucci and Peccati (2011), Peccati and Taqqu (2011), Seneta (1976), Taqqu (1975), Taqqu (1979) and Yadrenko (1983).

2.1 Mathematical notations related to d-dimensional Euclidean space

This section presents basic notations related to d-dimensional Euclidean space which are being used through out this thesis.

Let $\mathbb{R}^d$ denote a real Euclidean space of dimension $d \geq 2$. In what follows, $|\cdot|$ and $\|\cdot\|$ denote the Lebesgue measure and the Euclidean distance in $\mathbb{R}^d$, respectively. Let $\mathbb{R}^d_+ = \left\{x \in \mathbb{R}^d : x_i \geq 0, i = 1, \ldots, d\right\}$ denote the non-negative octant of $\mathbb{R}^d$. For $X \subseteq R^m$ we denote its Borel σ -algebra by $\mathcal{B}(X)$.

Let $B(r) = \left\{x \in \mathbb{R}^d : \|x\| < r\right\}$ and $s_{d-1}(r) = \left\{x \in \mathbb{R}^d : \|x\| = r\right\}$ denote a d-dimensional ball and a $(d-1)$-dimensional sphere in $\mathbb{R}^d$ having radius $r > 0$ and with center at the origin. The $SO(d)$ denotes the group of rotations in $\mathbb{R}^d$. For $d \geq 2$, the Lebesgue measure/volume of the d-dimensional ball in $\mathbb{R}^d$, is $|B(r)| = \pi^{\frac{d}{2}}\Gamma^{-1}\left(\frac{d}{2}+1\right) r^d$ and the Lebesgue measure/surface area of the $(d-1)$-dimensional sphere in $\mathbb{R}^d$, is $|s_{d-1}(r)| = 2\pi^{\frac{d}{2}}\Gamma^{-1}\left(\frac{d}{2}\right) r^{d-1}$, where $\Gamma(z)$ denotes the Gamma function.

The Cartesian coordinate system, $x = (x_1, x_2, ..., x_d)$ in $\mathbb{R}^d$ is related with the spherical coordinate system, (r, ϕ), where the radius $r \geq 0$ and $\phi = (\phi_1, \ldots, \phi_{d-1}) \in s_{d-1}(1)$ and $0 \leq \phi_1 < 2\pi, 0 \leq \phi_j < \pi, j = 2, \ldots, d-1$, by the following formulae

$$x_1 = r\sin(\phi_{d-1})\ldots\sin(\phi_2)\sin(\phi_1), \quad x_2 = r\sin(\phi_{d-1})\ldots\sin(\phi_2)\cos(\phi_1)$$

$$x_3 = r\sin(\phi_{d-1})\ldots\sin(\phi_3)\cos(\phi_2), \ldots, x_{d-1} = r\sin(\phi_{d-1})\cos(\phi_{d-2}), x_d = r\cos(\phi_{d-1})$$

Let $dx = dx_1 \ldots dx_d$ denote an element of the Lebesgue measure in $\mathbb{R}^d$ and $d\sigma(x) = r^{d-1}\sin^{d-2}(\phi_{d-1})\cdots\sin(\phi_2)d\phi_{d-1}\cdots d\phi_2 d\phi_1$ denote an element of the Lebesgue measure on the $(d-1)$-dimensional sphere $s_{d-1}(r)$.

Let the Kronecker delta function be defined by

$$\delta_a^b = \begin{cases} 0, & \text{if } a \neq b, \\ 1, & \text{if } a = b. \end{cases}$$

2.2 HEALPix ordering schemes

This section introduces basic material pertaining to the HEALPix ordering schemes.

HEALPix stands for the Hierarchical Equal Area isoLatitude Pixelisation which is a uniform grid on the unit sphere $s_2(1)$. HEALPix is a spherical data representation method which efficiently stores and organizes cosmic microwave background radiation (CMB) data on the sphere. The stored CMB data using HEALPix format are available as flexible image transport system (FITS) files. Each CMB pixel has a set of attributes such as distinct location, temperature intensity and polarisation data.

In comparison with the other spherical data representation formats, HEALPix has many advantages which make it applicable for high resolution data. They are,

- HEALPix has equal area pixels useful for equal area random sampling;
- HEALPix has hierarchical tessellations of the sphere useful for zooming in small details quickly;
- HEALPix has iso-Latitude rings of pixels useful for computing fast spherical harmonic transform.

The HEALPix structure initially divides the unit sphere $s_2(1)$ into 12 equiareal quadrilateral base pixels belonging to 3 rings. Figure 2.1 depicts the planar projection of the HEALPix base pixel structure consisting of 12 squares for $j = 0$ where j is the resolution parameter. Here, the pixel centers belong to an equatorial and two polar rings. Then, for $j > 0$, each of these equiareal quadrilateral base pixels are further subdivided into 4-equiareal quadrilateral child pixels until the required resolution is achieved.

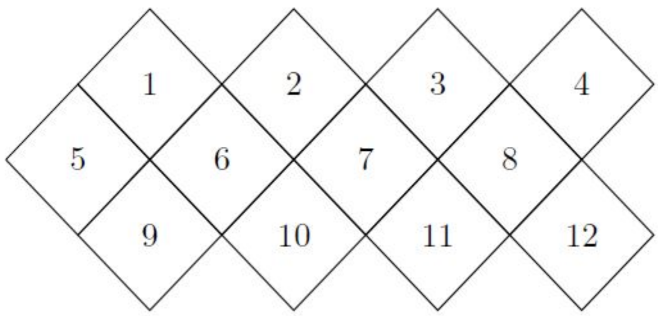


Figure 2.1: The planar projection of the HEALPix base pixel tessellation at $j = 0$

Let j denote the resolution parameter, N_{side} denote the grid resolution parameter and N_{pix} denote the total number of pixels on $s_2(1)$. They are related by the formulae,

$N_{side} = 2^j$ and $N_{pix} = 12 \times (N_{side})^2$ where $j \geq 0$.

Let Φ_{pix} denote the area of a equiareal quadrilateral base pixel on $s_2(1)$. Then, $\Phi_{pix} = \dfrac{\pi}{(3 \times (N_{side})^2)}$.

Let N_{ring} denote the total number of rings on $s_2(1)$ for given $j \geq 0$. Then, $N_{ring} = (4 \times N_{side}) - 1$.

Remark 2.1. The geometrical properties of the HEALPix structure has divided the unit sphere $s_2(1)$ in three parts: upper, middle and lower parts. The rings belonging to the upper part are between the north pole and the upper boundary of the equatorial region whereas the rings belonging to the lower part are between the lower boundary of the equatorial region and the south pole. Each upper and lower part consists of $(N_{side} - 1)$ rings. There are $((2 \times N_{side}) + 1)$ rings in the middle part and all of them have an equal number of pixels given by $(4 \times N_{side})$.

Example 2.1. *N_{pix} values and the corresponding N_{ring} values for the first few j values are*

$j = 1$: $N_{side} = 2$, $N_{pix} = 48$ and $N_{ring} = 7$;

$j = 2$: $N_{side} = 4$, $N_{pix} = 192$ and $N_{ring} = 15$;

$j = 3$: $N_{side} = 8$, $N_{pix} = 768$ and $N_{ring} = 31$.

The HEALPix structure has two major ordering/numbering schemes. They are nested ordering scheme and ring ordering scheme. According to the ring ordering scheme, pixel indices are located in the anticlockwise direction in each ring winding down from the north pole to the south pole. This has enabled to perform fast Fourier transform with spherical harmonics. Figure 2.2 shows the planar projection of the HEALPix ring ordered structure at $j = 1$ (Fryer and Olenko (2019)). It is clear from Figure 2.2 that the HEALPix indices are increasing in the ascending order from left to right consecutively through the rings.

Next, according to the nested ordering scheme, pixel indices are located in a tree-like structure which grows within the consecutively ordered branches. The roots of these ordered branches are the 12 equiareal quadrilateral base pixels. This ordering scheme has facilitated efficient development of applications involving nearest-neighbour searches. Also, it has enabled speedy computation of the fast Haar wavelet transform on the HEALPix structure. Figure 2.3 depicts the planar projection of the HEALPix nested ordered structure at $j = 1$ (Fryer and Olenko (2019)).

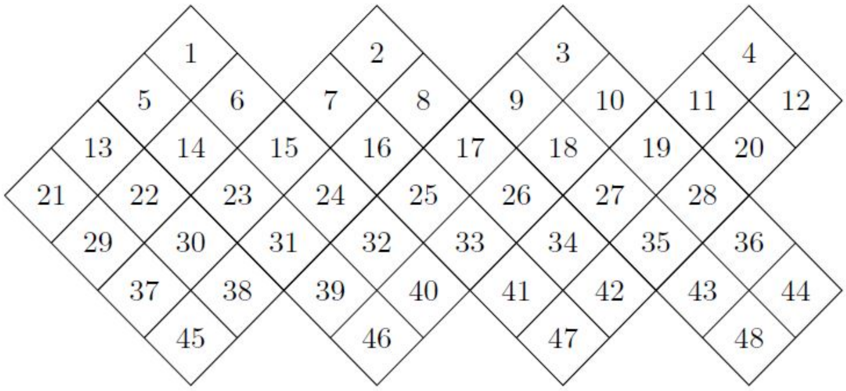


Figure 2.2: The planar projection of the HEALPix ring ordering scheme at $j = 1$

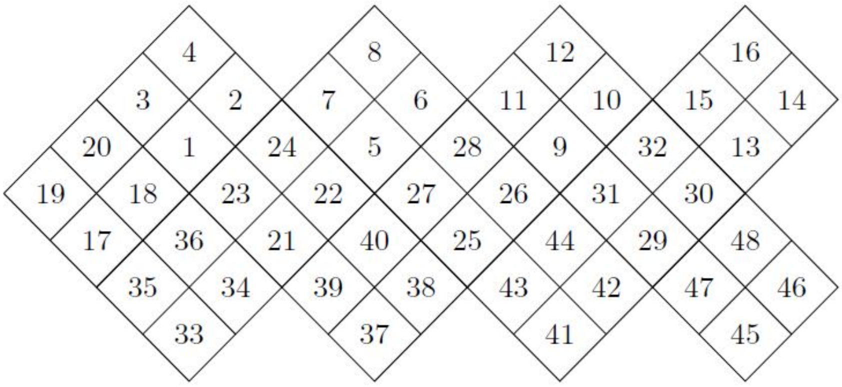


Figure 2.3: The planar projection of the HEALPix nested ordering scheme at $j = 1$

2.3 Spherical harmonics

This section presents basic notations and definitions related to spherical harmonics and their properties. Spherical harmonics are special functions defined on the surface of the sphere. Most of the materials included in this section are based on Ivanov and Leonenko (1989) and Marinucci and Peccati (2011).

In this thesis, we consider the complex spherical harmonics for $d = 3$ in $\mathbb{R}^3$ as follows.

Definition 2.1. For every integer $l = 0, 1, 2, \ldots$, and $m = 0, \pm 1, ..., \pm l$, the spherical harmonic function $Y_l^m(\cdot)$ is defined as

$$Y_l^m(\theta, \varphi) = (-1)^m \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta) e^{im\varphi}, \tag{2.1}$$

where $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi)$ and $P_l^m(\cdot)$ is the associated Legendre polynomials having degree l and order m.

Definition 2.2. The associated Legendre functions of the first kind of degree l and order m denoted by $P_l^m(x)$, $x \in [-1,1]$ for $l \geq 0$, $m = 0, \ldots, l$ are defined as

$$P_l^m(x) = (-1)^m \left(1 - x^2\right)^{m/2} \frac{d^m}{dx^m} P_l(x),$$

where the l-th Legendre polynomials are $P_l(x) = \dfrac{1}{2^l l!} \dfrac{d^l}{dx^l} \left(x^2 - 1\right)^l$.

The spherical harmonics have the following properties:

(i) Orthonormality: For all $l, l^{'} \geq 0$, $-l \leq m \leq l$ and $l^{'} \leq m^{'} \leq l^{'}$,

$$\int_0^\pi \int_0^{2\pi} Y_l^m(\theta,\varphi) \overline{Y_{l'}^{m'}(\theta,\varphi)} \sin\theta d\theta d\varphi = \delta_l^{l'} \delta_m^{m'},$$

(ii) Symmetry: $\overline{Y_l^m(\theta,\varphi)} = (-1)^m Y_l^{(-m)}(\theta,\varphi)$,

(iii) Addition formula: $\sum_{m=-l}^{l} Y_l^m(\theta,\varphi)\overline{Y_l^m(\theta,\varphi)} = \dfrac{2l+1}{4\pi}$,

(iv) Relation with Legendre polynomials: $Y_l^0(\theta,\varphi) = \sqrt{\dfrac{2l+1}{4\pi}} P_l(\cos\theta)$,

(v) $Y_l^0(0,0) = \sqrt{\dfrac{2l+1}{4\pi}}$,

(vi) Spatial inversion: $Y_l^m(\pi - \theta, \pi + \varphi) = (-1)^l Y_l^m(\theta,\varphi)$.

Example 2.2. *The first spherical harmonics are given as follows.*

- *For* $l = 0$, $Y_0^0(\theta,\varphi) = \dfrac{1}{\sqrt{4\pi}}$;
- *For* $l = 1$, $Y_1^0(\theta,\varphi) = \sqrt{\dfrac{3}{4\pi}} \cos\theta$ *and* $Y_1^1(\theta,\varphi) = -\sqrt{\dfrac{3}{8\pi}} \sin\theta e^{i\varphi}$;
- *For* $l = 2$, $Y_2^0(\theta,\varphi) = \dfrac{1}{2}\sqrt{\dfrac{5}{4\pi}} \left(3\cos^2\theta - 1\right)$, $Y_2^1(\theta,\varphi) = -\sqrt{\dfrac{15}{8\pi}} \sin\theta\cos\theta e^{i\varphi}$, *and* $Y_2^2(\theta,\varphi) = \dfrac{1}{4}\sqrt{\dfrac{15}{2\pi}} \sin^2\theta e^{2i\varphi}$.

Theorem 2.1. *(Peter-Weyl Theorem on the sphere)*
Let $f(\cdot)$ be a square integrable function defined on $s_2(1)$. Then, it holds

$$f(\theta,\varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_l^m Y_l^m(\theta,\varphi),$$

where $a_l^m = \int_{s_2(1)} f(\theta,\varphi)\overline{Y_l^m(\theta,\varphi)} \sin\theta d\theta d\varphi$, and the series converges in the square integrable sense.

2.4 Hermite polynomials

This section gives basic notations and definitions related to Hermite polynomials and their properties. Hermite expansions are important in investigating the asymptotics of nonlinear functionals of random fields. It is well known that their limiting distributions are governed by the leading terms of Hermite expansions. Most of the materials included in this section are based on Ivanov and Leonenko (1989), Leonenko and Olenko (2014) and Taqqu (1975).

Definition 2.3. Let $H_q(\cdot)$, $q \geq 0$, denote the q^{th} order Hermite polynomials on $\mathbb{R}$. Then the Hermite polynomials are defined by the following relation.

$$H_q(x) = (-1)^q \exp\left(\frac{x^2}{2}\right) \frac{d^q}{dx^q} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

Next, we consider some examples of Hermite polynomials.

Example 2.3. *The first few Hermite polynomials are given by the following expressions.*

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x, \quad \textit{and} \quad H_4(x) = x^4 - 6x^2 + 3.$$

The q^{th} order Hermite polynomial is a polynomial of degree q. Now, we consider some properties of Hermite polynomials.

Let $L_2(\mathbb{R}, \phi(x)\mathrm{d}x)$ be a Hilbert space, where $\phi(x) = \dfrac{e^{-x^2/2}}{\sqrt{2\pi}}$, $x \in \mathbb{R}$. The Hermite polynomials have the following properties.

(i) Orthogonality: For all m, $n \geq 0$,

$$\int_{-\infty}^{\infty} H_m(x) H_n(x) \phi(x) dx = \delta_m^n m!,$$

(ii) Symmetry: $H_n(-x) = (-1)^n H_n(x)$,

(iii) Recurrence relation: For $n \geq 1$,

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x) \quad \text{and} \quad H_n'(x) = nH_{n-1}(x).$$

(iv) Explicit expression: For $n \geq 0$,

$$H_n(x) = n! \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k x^{n-2k}}{k! 2^k (n-2k)!},$$

where $\lfloor \cdot \rfloor$ denotes the floor function.

Another important property of the Hermite polynomial is given by the following lemma.

Lemma 2.1. *Let $(X_1, \ldots, X_{2p})$ be a $2p$-dimensional zero mean Gaussian vector such that*

$$\mathrm{E}(X_j X_k) = \begin{cases} 1, & \text{if } k = j; \\ r_j, & \text{if } k = j + p \text{ and } 1 \leq j \leq p; \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\mathrm{E}\left(\prod_{j=1}^{p} H_{k_j}(X_j) H_{m_j}(X_{j+p})\right) = \prod_{j=1}^{p} \delta_{k_j}^{m_j} k_j! r_j^{k_j}.$$

Let $G(\cdot)$ be an arbitrary function defined in $L_2(\mathbb{R}, \phi(x)\mathrm{d}x)$. Then, the function $G(\cdot)$ can be expressed in terms of Hermite polynomials for $j \geq 0$ and it possesses the $L_2(\mathbb{R}, \phi(x)dx)$ convergent expansion,

$$G(x) = \sum_{n=0}^{\infty} \frac{C_n H_n(x)}{n!}, \quad C_n = \int_{\mathbb{R}} G(x) H_n(x) \phi(x) dx.$$

Definition 2.4. Let the function $G(x)$ be in $L_2(\mathbb{R}, \phi(x)\mathrm{d}x)$ and suppose there exists an integer $\kappa \geq 0$ such that $C_n = 0$, for all $0 \leq n \leq (\kappa - 1)$, but $C_\kappa \neq 0$. The positive integer κ is termed the Hermite rank of $G(x)$ and is denoted by $H\mathrm{rank}G$.

2.5 Probability theory

This section presents basic notations, definitions, propositions and theorems related to the probability theory.

Definition 2.5. The ordered triple $(\Omega, \mathcal{F}, P)$ is called a probability space if

- Ω is a non-empty set (the space of outcomes),
- $\mathcal{F}$ is a σ-algebra of subsets of Ω (the family of events),
- $P : \mathcal{F} \to [0, 1]$ is a probability measure.

Definition 2.6. Given a probability triple $(\Omega, \mathcal{F}, P)$, a random variable is a function X from Ω to the real numbers $\mathbb{R}$, such that $\{\omega \in \Omega; X(w) \leq x\} \in \mathcal{F}$ for all $x \in \mathbb{R}$.

The cumulative distribution function (CDF) or simply distribution function of X is given by $F_X(x) = P(X \leq x)$. For a continuous random variable X, $f_X(\cdot)$ is called the probability density function (PDF) of the random variable X and it is related with the CDF by $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$.

Definition 2.7. A random variable X has expectation if $\int XdP$ is finite. The expected value of X is then $\mathrm{E}(X) = \int_\Omega XdP$. If X has a continuous density f_X then, $\mathrm{E}\left(\varphi(X)\right) = \int_{-\infty}^{+\infty} \varphi(y) f_X(y)dy$.

We state $X \in L_1(\Omega, \mathcal{F}, P)$, or $X \in L_1$, if X has expected value. We say that X is square integrable if $X^2 \in L_1$, and state $X \in L_2(\Omega, \mathcal{F}, P)$, or $X \in L_2$.

Definition 2.8. If $X \in L_2$, its variance is defined by $\mathrm{Var}(X) = \mathrm{E}(X - \mathrm{E}(X))^2$.

Given $X, Y \in L_2$, their inner product can be defined by $\langle X, Y\rangle = \mathrm{E}(X\overline{Y})$. The norm of $X \in L_2$ is $\|X\| = \sqrt{\langle X, X\rangle}$.

Definition 2.9. For a given probability space $(\Omega, \mathcal{F}, P)$, an n-dimensional random vector is a function $X : \Omega \to \mathbb{R}^n$ such that each X_i is a random variable. Here X_i is the i^{th} coordinate of $X(\omega) = (X_1(\omega), \ldots, X_n(\omega))$.

Definition 2.10. If the components $X_1, ..., X_n$ of a random vector X have expectation, then we say that $X \in L_1$ and $\mathrm{E}(X) = (\mathrm{E}(X_1), ..., \mathrm{E}(X_n))^T$. When $X \in L_2$ then each of the expectations $\mathrm{E}(X_i X_j)$ exist. The covariance matrix of X is defined as

$$\mathrm{Cov}(X) = \mathrm{E}(X - \mathrm{E}(X))(X - \mathrm{E}(X))^T.$$

Theorem 2.2. *(Central Limit Theorem) If $X_1, ..., X_n \in L_2$ are iid with mean m and variance σ^2, then*

$$\lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n} X_i - m}{\sigma/\sqrt{n}} \leq x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy,$$

uniformly in $x \in \mathbb{R}$.

Definition 2.11. Let $f \in L_1(\mathbb{R})$. Then

$$\hat{f}(y) = \int_{-\infty}^{\infty} e^{-2\pi ixy} f(x)dx,$$

is called the Fourier transform of f.

2.6 Stochastic processes

This section gives basic results related to stochastic processes.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. It will be always assumed that all random elements are defined on this probability space.

Definition 2.12. A stochastic process $X(t)$, $t \in T$, is a collection of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. Sometimes, it will also be denoted as X_t.

The stochastic process $X(t), t \in T$, is called a discrete-time stochastic process if $T = \{0, 1, 2, ...\}$. It is called a a continuous-time stochastic process if $T = [0, \infty)$.

Definition 2.13. The probability space $(\Omega, \mathcal{F}, P)$ is a filtered probability space if there exist a sequence of σ-algebras $\{\mathcal{F}_t\}_{t \in T}$ included in $\mathcal{F}$ such that it is an increasing collection, that is, $F_s \subseteq F_t$ for all $s \leq t$, $s, t \in T$.

Definition 2.14. A stochastic process $X(t)$, $t \in T$ defined on a filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\})$ is called adapted if $X(t)$ is $\mathcal{F}_t$-measurable for any $t \in T$.

Let's assume that a standard probability space is given where no filtration is defined on the space. It means that for time t, all the information present is generated from the stochastic process $X(t)$ only. Then, the standard filtration originated from the stochastic process $X(t), t \in T$, is given by $\mathcal{F}_t = \sigma(\{X_s : s \leq t, s \in T\})$.

The Gaussian processes are an important class of stochastic processes. For these stochastic processes $X(t)$, all stochastic vectors $(X(t_1), ..., X(t_n))$ with $(t_1, ..., t_n) \in T$ are normally distributed.

Definition 2.15. For a stochastic process $X(t)$, $t \in T$, the function $m(t) = \mathrm{E}(X(t))$ is called the mean value function and the function $\sigma^2(t) = \mathrm{Var}(X(t))$ is called the variance function of the process. The function $K(s, t) = \mathrm{Cov}(X(s), X(t))$ defined for $s \in T$ and $t \in T$ is called a covariance function of the process.

Definition 2.16. A stochastic process $X(t)$, $t \in T$ is called stationary if the finite dimensional distributions are translation invariant, i.e. the random vector $(X(t_1 + c), ..., X(t_n + c))$ has the same distribution as the random vector $(X(t_1), ..., X(t_n))$ for all integer n, and $t_1, ..., t_n$, $c \in T$.

Definition 2.17. A stochastic process $X(t)$ is said to be wide sense stationary or covariance stationary if $X(t)$ has finite second moments for any t and its covariance function $\mathrm{Cov}(X(t), X(t+h))$ depends only on $h \in T$ for all $t \in T$.

Definition 2.18. Let's assume that T is totally ordered. A stochastic process $X(t)$ is said to have independent increments if the random variables $(X(t_2) - X(t_1), X(t_3) - X(t_2), ..., X(t_n) - X(t_{n-1}))$ are independent for any integer n and any choice of the sequence $\{t_1, t_2, ..., t_n\}$ in T with $t_1 < t_2 < ... < t_n$.

Definition 2.19. A stochastic process $X(t)$ is said to have stationary increments if the distribution of the random variable $X(t+h) - X(t)$ depends only on h and not on t.

2.7 Random fields

This section demonstrates basic notations, definitions, propositions and theorems related to random fields.

Let $T \subseteq \mathbb{R}^d$ be a multidimensional set. Then, for $T \subseteq \mathbb{R}^d$, we denote the Borel σ-algebra by $\mathcal{B}(T)$.

Definition 2.20. A random field is defined as a function $\xi(\omega, x) : \Omega \times T \to \mathbb{R}^m$ such that $\xi(\omega, x)$ is a random variable for each $x \in T$. A random field will also be denoted as $\xi(x), x \in \mathbb{R}^d$.

The random field $\xi(x)$ is a random process when $d = 1$. The random field $\xi(x) \in T$ is called a scalar random field for $d > 1$ and $m = 1$. It is called a vector random field for $m > 1$.

This thesis focuses on scalar random fields. The function $\xi(x), x \in T$, is termed as a realization of the random field for fixed $\omega \in \Omega$.

Definition 2.21. A second order random field over $T \subset \mathbb{R}^d$ is a function $\xi(x) : T \to L_2(\Omega, \mathcal{F}, P)$.

Definition 2.22. The finite dimensional distributions of the random field $\xi(x)$, $x \in T$, are defined as a set of distributions $\mathrm{P}\left(\xi(x_i) \in B_i, i = 1, ..., r\right)$ where $B_i \in \mathcal{B}$, $i = 1, \ldots, r$, $r \in \mathbb{N}$.

Definition 2.23. Two random fields $\xi(x)$ and $\eta(x), x \in T$, are said to be stochastically equivalent if the finite-dimensional distributions of the random fields $\xi(x)$ and $\eta(x)$ coincide. It will be denoted by $\xi(x) \stackrel{d}{=} \eta(x)$, $x \in T$.

Definition 2.24. A scalar random field $\xi(x) \in T$ is called Gaussian if $\xi(x) \in T$ is a Gaussian system of random variables.

The finite-dimensional distributions of a Gaussian random field $\xi(x)$, $x \in T$, are determined by its mean and covariance functions. The mean function is defined by $\mathrm{E}\left(\xi(x)\right) = \mu(x)$, $x \in T$. The covariance function $B(x,y) = \mathrm{Cov}\left(\xi(x), \xi(y)\right)$ is defined by $B(x,y) = \mathrm{E}\left(\left(\xi(x) - \mu(x)\right)\left(\xi(y) - \mu(y)\right)\right)$ where $x, y \in T$. Here, when $x = y$, the function $B(x,x) = \mathrm{Var}\left(\xi(x)\right)$ is known as the variance of the random field $\xi(x)$, $x \in T$.

Definition 2.25. A real random field $\xi(x)$, $x \in T$, with a finite second moment is said to be mean-square continuous at point x_0 if $\mathrm{E}\,|\xi(x) - \xi(x_0)|^2 \to 0$ as $||x - x_0|| \to 0$. If this relation holds for any $x_0 \in T$, the random field $\xi(x)$, $x \in T$, is called mean square continuous on T.

A random field $\xi(x)$, $x \in T$, satisfying the condition $\mathrm{E}\left(\xi^2(x)\right) < \infty$ is mean square continuous if and only if the function $\mathrm{Cov}(\xi(x), \xi(y))$ is continuous along the diagonal $\{(x,y) \in T \times T : x = y\}$.

Theorem 2.3. *(Karhunen (1947))*
Let $\xi(x)$, $x \in T \subseteq \mathbb{R}^d$ be a random field with $\mathrm{E}\left(\xi(x)\right) = 0$, $\mathrm{E}\,|\xi(x)|^2 < \infty$. If its covariance can be expressed as

$$B(x,y) = \mathrm{E}\left(\xi(x)\xi(y)\right) = \int_\Lambda f(x,\lambda)\overline{f(y,\lambda)}F(d\lambda), \quad f(x,.) \in L_2(\Lambda), x, y \in T,$$

then there exists an orthogonal complex-valued random measure Z on Λ with the control measure F such that

$$\xi(x) = \int_\Lambda f(x,\lambda)Z(d\lambda), \tag{2.2}$$

where $\mathrm{E}\,|Z(\Delta)|^2 = F(\Delta)$, $\Delta \in \mathcal{B}(\Lambda)$, and the stochastic integral in (2.2) is viewed as an $L_2(\Lambda)$ integral.

Next, we focus on the spectral theory of homogeneous and isotropic random fields.

Definition 2.26. A real random field $\xi(x)$, $x \in T$, satisfying $\mathrm{E}\left(\xi^2(x)\right) < \infty$ is called homogeneous in the wide sense if its mean function $m(x) = \mathrm{E}\left(\xi(x)\right)$ and the covariance function $B(x,y) = \mathrm{Cov}(\xi(x), \xi(y))$ are invariant with respect to the Abelian group $G = (\mathbb{R}^d, +)$ of shifts in $\mathbb{R}^d$. i.e. $m(x) = m(x+\tau)$, $B(x,y) = B(x+\tau, y+\tau)$ for any $x, y, \tau \in T$.

Definition 2.27. A random field $\xi(x)$, $x \in T$, satisfying $\mathrm{E}\left(\xi^2(x)\right) < \infty$ is called isotropic in the wide sense if its mean function $m(x)$ and the covariance function $B(x,y)$ are invariant with respect to the group of rotations $\mathrm{SO}(d)$. i.e. $m(x) = m(gx), B(x,y) = B(gx, gy)$ for any $x, y \in T, g \in \mathrm{SO}(d)$.

Remark 2.2. Let $\xi(x)$, $x \in T$, be a homogeneous isotropic random field. Then its mean function, $\mathrm{E}\left(\xi(x)\right) = const$ and the covariance function $B(x,y)$, $x, y \in T$, depends only on the Euclidean distance $\|x-y\|$ between the points x and y. i.e. $B(x,y) = B(\|x-y\|)$.

For $\nu > -\frac{1}{2}$, the Bessel function of the first kind of order ν is given by

$$J_\nu(z) = \sum_{l=0}^{\infty}(-1)^l \left(\frac{z}{2}\right)^{2l+\nu} [l!\Gamma(l+\nu+1)]^{-1}, \quad z > 0,$$

and the spherical Bessel function by

$$Y_1(z) = \cos z, \quad Y_d(z) = 2^{(d-2)/2}\Gamma\left(\frac{d}{2}\right) J_{(d-2)/2}(z) z^{(2-d)/2}, \quad z \geq 0, d \geq 2.$$

Let $\xi(x)$, $x \in \mathbb{R}^d$, be a measurable mean-square continuous zero-mean homogeneous isotropic real-valued Gaussian random field. Then, the function $B(r)$ is the covariance function of this field if and only if there exists a finite measure $G(\cdot)$ on $(\mathbb{R}_+, (\mathcal{B}(\mathbb{R}_+))$ such that

$$\mathrm{B}(r) = \mathrm{Cov}(\xi(x), \xi(y)) = \int_0^\infty Y_d(rz) G(\mathrm{d}z), \tag{2.3}$$

for $r = \|x-y\|$, $x, y \in \mathbb{R}^d$ and $z \geq 0$ with $G(\mathbb{R}_+) = B(0) < \infty$. The function $G(\cdot)$ is called the isotropic spectral measure.

Example 2.4. *The covariance functions of the random fields in $\mathbb{R}^2$ and $\mathbb{R}^3$ are of great importance. Therefore for $d = 2$, the equation (2.3) takes the form,*

$$B(r) = \int_0^\infty J_0(rz) G(dz),$$

and for $d = 3$, it takes the form,

$$B(r) = \int_0^\infty \frac{\sin(rz)}{rz} G(dz).$$

Definition 2.28. Let $G(\cdot)$ be the isotropic spectral measure of the random field $\xi(x)$, $x \in \mathbb{R}^d$. The spectral measure $G(\cdot)$ is said to be absolutely continuous if there exists a function $f(\cdot)$ such that

$$G(z) = 2\pi^{d/2}\Gamma^{-1}\left(\frac{d}{2}\right)\int_0^z u^{d-1}f(u)du, \quad z \geq 0, \quad u^{d-1}f(u) \in L_1(\mathbb{R}_+)$$

where $f(\cdot)$ is called the isotropic spectral density of the field $\xi(x)$, $x \in \mathbb{R}^d$.

In this case, the covariance function can be represented by

$$B(r) = 2\pi^{d/2}\Gamma^{-1}\left(\frac{d}{2}\right)\int_0^\infty Y_d(ru)u^{d-1}f(u)du. \tag{2.4}$$

Definition 2.29. The Gaussian random field $\xi(x)$, $x \in \mathbb{R}^d$, with an absolutely continuous spectrum has the isonormal spectral representation given by

$$\xi(x) = \int_{\mathbb{R}^d} e^{i(\lambda,x)}\sqrt{f(\|\lambda\|)}W(d\lambda),$$

where $W(\cdot)$ is the complex Wiener white noise random measure on $\mathbb{R}^d$.

In this thesis, we consider long-range dependent random processes and fields. In the asymptotic theory of long-range dependent random fields, slowly varying functions play an important role. They are used to represent fluctuations in the covariance and spectral density functions.

Definition 2.30. A measurable function $\mathcal{L} : (0,\infty) \to (0,\infty)$ is called slowly varying at infinity if for all $\lambda > 0$,

$$\lim_{t\to\infty}\frac{\mathcal{L}(\lambda t)}{\mathcal{L}(t)} = 1.$$

Each slowly varying function has the following representation (Karamata (1930)).

Theorem 2.4. *(Representation Theorem)*
Let $\mathcal{L}(\cdot)$ be a slowly varying function. Then, for all $t \geq a$, $a > 0$, $\mathcal{L}(\cdot)$ can be represented in the following form

$$\mathcal{L}(t) = \exp\left(\zeta_1(t) + \int_a^t \frac{\zeta_2(u)}{u}du\right),$$

where $\zeta_1(\cdot)$ and $\zeta_2(\cdot)$ are two bounded, measurable functions satisfying $\zeta_1(t) \to b$, $b \in \mathbb{R}$, and $\zeta_2(t) \to 0$ as $t \to \infty$.

Example 2.5. *The following functions are slowly varying for $q_1, q_2, q_3 > 0$ and $\beta \in \mathbb{R}$:*

$$\mathcal{L}(t) \equiv 1, \quad \mathcal{L}(t) = (\log(q_1 + t))^{q_2}, \quad \mathcal{L}(t) = \log\log(t + q_1),$$

$$\mathcal{L}(t) = \log^{\beta} t, \quad \mathcal{L}(t) = \frac{t^{q_2 q_3}}{(q_1 + t^{q_2})^{q_3}}, \quad \mathcal{L}(t) = \exp\left((\log t)^{\frac{1}{3}} \cos\left((\log t)^{\frac{1}{3}}\right)\right).$$

A slowly varying function $\mathcal{L}(\cdot)$ holds the following important properties.

- For $\gamma > 0$, $t^{\gamma}\mathcal{L}(t) \to \infty$ and $t^{-\gamma}\mathcal{L}(t) \to 0$ as $t \to \infty$,
- $\dfrac{\log(\mathcal{L}(t))}{\log(t)} \to 0$ as $t \to \infty$,
- For every $\gamma \in \mathbb{R}$, $(\mathcal{L}(t))^{\gamma}$ is also a slowly varying function.

Definition 2.31. A random field $\xi(x)$, $x \in \mathbb{R}^d$, is said to be long-range dependent if its covariance function $B(\cdot)$ is a non-integrable function such that

$$\int_{\mathbb{R}^d} |B(x)| dx = \infty.$$

In most of the cases, we consider the covariance function $B(\cdot)$ that is a hyperbolically decaying non-integrable function represented as $B(x) = \|x\|^{-\alpha}\mathcal{L}(\|x\|)$, where $\mathcal{L}(\cdot)$ is a slowly varying function and $\alpha \in (0, d)$.

The isotropic spectral density function $f(\cdot)$ of a long-range dependent random field takes the form

$$f(\|\lambda\|) = c_0(d, \alpha)\|\lambda\|^{\alpha-d}\mathcal{L}(1/\|\lambda\|), \quad \|\lambda\| \to 0,$$

where $c_0(d, \alpha) = \Gamma\left(\frac{d-\alpha}{2}\right)/2^{\alpha}\pi^{d/2}\Gamma\left(\frac{\alpha}{2}\right)$.

The Abelian and Tauberian theorems are utilized in probability theory and statistics to study and associate the asymptotic behaviour of the covariance function and the spectral density, see, for example, Leonenko and Olenko (2013).

Let's consider the isotropic spectral density function $f(\cdot)$ as the original function and the covariance function $B(\cdot)$ as the transform $\mathcal{T}(\cdot)$ of $f(\cdot)$ using (2.4). Then, a theorem is of the Abelian type if it deduces the asymptotic behaviour of $\mathcal{T}(f)$ from the properties of $f(\cdot)$ and a theorem is of the Tauberian type if it deduces the asymptotic behaviour of $f(\cdot)$ using the properties of $\mathcal{T}(f)$.

More precisely, the Tauberian and Abelian theorems give the relationship between the asymptotic behaviour of the covariance function $B(\cdot)$ at the infinity and the singularity properties of the isotropic spectral density $f(\cdot)$ at zero as follows.

Theorem 2.5. *Let $\mathcal{L}(\cdot)$ be a slowly varying function. Suppose that there exists an isotropic spectral density $f(z), z \in [0, \infty)$, such that $f(z)$ is decreasing for all $z \in (0, \varepsilon], \varepsilon > 0$. Then for $0 < \alpha < d$, the following statements are equivalent* *(a)* $r^{\alpha}\mathrm{B}(r) \sim \mathcal{L}(r),\ r \to \infty;$

(b) $z^{d-\alpha} f(z) \sim L\left(\frac{1}{z}\right)/c_1(d,\alpha), z \to 0,$ *where* $c_1(d,\alpha) = \dfrac{2^{\alpha}\pi^{d/2}\Gamma\left(\frac{\alpha}{2}\right)}{\Gamma\left(\frac{d-\alpha}{2}\right)}.$

For example, when $d = 3$, the constant $c_1(3,\alpha)$ can be simplified to the form

$$c_1(3,\alpha) = \frac{4\pi\Gamma(\alpha)\cos\left(\frac{\alpha\pi}{2}\right)}{(1-\alpha)}.$$

Example 2.6. *Consider the Bessel covariance function given by*

$$B(r) = \frac{1}{(1+r^2)^{\beta/2}}, \quad \beta > 0, r \geq 0.$$

Then, its isotropic spectral density is given by

$$f(z) = \left(\pi^{d/2} 2^{(d+\beta-2)/2}\Gamma\left(\frac{\beta}{2}\right)\right)^{-1} K_{(d-\beta)/2}(z) z^{(\beta-d)/2}, \quad z \geq 0,$$

where $K_{\nu}(\zeta)$ *is the modified Bessel function of the second kind.*

These functions satisfy the statements of Theorem 2.5.

2.8 Spherical random fields

This section discusses basic material related to the theory of spherical random fields. Spherical random fields are the random fields defined on a sphere. This thesis mainly considers random fields on the unit sphere. Therefore, in most of the cases $r = 1$ is used. Most of the materials included in this section are based on Marinucci and Peccati (2011) and Yadrenko (1983).

For two points, $P = (\theta_1, \varphi_1)$ and $Q = (\theta_2, \varphi_2)$ on the sphere $s_2(r)$, let Θ denote the angle formed between them originating from the centre/origin of the sphere. Then we call, Θ as the angular distance between the two points P and Q in $s_2(r)$.

Definition 2.32. A random function $\tilde{T}(\omega, x)$, $x \in s_2(r)$, defined on a complete probability space $(\Omega, \mathcal{F}, P)$ is called a spherical random field defined on the sphere $s_2(r)$,

$$\tilde{T}(x) = \tilde{T}(\omega, x) = T(r,\theta,\varphi), \quad x \in s_2(r), \theta \in [0,\pi], \varphi \in [0, 2\pi), r > 0.$$

Here, the notation $\tilde{T}(x) = T(r,\theta,\varphi)$, $x \in \mathbb{R}^3$ will be used to indicate the dependence of the random field on Euclidean coordinates. In the following, we will consider a mean-square continuous, real-valued spherical random field $T(r, \theta, \varphi)$, with zero-mean and finite second-order moments.

Definition 2.33. Let $\tilde{T}(x)$, $x \in s_2(r)$, be a real-valued spherical random field. The spherical field $\tilde{T}(\cdot)$ is called strongly isotropic if, for all $k \in \mathbb{N}$, $x_1, \ldots, x_k \in s_2(r)$ and $g \in SO(3)$, the joint distributions of the random variables $\tilde{T}(x_1), \ldots, \tilde{T}(x_k)$ and $\tilde{T}(gx_1), \ldots, \tilde{T}(gx_k)$ possesses the same law.

Further, the spherical field $\tilde{T}(\cdot)$ is called 2-weakly isotropic if for any rotation $g \in SO(3)$ and $x, x_1, x_2 \in s_2(r)$, the following holds

$$E\left(\tilde{T}(x)\right) = E\left(\tilde{T}(gx)\right), \quad E\left(\tilde{T}(x_1)\tilde{T}(x_2)\right) = E\left(\tilde{T}(gx_1)\tilde{T}(gx_2)\right).$$

In other terms, we say that a real-valued spherical field $T(r, \theta, \varphi)$ with zero-mean is isotropic, if the the covariance function $E\left(\tilde{T}(r, \theta_1, \varphi_1)\tilde{T}(r, \theta_2, \varphi_2)\right) = B(\cos\Theta)$, depends only on the angular distance Θ between two points.

An isotropic real-valued random field $\tilde{T}(\cdot)$ with zero-mean can be expanded as a Laplace series in the mean square sense. That is, $\tilde{T}(\cdot)$ admits the following spectral representation

$$T(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(\theta, \varphi) a_l^m(r), \tag{2.5}$$

where the functions $Y_l^m(\theta, \varphi)$ represent the spherical harmonics defined in (2.1). The random variables $a_l^m(r)$ are derived through the inversion arguments as the mean-square stochastic integrals given by

$$a_l^m(r) = \int_0^{\pi} \int_0^{2\pi} T(r, \theta, \varphi)\overline{Y_l^m(\theta, \varphi)} r^2 \sin\theta d\theta d\varphi = \int_{s_2(1)} \tilde{T}(ru)\overline{Y_l^m(u)}\sigma_1(du), \tag{2.6}$$

where $-l \leq m \leq l$, $l \in \mathbb{N}_0$, $u = \frac{x}{\|x\|} \in s_2(1)$ and $r = \|x\|$.

Proposition 2.1. *The spherical random field $T(r, \theta, \varphi)$ is isotropic if there exist a sequence $\{C_l(r), l \in \mathbb{N}_0\}$ of non-negative real functions such that*

$$E\left(a_l^m(r)\overline{a_{l'}^{m'}(r)}\right) = \delta_l^{l'} \delta_m^{m'} C_l(r), \quad E|a_l^m(r)|^2 = C_l(r), \tag{2.7}$$

where $-l \leq m \leq l$, $-l' \leq m' \leq l'$ and $l, l' \in \mathbb{N}$.

The sequence $\{C_l(r), l \in \mathbb{N}_0\}$ is called the angular power spectrum of the isotropic random field $T(r, \theta, \varphi)$.

Then, from the equations (2.5), (2.6), (2.7) and the addition theorem for spherical harmonics

$$Cov(T(r,\theta_1,\varphi_1),T(r,\theta_2,\varphi_2)) = \frac{1}{4\pi}\sum_{l=0}^{\infty}(2l+1)C_l(r)P_l(\cos\Theta),$$

where $\sum_{l=0}^{\infty}(2l+1)C_l(r) < \infty$ for every fixed $r > 0$.

Then, a homogeneous, isotropic, mean-square continuous spherical random field with zero-mean possesses the spectral representation given by

$$\widetilde{T}(x) = T(r,\theta,\varphi) = \sum_{l=1}^{\infty}\sum_{m=-l}^{l} Y_l^m(\theta,\varphi)a_l^m(r),$$

where

$$a_l^m(r) = \sqrt{2}\pi\int_0^{\infty}\frac{J_{l+\frac{1}{2}}(\mu r)}{\sqrt{(\mu r)}}Z_l^m(d\mu).$$

Here, Z_l^m, $-l \le m \le l$, $l \in \mathbb{N}$, is a family of complex-valued random measures on $(\mathbb{R}_+, (\mathcal{B}(\mathbb{R}_+))$ satisfying

$$\mathrm{E}\left(Z_l^m(A)\right) = 0, \quad \mathrm{E}\left(Z_l^m(A)\overline{Z_{l'}^{m'}}(B)\right) = \delta_l^{l'}\delta_m^{m'}G(A\cap B).$$

Chapter 3

Analysis of spherical monofractal and multifractal random fields

This chapter is based on the article, Leonenko, N., Nanayakkara, R. and Olenko, A. Analysis of spherical monofractal and multifractal random fields, which has been published in *Stochastic Environmental Research and Risk Assessment Journal*, 35:681–701, (2021).

Due to artistic reasons, the format of this paper was changed in accordance with the style of the thesis. This did not change the main contents of the paper, but gave rise to slight changes in the paper layout.

3.1 Introduction

Recent years have witnessed an enormous amount of attention, in the environmental, earth science, biological and astrophysical literature, on investigating spherical random fields. Excellent overviews of some novel geostatistics directions and applications can be found in Christakos (2017), Jeong et al. (2017), Marinucci and Peccati (2011), Porcu et al. (2018) and references therein. From a statistical point of view, random fields on Euclidean spaces is a rather well studied area. However, the majority of available results is not directly translatable to manifolds (where a sphere is an obvious first important candidate for investigations) and requires new stochastic models and tools, see, for example, Emery and Porcu (2019), Emery et al. (2019), Lang and Schwab (2015), Malyarenko (2012) and Marinucci and Peccati (2011). This research investigates multifractal properties of spherical random fields and provides practical methodology and examples of applications to actual data.

The concept of multifractality initially emerged in the context of physics. B. Mandelbrot showed the significance of scaling relations in turbulence modelling. Subsequently this concept developed to mathematical models and examining their fine scale characteristics. A multifractal pattern is a type of a fractal pattern that scales with multiple scaling rules in contrast to monofractals that have only scaling rule. A fractal dimension explores the change in characteristics with respect to the change in the scale used. In general, a multifractal scheme is a fractal scheme where its dynamics cannot be explained by a single fractal dimension. More details and references can be found in Harte (2001).

Multifractal structures are typical in nature. Multifractal models have been extensively used in the fields of geophysics, genomics, image modelling, finance, hydrodynamic turbulence, meteorology, internet traffic, etc., see references in Ruiz-Medina et al. (2008) and Anh et al. (2008). Multifractal behaviour has been discovered in stochastic processes as well, see Angulo and Esquivel (2015) and Grahovac and Leonenko (2014). Multifractal products of stochastic processes have been investigated by Mannersalo et al. (2002) with applications of time series in economics and teletraffic. New teletraffic models have been explored and random multifractal measure constructions by considering the stationarity of the processes' increments were proposed. This methodology has been first introduced in the groundwork Kahane (1987) on multiplicative chaos and T-martingales of positive type. Jaffard (1999) has shown the multifractal nature of specific Lévy processes and demonstrated that the multifractal spectra of such processes depicts a linear pattern rather than a concave pattern which has been noticed in the actual teletraffic data. Molchan (1996) has studied the multifractal properties such as scaling exponents of the structure function and the Rényi function of random cascade measures under various conditions. Multifractal analysis has been an important technique in the examination of singular measures and the multifractal spectrum of the random measures based on self-similar processes (Falconer (1994)). The main methods to construct random multifractal structures are based on stochastic processes, branching processes and binomial cascades (Riedi (2002)). The Rényi function plays an important role in the analysis of multifractal random fields. There are several scenarios where the Rényi function was computed for the one-dimensional case and time-series. However, there are very few multidimensional models where it is given in an explicit form. Leonenko and Shieh (2013) computed the Rényi function for three classes of multifractal random fields on the sphere. It showed some major schemes with regard to the Rényi function which reveal the multifractality of random fields that are homogeneous and isotropic.

The cosmic microwave background radiation (CMB) is the radiation from the universe since 380,000 years from the Big Bang. This elongated time period is very short compared to the age of the universe which is of 14 billion years. The CMB is an electromagnetic radiation residue from it's earliest stage. The CMB depicts variations which corresponds to different regions and represents the roots for all future formation including the solar system, stars and galaxies. At the beginning, the universe was very hot and dense and formation of atoms was impossible. The atoms were split as electrons and protons. That time the universe constituted of a plasma or ionised gas. Then the universe started to expand and cool down. Thus, it had been possible for the atoms to reconcile. This phenomenon is known as "Epoch of combination" and since that time photons have been able to move freely escaping from the opaque of the early universe. The first light which eliminated from this process is termed as the cosmic microwave background, see Planck Satellite (2021).

In 2009, the European Space Agency launched the mission Planck to study the CMB. The frequency range captured by the Planck is much wider and its resolution is higher than that of the previous space mission WMAP. The CMB's slight variations were measured with a high precision, see European Space Agency (2021b). One of the aims of the mission Planck was to verify the standard model of cosmology using this achieved greater resolution and to find out fluctuations from the specified standard model of cosmology. According to the standard model of the CMB, the universe is homogeneous and isotropic. This means that almost every part of the universe has very similar properties and that they do not differ based on the direction of the space. However, various research argue that it's not the case, see Hill (2018), Kogut et al. (1996), Marinucci (2004), Minkov et al. (2019), Novikov et al. (2000) and Starck et al. (2004). The motivation of this chapter is to develop several multifractal models and the corresponding statistical methodology and use them and other existing models to study whether the CMB data has a multifractal behaviour.

The aim of this chapter is to present and study three known multifractal models for random fields defined on the sphere and suggest several simpler models for which the Rényi function can be explicitly computed.

The first novelty of the obtained results is that for all these models, we derive singularity spectrum and study dependence of their Rényi functions on the scaling parameter. We provide several plots that illustrate typical multifractal behaviour of the models. Note, that even for the three known models their singularity spectrum was not computed and analysed before. Secondly, in Section 3.5 and the example in Section 3.7 we demonstrate the direct probability approach that can be employed to check whether assumptions on models' parameters guarantee the form of the Rényi function. This approach is less general than the one that is based on martingales for $q \in [1, 2]$, see Leonenko and Shieh (2013) and Mannersalo et al. (2002). The advantage is that the proposed methodology is simple and can also be used for $q > 2$. Third, Section 3.7 suggests four simple new models, explicitly computes their singularity spectrum and Rényi functions and investigates their properties. Finally, we discuss the methodology of computing the Rényi functions and provide various numerical studies of the actual CMB data.

The proposed models and methodology can find various applications to other spherical data. The obtained results and discussion in the chapter provide detailed guidance how the multifractal modelling and analysis can be done for general spherical data. It could be very useful for various earth, environmental and image analysis problems. In particular, the recent paper by Fryer and Olenko (2019) discusses methodology and provides R code for transforming various spherical and directional data to the HEALPix format of the CMB data. Then, the results of this research can be directly applied.

The plan of the chapter is as follows. Section 3.2 provides main notations and definitions related to the theory of random fields. Section 3.3 introduces spherical random fields. Section 3.4 gives results related to the theory of multifractality and the Rényi function. Section 3.5 describes the direct probability approach to get the conditions on the limit random measure μ. Section 3.6 provides results for three known models of the Rényi function for spherical random fields of the exponential type. Section 3.7 proposes new models based on power transformations of Gaussian fields. Section 3.8 presents numerical studies including computing and fitting the empirical Rényi functions for CMB data from different sky windows and models. The conclusions and some new problems are given in Section 3.9. Finally, proofs of all key results can be found in Section 3.10.

All numerical studies were conducted by using Maple 2019.0 and R 3.6.3 software, in particular, the R packages 'rcosmo' (Fryer et al. (2020), Fryer et al. (2019)) and 'RandomFields' (Schlather et al. (2019)). A reproducible version of the code in this chapter is available in Appendix A.

3.2 Main notations and definitions

This section presents background materials in the random fields theory and multifractal analysis methodology. Most of the material included in this and next two sections are based on Lang and Schwab (2015), Leonenko (1999), Malyarenko (2012), Mannersalo et al. (2002) and Marinucci and Peccati (2011).

Let $S \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a multidimensional set, $\|\cdot\|$ denote the Euclidean distance in $\mathbb{R}^n$, $s_{n-1}(1) = \{u \in \mathbb{R}^n : \|u\| = 1\}$, and $SO(n)$ be the group of rotations in $\mathbb{R}^n$. The notation $|\cdot|$ will be used for the Lebesgue measure on $\mathbb{R}^n$. $\{\cdot\} \stackrel{d}{=} \{\cdot\}$ will stand for the equality of finite dimensional distributions.

The Kronecker delta is a function defined as:

$$\delta_i^j = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$$

For $\nu > -\frac{1}{2}$, we use the Bessel function of the first kind of order ν

$$J_\nu(z) = \sum_{m=0}^{\infty} (-1)^m \left(\frac{z}{2}\right)^{2m+\nu} [m!\Gamma(m+\nu+1)]^{-1}, \quad z > 0,$$

where $\Gamma(\cdot)$ is the gamma function.

Definition 3.1. A random field is a function $\xi(\omega, x) : \Omega \times S \to \mathbb{R}^m$ such that $\xi(\omega, x)$ is a random vector for each $x \in S$. For simplicity it will also be denoted by $\xi(x)$, $x \in S$.

When $n = 1$, $\xi(x)$ is a random process. When $S \subseteq \mathbb{R}^n$, $n > 1$, then $\xi(x)$ is termed as a random field. It is called a vector random field for $m > 1$. In this chapter, we concentrate on scalar random fields $\xi(x)$, $x \in S$, $n > 1$ and $m = 1$.

If $\{\xi(x_1), ..., \xi(x_N), x_1, ..., x_N \in S\}$ is a set of random variables belonging to a Gaussian system for each $N \geq 1$, then $\xi(x), x \in S$, is called Gaussian.

We assume that all random variables $\xi(x)$ are defined on the same probability space $(\Omega, \mathcal{F}, P)$.

Definition 3.2. A second order random field is a random function $\xi : S \to L_2(\Omega, \mathcal{F}, P)$, $S \subset \mathbb{R}^n$.

In other words, the random variables $\xi(x)$, $x \in S$, satisfy $E|\xi(x)|^2 < +\infty$. Thus, a second order random field over S is a family $\{\xi(x), x \in S\}$ of square integrable random variables.

Definition 3.3. A second order random field $\xi(x), x \in \mathbb{R}^n$, is homogeneous (in the wide sense) if its mathematical expectation $m(x) = E[\xi(x)]$ and covariance function $B(x, y) = cov(\xi(x), \xi(y))$ are invariant with respect to the Abelian group $G = (\mathbb{R}^n, +)$ of shifts in $\mathbb{R}^n$, that is

$$m(x) = m(x + \tau),\ B(x, y) = B(x + \tau, y + \tau),$$

for any $x, y, \tau \in \mathbb{R}^n$.

That is, for homogeneous random fields $E[\xi(x)] = const$, and the covariance function $B(x, y) = B(x - y)$ depends only on the difference $x - y$.

The covariance function $B(x - y)$ of a homogeneous random field is a non-negative definite kernel on $\mathbb{R}^n \times \mathbb{R}^n$, that is, for any $r \geq 1$, $x^{(j)} \in \mathbb{R}^n$, $z_j \in \mathbb{C}$, $j = 1, ..., r$,

$$\sum_{i,j=1}^{r} B(x^{(i)} - x^{(j)}) z_i \bar{z_j} \geq 0.$$

If the covariance function $B(x)$ is continuous at $x = 0$, then the field is mean-square continuous for each $x \in \mathbb{R}^n$ and vice versa.

Definition 3.4. The second order random field $\xi(x)$ is isotropic (in the wide sense) on $\mathbb{R}^n$ if its mathematical expectation and covariance function are invariant with respect to the group of rotations on the sphere, i.e.

$$m(x) = m(gx), \; B(x, y) = B(gx, gy),$$

for every $x, y, \tau \in \mathbb{R}^n$ and $g \in SO(n)$.

In the following we will be considering real-valued second order random fields. It will be also assumed that $E[\xi(x)] = 0$ without loss of generality.

If a real-valued second order random field $\xi(x), x \in \mathbb{R}^n$ is homogeneous and isotropic, then its mathematical expectation and the covariance function depend only on the Euclidean distance $\rho_{xy} = \|x-y\|$ between x and y. It means that its mathematical expectation $m(x)$ and covariance function $B(x, y)$ are invariant with respect to shifts, rotations and reflections in $\mathbb{R}^n$.

Definition 3.5. A stochastic process $\{X(t), t \geq 0\}$ is self-similar if for any non-random constant $a > 0$, there exists non-random constant $b > 0$ such that $\{X(at)\} \stackrel{d}{=} \{bX(t)\}$.

For self-similar, continuous at 0 and non-trivial $X(t)$, the constant b must be equal a^H, $a > 0$, where $H \geq 0$. Thus, $\{X(at)\} \stackrel{d}{=} \{a^H X(t)\}$. The constant H is known as the Hurst parameter. The process $\{X(t), t \geq 0\}$ is called *H-ss* (self-similar) or *H-sssi* (self-similar stationary increments) if its increments are stationary.

The concept of multifractal processes was motivated by establishing the following scaling rule of self-similar processes.

Definition 3.6. A stochastic process $X(t)$ is multifractal if it holds $\{X(ct)\} \stackrel{d}{=} \{M(c)X(t)\}$, where $M(c)$ is a random variable independent of $X(t)$ for every $c > 0$ and the distribution of $M(c)$ does not depend on t.

The process is self-similar if $M(c)$ is non-random for every $c > 0$ and $M(c) = c^H$. The scaling factor $M(c)$ satisfies $\{M(ab)\} \stackrel{d}{=} \{M_1(a)M_2(b)\}$ for every selection of constants a and b and random M_1 and M_2 that are independent copies of M. This establishes the characteristic of the deterministic factor *H-ss* processes $(ab)^H = a^H b^H$.

Another definition of multifractality is

Definition 3.7. A stochastic process $X(t)$ is multifractal if there exist non-random functions $c(q)$ and $\tau(q)$ such that for all $t, s \in \mathcal{T}, q \in \mathcal{Q}$,

$$E|X(t) - X(s)|^q = c(q)|t - s|^{\tau(q)},$$

where $\mathcal{T}$ and $\mathcal{Q}$ are intervals on the real line with positive length and $0 \in \mathcal{T}$.

The function $\tau(q)$ is known as the scaling function. The interval $\mathcal{Q}$ may include negative values. Instead of the increments of the process, the definition can also be established on the moments of the process. i.e. $E|X(t)|^q = c(q)t^{\tau(q)}$. Above definitions coincide if the increments are stationary. If $\{X(t)\}$ is H-*sssi*, then it holds that $\tau(q) = Hq$.

3.3 Spherical random fields

This section introduces some basic notations of the theory of random fields on a sphere. The sphere is a simplest case of a manifold in $\mathbb{R}^n$. For simplicity, we consider only the case $n = 3$.

Let us denote the 3-dimensional unit ball as $B^3 = \{x \in \mathbb{R}^3 : \|x\| \leq 1\}$. The spherical surface in $\mathbb{R}^3$ with a given radius $r > 0$ is $s_2(r) = \{x \in \mathbb{R}^3 : \|x\| = r\}$, with the corresponding Lebesgue measure on the sphere $\sigma_r(du) = \sigma_r(d\theta \cdot d\varphi) = r^2 \sin\theta d\theta d\varphi$, $(\theta, \varphi) \in s_2(1)$. For two points on $s_2(r)$ we use Θ to denote the length of the angle formed between two rays originating at the origin and pointing at these two points. Θ is called the angular distance between these two points.

A spherical random field $T = \{T(r, \theta, \varphi) : 0 \leq \theta \leq \pi, 0 \leq \varphi < 2\pi, r > 0\}$ is a random function, which is defined on the sphere $s_2(r)$. We deal with a spherical real-valued mean-square continuous random field T with a constant mean and finite second order moments.

Definition 3.8. A real-valued second order random field $T(x), x \in s_2(r)$, with $E[T(x)] = 0$ is isotropic if $E[T(x_1)T(x_2)] = B(\cos\Theta)$, $x_1, x_2 \in s_2(r)$, depends only on the angular distance Θ between x_1 and x_2.

For the considered mean-square continuous isotropic random fields, the covariance function $B(\cos\Theta)$ is a continuous function on $[0, \pi)$.

An isotropic spherical random field on $s_2(r)$ can be expanded in a Laplace series in the mean-square sense.

$$T(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(\theta, \varphi) a_l^m(r), \tag{3.1}$$

where $\{Y_l^m(\theta, \varphi)\}$ represents the spherical harmonics defined as

$$Y_l^m(\theta, \varphi) = c_l^m \exp{(im\varphi)} P_l^m(\cos\theta), \quad l = 0, 1, ..., \ m = 0, \pm 1, ..., \pm l,$$

with

$$c_l^m = (-1)^m \left[\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2},$$

the associated Legendre polynomials $P_l^m(\cdot)$ having degree l and order m

$$P_l^m(x) = (-1)^m \left(1-x^2\right)^{m/2} \frac{d^m}{dx^m} P_l(x),$$

and the l-th Legendre polynomials, see Leonenko and Shieh (2013),

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} \left(x^2-1\right)^l.$$

The spherical harmonics have the following properties

$$\int_0^\pi \int_0^{2\pi} \overline{Y_l^m(\theta,\varphi)} Y_{l'}^{m'}(\theta,\varphi) \sin\theta d\varphi d\theta = \delta_l^{l'} \delta_m^{m'},$$

$$\overline{Y_l^m(\theta,\varphi)} = (-1)^m Y_l^{(-m)}(\theta,\varphi).$$

The notation $\tilde{T}(x) = T(r,\theta,\varphi), x \in \mathbb{R}^3$, will be used to highlight the random field's dependence on Euclidean coordinates.

The random coefficients of the Laplace series can be computed as the mean-square stochastic integrals via the inversion arguments as

$$a_l^m(r) = \int_0^\pi \int_0^{2\pi} T(r,\theta,\varphi) \overline{Y_l^m(\theta,\varphi)} r^2 \sin\theta d\theta d\varphi = \int_{s_2(1)} \tilde{T}(ru) \overline{Y_l^m(u)} \sigma_1(du), \tag{3.2}$$

where $u = \frac{x}{\|x\|} \in s_2(1), \quad r = \|x\|$.

The covariance functions $E(T(r,\theta,\varphi)T(r,\theta',\varphi'))$ of the isotropic random fields depend only on the angular distance $\Theta = \Theta_{PQ}$ between the points $P = (r,\theta,\varphi)$ and $Q = (r,\theta',\varphi')$. For spherical isotropic random fields it possesses

$$E a_l^m(r) \overline{a_{l'}^{m'}}(r) = \delta_l^{l'} \delta_m^{m'} C_l(r), \quad E|a_l^m(r)|^2 = C_l(r), \quad m = 0, \pm 1, ..., \pm l. \tag{3.3}$$

The angular power spectrum of the isotropic random field $T(r,\theta,\varphi)$ is defined as the functional series $\{C_0(r), C_1(r), ..., C_l(r), ...\}$.

From (3.1) - (3.3) and the addition theorem for spherical harmonic functions we obtain

$$Cov(T(r,\theta,\varphi), T(r,\theta',\varphi')) = \frac{1}{4\pi} \sum_{l=0}^{\infty} (2l+1) C_l(r) P_l(\cos\Theta),$$

where for every $r > 0$ it holds $\sum_{l=0}^{\infty}(2l+1)C_l(r) < \infty$.

If $T(r,\theta,\varphi)$ is an isotropic Gaussian field defined on the sphere $s_2(r)$, then the coefficients $a_l^m(r)$ are independent Gaussian random variables that are complex-valued with $Ea_l^m(r) = 0$.

For the homogeneous and isotropic random field $\tilde{T}(x)$, $x \in \mathbb{R}^3$, it holds (Marinucci and Peccati (2011))

$$Ea_l^m(r)\overline{a_{l'}^{m'}}(s) = \delta_l^{l'}\delta_m^{m'}C_l(r,s), \ r > 0, s > 0,$$

where

$$C_l(r,s) = 2\pi^2 \int_0^\infty \frac{J_{l+\frac{1}{2}}(\mu r)J_{l+\frac{1}{2}}(\mu s)}{(\mu r)^{1/2}(\mu s)^{1/2}} G(d\mu), \quad l = 0,1,2,...,$$

and $G(\cdot)$ is a finite measure defined on the Borel sets of $[0,\infty)$ satisfying

$$\sigma^2 = Var\{\tilde{T}(x)\} = \int_0^\infty G(d\mu) < \infty, \quad x \in \mathbb{R}^3.$$

3.4 Rényi function and multifractal spectrum

This section introduces basic notations, definitions and concepts regarding the multifractal theory and Rényi functions.

The Rényi function which is also known as the index of diversity is used in multifractal analysis to assess the randomness of many natural phenomena. It can be used to detect the multifractal behaviour of a given random process. The Rényi function computes how the measure/mass/intensity on a surface varies with the resolution or the block size of an image. That is, it calculates the change in detail of a pattern according to the change in scale. The Rényi function characterises the distortion in the mean of a pattern's probability distribution of pixel values. Rényi functions of non-fractal and monofractal processes exhibit a flatter curve than ones of multifractal processes. Rényi functions of multifractal processes typically have quadratic shapes that suggest the presence of different fractal dimensions.

Consider a random field $\Lambda(x,\omega)$, $x \in \mathbb{R}^3, \omega \in \Omega$, that is measurable, homogeneous and isotropic (HIRF) on the 3-dimensional Euclidean space $\mathbb{R}^3$. It will be called the mother field. For simplicity it will be denoted as $\Lambda(x) = \Lambda(x,\omega)$.

Condition 3.1. *Let a random field* $\Lambda(x)$, $x \in \mathbb{R}^3$, *satisfy*

$$E[\Lambda(x)] = 1, \quad \Lambda(x) > 0, \quad Cov(\Lambda(x), \Lambda(y)) = R_\Lambda(\|x - y\|) = \sigma_\Lambda^2 \rho_\Lambda(\|x - y\|),$$

where $\rho(0) = 1$ *and* $\sigma_\Lambda^2 < \infty$.

Let $\Lambda^{(i)}(x), x \in \mathbb{R}^3, i = 0, 1, 2, ...,$ be a sequence of independent copies of the random field $\Lambda(\cdot)$. We consider the re-scaling of $\Lambda(\cdot)$ defined as $\Lambda^{(i)}(b^i x)$, where $b > 1$ is a constant called a scaling factor and $b^i x$ is the product of a vector x by a scalar b^i.

A finite-product field on B^3 is defined by

$$\Lambda_k(x) = \prod_{i=0}^{k} \Lambda^{(i)}(b^i x), \quad k = 1, 2, ...$$

Then one can introduce the random measure $\mu_k(\cdot)$ on the Borel σ-algebra $\mathcal{B}$ of a unit ball B^3 by

$$\mu_k(A) = \int_{y \in A} \Lambda_k(y) dy, \ A \in \mathcal{B}, \ k = 0, 1, 2, ...,$$

where $\mu_k(A)$ is defined in L_p sense (Denisov and Leonenko (2016)).

We denote by $\mu_k \xrightarrow{d} \mu, \ k \to \infty$, the weak almost surely convergence of the measures μ_k to some measure μ. It means that for all continuous functions $g(y), y \in B^3$, it holds with probability 1 that

$$\int_{B^3} g(y) \mu_k(dy) \to \int_{B^3} g(y) \mu(dy), \ k \to \infty,$$

where the integrals are defined in L_p sense (Denisov and Leonenko (2016)).

Remark 3.1. The weak almost surely convergence of random measures implies that for a finite or countable family of sets A_j from $\mathcal{B}$, with probability 1,

$$\mu_k(A_j) \to \mu(A_j), \quad k \to +\infty,$$

for all j, see Kahane (1987) and Mannersalo et al. (2002). Moreover, it was shown in Denisov and Leonenko (2016), Leonenko and Shieh (2013) and Mannersalo et al. (2002) that for mother random fields with $\rho_\Lambda(r)$ possessing an exponentially decaying bound, the random variables $\mu_k(B^3)$ converge to $\mu(B^3)$ in L_2 (and hence in L_q for $q \in (0, 2]$) when $k \to +\infty$. In the following, for all models considered in this chapter, it will be assumed that $|\rho_\Lambda(r)| \leq Ce^{-\gamma r}$ for some positive constants C, and γ.

Definition 3.9. The Rényi function of a random measure μ is a non-random function defined by

$$T(q) = \liminf_{m\to\infty} \frac{\log_2 E\sum_l \mu(B_l^{(m)})^q}{\log_2 |B_l^{(m)}|},$$

where $\{B_l^{(m)}, l = 0, 1, ..., 2^m - 1,\ m = 1, 2, ...,\}$ denotes the mesh formed by the m^{th} level dyadic decomposition of the unit ball B^3.

The key result about the form of the Rényi function is the following theorem.

Theorem 3.1. (Leonenko and Shieh (2013)) *Suppose that Condition 3.1 holds.*

(i) Assume that the correlation function $\rho_\Lambda(\|x-y\|) = \rho(r)$ of the field $\Lambda(\cdot)$ satisfies the following condition

$$|\rho_\Lambda(r)| \le Ce^{-\gamma r},\ r > 0, \tag{3.4}$$

for some positive constants C and γ. Then, for the scaling factor $b > \sqrt[3]{1+\sigma_\Lambda^2}$, the measures $\mu_k \xrightarrow{d} \mu$, $k \to \infty$, on B^3.

(ii) If for some range $q \in Q = [q_-, q_+]$, both $E^q\Lambda(0) < \infty$ and $E\mu^q(B^3) < \infty$, then the Rényi function $T(q)$ of μ is given by

$$T(q) = q - 1 - \frac{1}{3}\log_b E\Lambda^q(0), \quad q \in Q.$$

Similarly, for spherical random fields on $s_2(1)$, one can introduce an analogous approach.

Condition 3.2. *Let the random field $\tilde{\Lambda}(x)$, $x \in s_2(1)$, satisfy*

$$E\tilde{\Lambda}(x) = 1, \quad Var\tilde{\Lambda}(x) = \sigma_{\tilde{\Lambda}}^2 < \infty, \quad \tilde{\Lambda}(x) > 0,$$

$$Cov(\tilde{\Lambda}(\theta,\varphi), \tilde{\Lambda}(\theta',\varphi')) = \frac{1}{4\pi}\sum_{l=0}^{\infty}(2l+1)C_l P_l(\cos\theta), \quad \sum_{l=0}^{\infty}(2l+1)C_l < \infty.$$

Let $\tilde{\Lambda}^{(i)}(x)$, $x \in s_2(1)$, $i = 0, 1, 2, ...$, be a sequence of independent copies of the field $\tilde{\Lambda}(\cdot)$. Let us use the following spherical coordinate notations for points on $s_2(1) : x = (1, \theta, \varphi) \in s_2(1)$. Consider $\tilde{\Lambda}^{(i)}(b^i \times x)$, where $b > 1$ is a scaling factor, $b^i \times x := (1, b^i \underset{\pi}{\times} \theta, b^i \underset{2\pi}{\times} \varphi)$ $\in s_2(1)$, and the modulus algebra is used to compute the products $b^i \underset{\pi}{\times} \theta$ and $b^i \underset{2\pi}{\times} \varphi$.

Define the finite product fields on $s_2(1)$ by

$$\tilde{\Lambda}_k(x) = \prod_{i=0}^{k} \tilde{\Lambda}^{(i)}(b^i \times x), \quad k = 1, 2, ...$$

Let us introduce the random measure $\mu_k(\cdot)$ on the Borel σ-algebra $\mathcal{B}$ of $s_2(1)$ as

$$\mu_k(A) = \int_A \tilde{\Lambda}_k(y)dy, \ k = 0, 1, 2, ..., \ A \in \mathcal{B}, \tag{3.5}$$

where $\mu_k(A)$ is defined in L_p sense (Denisov and Leonenko (2016)).

We denote by $\mu_k \xrightarrow{d} \mu$, $k \to \infty$, the weak convergence of the measures μ_k to some non-degenerate measure μ. It means that for all continuous functions $g(y), y \in s_2(1)$,

$$\int_{s_2(1)} g(y)\mu_k(dy) \to \int_{s_2(1)} g(y)\mu(dy), \quad k \to \infty.$$

The Rényi function of the random measure μ defined on $s_2(1)$ is defined as

$$T(q) = \liminf_{m\to\infty} \frac{\log_2 E \sum_l \mu(S_l^{(m)})^q}{\log_2 |S_l^{(m)}|}, \tag{3.6}$$

where $\{S_l^{(m)}, l = 0, 1, ..., 2^m - 1\}$ is the mesh constructed by m^{th} level dyadic decomposition of the spherical surface of $s_2(1)$.

Theorem 3.2. (Leonenko and Shieh (2013)) *Suppose that Condition 3.2 holds and the isotropic random field $\tilde{\Lambda}(\cdot)$ is the restriction to the sphere $s_2(1)$ of the HIRF $\Lambda(x), x \in \mathbb{R}^3$, with the correlation function $\rho_\Lambda(\|x - y\|) = \rho(r)$. Under similar assumptions to Theorem 3.1, the Rényi function $T(q)$ of the limit measure μ on $s_2(1)$ is given by*

$$T(q) = q - 1 - \frac{1}{2}\log_b E\Lambda^q(0), \quad q \in Q.$$

Remark 3.2. If x and y are two locations on the unit sphere $s_2(1)$ and Θ is the angle between them, then the Euclidean distance between these two points is $2\sin(\Theta/2)$, which gives a direct correspondence between the covariance function $\rho_\Lambda(\|x-y\|)$ in the Euclidean space and the covariance function $\rho(\cos\Theta) = \rho_\Lambda(2\sin\Theta/2)$ on the sphere. Thus, the restriction of the HIRF $\Lambda(x)$ to $s_2(1)$ is an isotropic spherical random field.

The multifractal or singularity spectrum is defined via the Legendre transform as

$$f(h) = \inf_q (hq - T(q)). \tag{3.7}$$

and is used to describe local fractal dimensions of random fields.

3.5 Conditions on measure μ

The random measure μ in the previous section was defined as a weak limit of the measures μ_k. Therefore, it would be difficult to check moment conditions on μ as its probability distribution is not explicitly known. This section provides sufficient conditions on the scaling factor b and the variance σ_Λ^2 that guarantee $E\mu^q(B^3) < \infty$. The general method to obtain such conditions for the range $q \in [1, 2]$ uses martingale L^2 convergence, see, for example, Mannersalo et al. (2002). The proof of the main result of this section in Section 3.10 demonstrates the direct probability approach, which is more elementary.

Theorem 3.3. *Let the mother field $\Lambda(x) > 0$, $x \in \mathbb{R}^3$, satisfy the conditions*

$$E\Lambda(x) = 1, \quad Var\Lambda(x) = \sigma_\Lambda^2 < +\infty, \quad Cov(\Lambda(x), \Lambda(y)) = \sigma_\Lambda^2 \rho_\Lambda(\|x - y\|),$$

$$|\rho_\Lambda(\tau)| \leq Ce^{-\gamma\tau}, \; \tau > 0,$$

and the scaling factor $b > \max(\sqrt[3]{1 + \sigma_\Lambda^2},\, e^{\frac{\sigma_\Lambda^2 C}{3}})$.

Then the measures $\mu_k \xrightarrow{d} \mu$, $k \to \infty$, and $E\mu^q(B^3) < +\infty$, for $q \in [1, 2]$.

Remark 3.3. The direct probability approach can be used to obtain conditions on the mother field that guarantee $E\mu_k^q(B^3) < +\infty$, for q in the range $[1, Q]$, where $Q > 2$.

For example, using the Lyapunov's inequality for $q \in [1, 4]$, see (Loéve, 1977, p.162),

$$E\mu_k^q(B^3) \leq (E\mu_k^4(B^3))^{q/4},$$

the conditions on b and σ_λ^2 that guarantee $E\mu_k^4(B^3) < +\infty$ are also sufficient for $E\mu^q(B^3) < +\infty$, $q \in [1, 4]$. Then, it follows from

$$E\mu_k^4(B^3) = \int_{B^3}\int_{B^3}\int_{B^3}\int_{B^3} \prod_{i=0}^{k} E\left(\prod_{j=1}^{4} \Lambda^{(i)}(y_i b^i)\right) \prod_{j=1}^{4} dy_i,$$

that one can impose some additional assumptions on the fourth order moments $E(\prod_{j=1}^4 \Lambda(y_j b^i))$ or cumulants of the mother field $\Lambda(\cdot)$. We will provide an example of such conditions in Section 3.7 for Model 4.

3.6 Rényi functions of exponential models

For the random fields on the sphere, there are three models where the Rényi function is known explicitly, see Leonenko and Shieh (2013). These models were obtained for exponential type spherical random fields. This section introduces these models, derives their singularity spectrum and studies dependence of their Rényi functions on the scaling parameter.

Model 1 Let the random field $\Lambda(x)$ be given as

$$\Lambda(x) = exp\left\{Y(x) - \frac{1}{2}\sigma_Y^2\right\},$$

where $Y(x)$, $x \in \mathbb{R}^3$, is a zero-mean Gaussian, measurable, separable random field with the covariance function $\sigma_Y^2 \rho_Y(r)$, $\rho_Y(0) = 1$.

The following result provides the conditions and the explicit form of the Rényi function for Model 1.

Theorem 3.4. (Leonenko and Shieh (2013)) *Let for Model* 1 *the correlation function satisfy*

$$0 < |\rho_Y(r)| \leq Ce^{-\gamma r},\ r > 0,$$

for some positive C *and* γ *and* $b > exp\{\frac{\sigma_Y^2}{3}\}$.

If $Y(x)$, $x \in s_2(1)$, *is a spherical isotropic random field that is a restriction of* $Y(x)$, $x \in \mathbb{R}^3$, *on the sphere* $s_2(1)$, *then the random measures* (3.5) *generated by the spherical fields* $\tilde{\Lambda}(x) = exp\left\{Y(x) - \frac{1}{2}\sigma_Y^2\right\}$, $x \in s_2(1)$, *converge weakly to the random measure* μ. *The corresponding Rényi function is*

$$T(q) = q\left(1 + \frac{\sigma_Y^2}{4\ln b}\right) - q^2\left(\frac{\sigma_Y^2}{4\ln b}\right) - 1, \quad q \in [1, 2]. \tag{3.8}$$

Remark 3.4. It's easier to define the covariance structure of a random field on the whole space $\mathbb{R}^3$ and then to consider it's restriction to the sphere rather than directly defining it and the corresponding covariances on the sphere.

Model 2 Let the random field $\Lambda(x)$ be of the form

$$\Lambda(x) = exp\left\{Z(x) - c_Z\right\}, \quad c_Z = -\ln\left(1 - \frac{1}{\lambda}\right)^{\beta},$$

where $Z(x)$, $x \in \mathbb{R}^3$, is a gamma-correlated HIRF with the correlation function $\rho_Z(r)$.

The field $Z(x)$ has the marginal density

$$p(u) = \frac{\lambda^\beta}{\Gamma(\beta)} u^{\beta-1} e^{-\lambda u}, \quad u, \lambda, \beta \in (0, +\infty), \tag{3.9}$$

and the bivariate density

$$p_0(u_1, u_2; \alpha) = \frac{(u_1 u_2/\alpha)^{\frac{\beta-1}{2}}}{\Gamma(\beta)(1-\alpha)} \exp\left\{-\frac{u_1+u_2}{1-\alpha}\right\} I_{\beta-1}\left(2\frac{\sqrt{u_1 \cdot u_2 \cdot \alpha}}{1-\alpha}\right), \tag{3.10}$$

where $I_v(z) = \sum_{k=0}^{\infty} \left(\frac{z}{2}\right)^{2k+v} (k!\Gamma(k+v+1))^{-1}$ is the modified Bessel function of the first kind, $\alpha \in [0,1]$, λ, β, and γ are constant parameters.

Then the covariance function of the mother random field is

$$\rho_\Lambda(r) = \left(\frac{e^{-2c_z}}{\left(1 - \frac{2}{\lambda} + \frac{2}{\lambda^2}(1 - \rho_Z(\tau))\right)^\beta} - 1\right)\left(\frac{e^{-2c_z}}{\left(1 - \frac{1}{\lambda}\right)^\beta} - 1\right)^{-1}.$$

The following result gives the Rényi function and the corresponding conditions for Model 2.

Theorem 3.5. (Leonenko and Shieh (2013)) *Suppose that for Model 2 the parameter $\lambda > 2$ and the correlation function satisfies*

$$0 < |\rho_Z(r)| \leq Ce^{-\gamma r}, \; r > 0,$$

for some positive constants C and γ. Then, for the parameters (β, λ) from the set

$$L_{\beta,\lambda} = \left\{(\beta, \lambda) : b > \left(1 + \frac{\frac{1}{\lambda^2}}{1 - \frac{2}{\lambda}}\right)^{\frac{\beta}{2}}, \lambda > 2, \beta > 0\right\},$$

the measures $\mu_k \xrightarrow{d} \mu$, $k \to \infty$. The Rényi function of μ is given by

$$T(q) = q\left(1 - \frac{\beta}{2}\log_b\left(1 - \frac{1}{\lambda}\right)\right) + \left(\frac{\beta}{2}\right)\log_b\left(1 - \frac{q}{\lambda}\right) - 1. \tag{3.11}$$

where $q \in Q = \{0 < q < \lambda\} \cap [1,2] \cap L_{\beta,\lambda}$.

Model 3 Let the mother random field be

$$\Lambda(x) = exp\left\{U(x) - c_U\right\},\ x \in \mathbb{R}^3,$$

where $U(x) = -Z^{-1}(x)$, and $Z(x)$, $x \in \mathbb{R}^3$, is a gamma-correlated HIRF with the densities given by (3.9) and (3.10) and the correlation function $\rho_Z(r)$.

Theorem 3.6. (Leonenko and Shieh (2013)) *Suppose that for Model 3 the correlation function satisfies*

$$0 < |\rho_Z(r)| \leq Ce^{-\gamma r},\ r > 0,$$

for some positive constants C and γ. Then, for any $(\beta, \lambda) \in L_{\beta,\lambda}$ and $b > \left(\frac{\Gamma(\beta)2^{\frac{\beta}{2}-1}K_\beta(2\sqrt{2\lambda})}{\lambda^{\beta/2}[K_\beta(2\sqrt{\lambda})]^2}\right)^{\frac{1}{2}}$ the measures $\mu_k \xrightarrow{d} \mu$ when $k \to \infty$. The Rényi function of measure μ is

$$T(q) = q\left(1 + \frac{c_U}{2\ln b}\right) - \frac{1}{2}\log_b\left(q^{\beta/2}K_\beta(2\sqrt{q\lambda})\right) - \left(1 + \frac{1}{2}\log_b\left(\frac{2\lambda^{\beta/2}}{\Gamma(\beta)}\right)\right), \tag{3.12}$$

where $q \in Q = [1, 2] \cap L_{\beta,\lambda}$, $K_\lambda(x)$ is the modified Bessel function of the third kind and

$$c_U = \ln\left(\frac{2\lambda^{\beta/2}K_\beta(2\sqrt{\lambda})}{\Gamma(\beta)}\right).$$

Let $\alpha(q)$ denote the q^{th} order singularity exponent and be defined by

$$\alpha(q) = \frac{d}{dq}T(q). \tag{3.13}$$

Then the multifractal spectrum defined by (3.7) can be expressed as a function of α by

$$f(\alpha(q)) = q \cdot \alpha(q) - T(q). \tag{3.14}$$

For Model 1 it is easy to see from (3.8) that

$$\alpha(q) = 1 + \frac{\sigma_Y^2}{4\ln(b)} - \frac{\sigma_Y^2}{2\ln(b)}q,$$

$$f(\alpha(q)) = 1 - \frac{\sigma_Y^2}{4\ln(b)}q^2, \quad q \in [1, 2]. \tag{3.15}$$

By (3.11) we obtain for Model 2

$$\alpha(q) = 1 - \frac{\beta}{2}\log_b\left(1 - \frac{1}{\lambda}\right) + \frac{\beta}{2\ln(b)(q - \lambda)},$$

$$f(\alpha(q)) = 1 + \frac{\beta}{2}\left(\frac{q}{\ln(b)(q-\lambda)} - \log_b\left(1 - \frac{q}{\lambda}\right)\right). \tag{3.16}$$

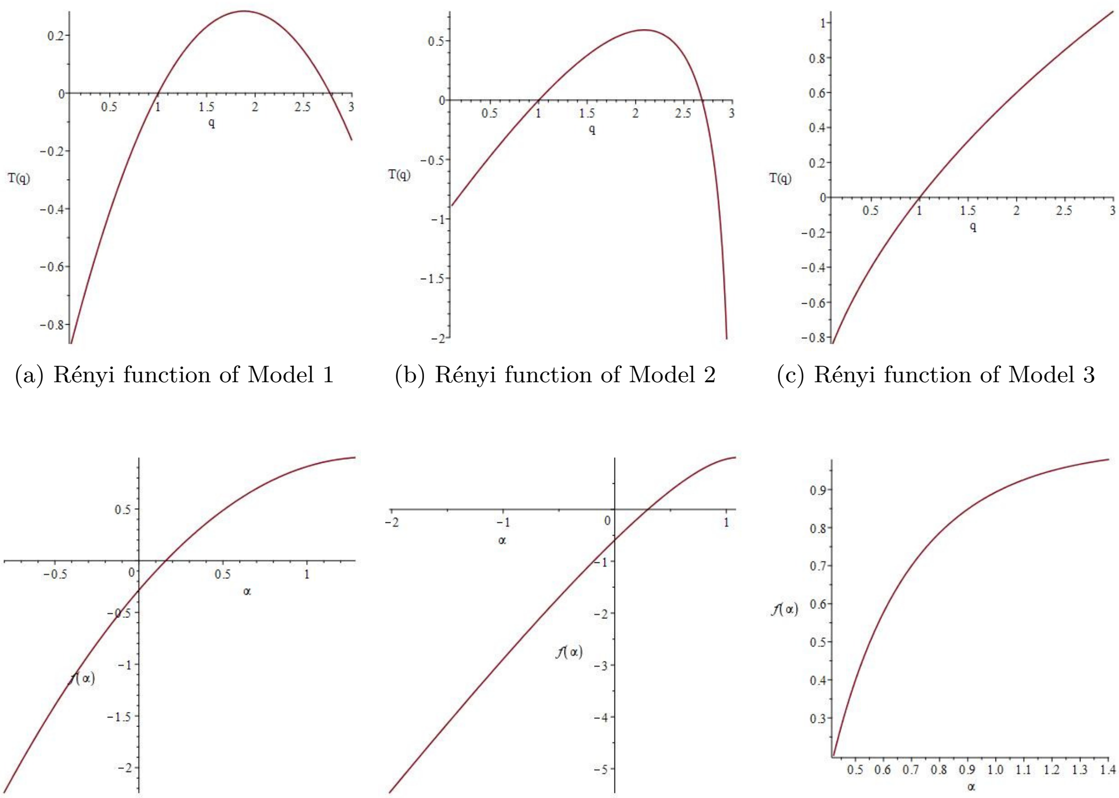
For Model 3 it follows from (3.12) and $K_\beta{}'(q) = -\frac{1}{2}\left(K_{\beta-1}(q) + K_{\beta+1}(q)\right)$, see 9.6.26 in Abramowitz and Stegun (1964), that

$$\alpha(q) = 1 + \frac{c_U}{2\ln(b)} - \frac{\beta}{4\ln(b)q} + \frac{\sqrt{\lambda}(K_{\beta-1}(2\sqrt{q\lambda}) + K_{\beta+1}(2\sqrt{q\lambda}))}{2\ln(b)K_\beta(2\sqrt{q\lambda})\sqrt{q}},$$

$$f(\alpha(q)) = 1 + \frac{\beta}{2}\log_b\left(\frac{2\lambda^{\beta/2}}{\Gamma(\beta)}\right) - \frac{\beta}{4\ln(b)} + \frac{1}{2}\log_b(q^{\beta/2}K_\beta(2\sqrt{q\lambda}))$$
$$+ \frac{\sqrt{q\lambda}(K_{\beta-1}(2\sqrt{q\lambda}) + K_{\beta+1}(2\sqrt{q\lambda}))}{2\ln(b)K_\beta(2\sqrt{q\lambda})}. \tag{3.17}$$

Summarising the above results we obtain

Theorem 3.7. *Let the corresponding conditions of Theorems* 3.4, 3.5 *and* 3.6 *are satisfied for Models* 1, 2 *and* 3. Then the multifractal spectra of these models are given by (3.15), (3.16) and (3.17) respectively.



(a) Rényi function of Model 1 (b) Rényi function of Model 2 (c) Rényi function of Model 3

(d) Spectrum of Model 1 (e) Spectrum of Model 2 (f) Spectrum of Model 3

Figure 3.1: Examples of Rényi functions and multifractal spectra for Models 1, 2 and 3

The plots shown in Figure 3.1 illustrate behaviours of the Rényi functions and multifractal spectra for Models 1, 2 and 3. For these numerical examples, we used the following values of the parameters: $b = 2$, $\sigma_Y = 1$, $\lambda = 3$, and $\beta = 2$. Notice that these values of b, λ and β satisfy the conditions in $L_{\beta,\lambda}$. We also selected $(0, 3)$ as the range of q values. It is slightly wider than the range $[1, 2]$ in the theorems and allows better visualisation of $T(q)$ and $f(\alpha)$, see Section 3.8.1 on the way to check its validity.

Figure 3.1 shows that the Rényi functions of the Models 1 and 2 have parabolic shapes while the Rényi function of the Model 3 is closer to a linear shape on the interval $(0, 3)$. Also, comparing the plots for Models 1, 2 and 3, we can see that the Rényi functions of Model 1 and 2 exhibit a concave down increasing and decreasing behaviour within $q \in (0, 3)$, whereas for Model 3 it increases. The multifractal spectra of Models 1, 2 and 3 show a concave down increasing behaviour within $q \in (0, 3)$.

3.7 Models based on power transformations of Gaussian fields

In the previous section, we considered three models based on an exponential transformation of Gaussian or gamma-correlated HIRF. This section proposes few much simpler scenarios where conditions of the theorems from Section 3.4 are satisfied.

First, note that the condition $\Lambda(x) > 0$ in Leonenko and Shieh (2013) can be relaxed to $\Lambda(x) > 0$ almost sure.

Model 4 Let $\Lambda(x) = Y^2(x)$, where $Y(x)$, $x \in \mathbb{R}^3$, is a zero-mean unit variance Gaussian HIRF with a covariance function $\rho_Y(\tau)$, $\tau \geq 0$.

For this model we obtain the following result, see the proof in Section 3.10.

Theorem 3.8. *Suppose that for Model 4, the correlation function of $Y(x)$ satisfies $|\rho_Y(r)| \leq Ce^{-\gamma r}, r > 0, \gamma > 0,$ and $b > \max(\sqrt[3]{1 + \sigma_\lambda^2}, e^{\sigma_\Lambda^2 C/3})$.*

Then the measures $\mu_k \xrightarrow{d} \mu$, $k \to \infty$, and the corresponding Rényi function is equal to

$$T(q) = q - 1 - \frac{1}{2} \log_b \left(\frac{2^q \Gamma(q + \frac{1}{2})}{\sqrt{\pi}} \right), \; q \in [1, 2]. \tag{3.18}$$

Example 3.1. *The approach developed in Section 3.5 can be used to obtain the moment conditions for $q \in [1, 4]$ in the case of Model 4. As it is shown in Section 3.10 it is enough to require that $b > e^{\frac{6(\max(\sigma_\Lambda C, 1))^4}{3}}$.*

Now we show that the assumption $\Lambda(x) > 0$ almost surely is indeed not restrictive and it is easy to construct a modification of Model 4 with $\Lambda(x) > 0$.

Model 4' Let

$$\Lambda(x) = Y^2(x) \cdot (1-\varepsilon) + \varepsilon,\ \varepsilon \in (0,1),$$

where $Y(x)$, $x \in \mathbb{R}^3$, is a zero-mean unit variance Gaussian HIRF with a covariance function $\rho_Y(\tau)$, $\tau \geq 0$.

It is easy to see that

$$E\Lambda(x) = (1-\varepsilon)EY^2(x) + \varepsilon = 1, \quad \sigma_\Lambda^2 = Var\Lambda(x) = 2(1-\varepsilon)^2 < +\infty,$$

$$Cov(\Lambda(x), \Lambda(y)) = 2(1-\varepsilon)^2 \rho_Y^2(\|x-y\|).$$

Hence, Model 4' satisfies Conditions 3.1 and 3.2 and $|\rho(r)| \leq Ce^{-\gamma r}$, $r > 0$, $\gamma > 0$, if $|\rho_Y(r)| \leq C' e^{-\gamma' r}$, $r > 0, \gamma' > 0$.

Therefore, we obtain

$$\begin{aligned} T_\varepsilon(q) &= q - 1 - \frac{1}{2}\log_b E((1-\varepsilon)Y^2(x) + \varepsilon)^q \\ &= q - 1 - \frac{q}{2}\log_b(1-\varepsilon) - \frac{1}{2}\log_b E\left(Y^2(x) + \frac{\varepsilon}{1-\varepsilon}\right)^q \end{aligned}$$

and

$$\lim_{\varepsilon \to 0} T_\varepsilon(q) = q - 1 - \frac{1}{2}\log_b E(Y^{2q}(x)),$$

which coincides with (3.18).

The next model generalizes Model 4 to an arbitrary even power of a Gaussian random field.

Model 5 Let $\Lambda(x) = Y^{2k}(x)$, $k \in \mathbb{N}$, where $Y(x)$, $x \in \mathbb{R}^3$, is a zero-mean Gaussian HIRF with the variance $\sigma^2 = \left(\frac{\sqrt{\pi}}{2^k \Gamma(k+\frac{1}{2})}\right)^{-\frac{1}{k}}$ and a covariance function $\rho_Y(r), r \geq 0$.

Theorem 3.9. *Suppose that for Model 5 the correlation function of $Y(x)$ satisfies $|\rho_Y(r)| \leq Ce^{-\gamma r}$, $r > 0$, $\gamma > 0$, and $b > \max\left(\sqrt[3]{1+\sigma_\Lambda^2},\ e^{\frac{\sigma_\Lambda^2 C}{3}}\right)$.*

Then the measures $\mu_k \xrightarrow{d} \mu$, $k \to \infty$, and the Rényi function is given by

$$T(q) = q - 1 - \frac{1}{2}\log_b EY^{2kq}(x) = q - 1 - \frac{1}{2}\log_b\left(\frac{2^{kq}\Gamma(kq+\frac{1}{2})}{\sqrt{\pi}}\right). \tag{3.19}$$

for $q \in [1,2]$.

The following model shows how vector-valued random fields can be used to construct mother fields. The chi-square random field $Y(\cdot)$, which is constructed from a vector Gaussian random field will be used.

Model 6 Let $\Lambda(x) = \frac{2}{k}Y(x)$, $k \in \mathbb{N}$, where $Y(x) \sim \chi^2(k)$, and the HIRF field $Y(x)$, $x \in \mathbb{R}^3$, has a covariance function $\rho_Y(r)$, $r \geq 0$.

Theorem 3.10. *Suppose that the correlation function in Model 6 satisfies the inequality* $|\rho_Y(r)| \leq Ce^{-\gamma r}$, $r > 0$, $\gamma > 0$, *and* $b > \max\left(\sqrt[3]{1+\sigma_\Lambda^2},\ e^{\frac{\sigma_\Lambda^2 C}{3}}\right)$.

Then the measures $\mu_k \xrightarrow{d} \mu$, $k \to \infty$, *and for* $q \in [1,2]$ *the Rényi function is equal to*

$$T(q) = q\left(1 - \frac{1}{2}\log_b\left(\frac{2}{k}\right)\right) - 1 - \frac{1}{2}\log_b\left(2^q\frac{\Gamma(q+\frac{k}{2})}{\Gamma(\frac{k}{2})}\right). \tag{3.20}$$

Note that $\Gamma^{'}(x) = \psi(x)\Gamma(x)$, where $\psi(x)$ is the digamma function defined by

$$\psi(x) = \int_0^\infty \left(\frac{e^{-t}}{t} - \frac{e^{-xt}}{1-e^{-t}}\right) dt.$$

Then, it follows from (3.13), (3.14) and (3.18) that for Model 4

$$\alpha(q) = 1 - \frac{1}{2}\log_b 2 - \frac{\psi(q+\frac{1}{2})}{2\ln 2},$$

$$f(\alpha(q)) = 1 + \frac{1}{2}\log_b\left(\frac{\Gamma(q+\frac{1}{2})}{\sqrt{\pi}}\right) - \frac{q\psi(q+\frac{1}{2})}{2\ln 2}. \tag{3.21}$$

Analogously, for Model 5 one gets from (3.19)

$$\alpha(q) = 1 - \frac{k}{2}\log_b 2 - \frac{k\psi(kq+\frac{1}{2})}{2\ln 2},$$

$$f(\alpha(q)) = 1 + \frac{1}{2}\log_b\left(\frac{\Gamma(kq+\frac{1}{2})}{\sqrt{\pi}}\right) - \frac{kq\psi(kq+\frac{1}{2})}{2\ln 2}. \tag{3.22}$$

Finally, it follows from (3.20) that for Model 6

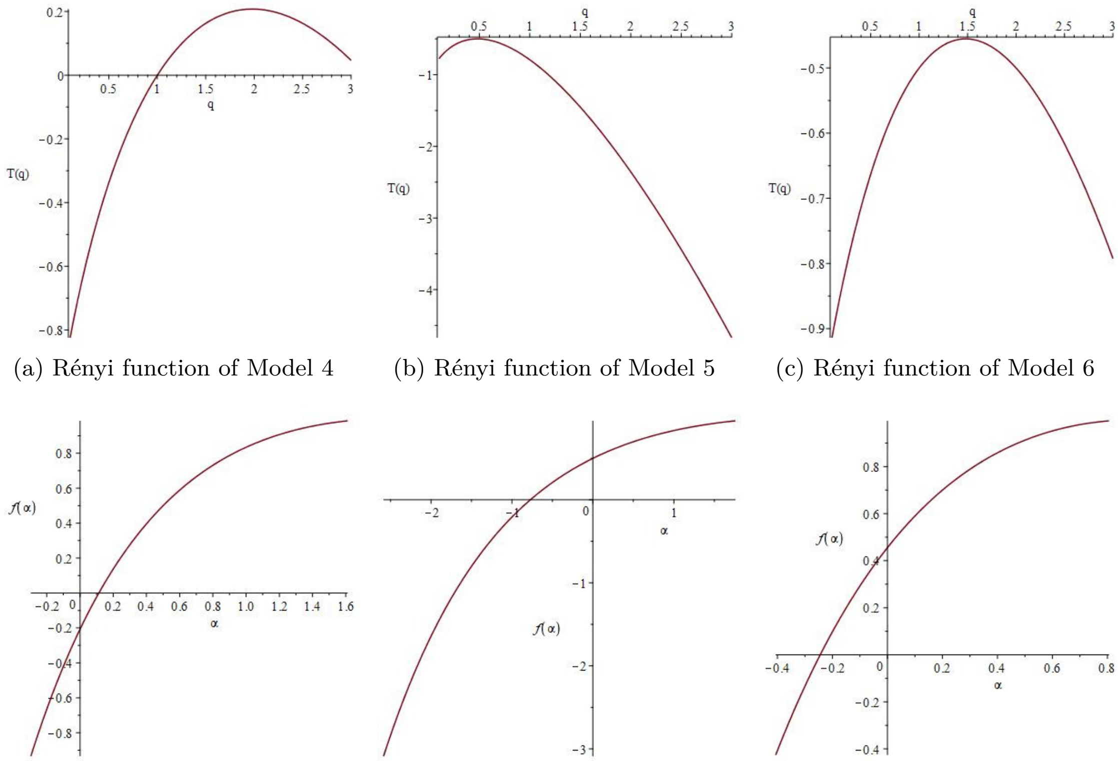
$$\alpha(q) = 1 - \frac{1}{2}\log_b\left(\frac{2}{k}\right) - \frac{1}{2}\log_b 2 - \frac{\psi(q+\frac{k}{2})}{2\ln 2},$$

$$f(\alpha(q)) = 1 + \frac{1}{2}\log_b\left(\frac{\Gamma(q+\frac{k}{2})}{\Gamma(\frac{k}{2})}\right) - \frac{q\psi(q+\frac{k}{2})}{2\ln 2}. \tag{3.23}$$

Summarising, we obtain

Theorem 3.11. *Let the corresponding conditions of Theorems 3.8, 3.9 and 3.10 are satisfied for Models 4, 5 and 6. Then the multifractal spectra of these models are given by (3.21), (3.22) and (3.23) respectively.*

The plots in Figure 3.2 demonstrate the behaviours of the Rényi functions and multifractal spectra of Models 4, 5 and 6. To plot Figure 3.2 we used similar settings and coordinate ranges as in Figure 3.1. The following values of parameters were chosen to produce the plots: $b = 2$, $k = 2$ and $\sigma_Y = 1$.



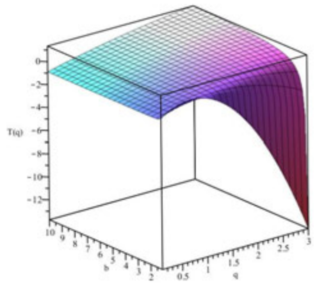
(a) Rényi function of Model 4 (b) Rényi function of Model 5 (c) Rényi function of Model 6

(d) Spectrum of Model 4 (e) Spectrum of Model 5 (f) Spectrum of Model 6

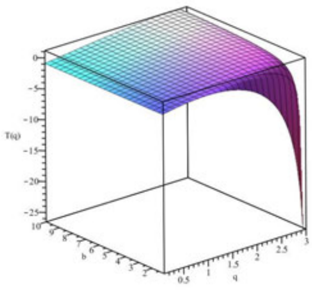
Figure 3.2: Examples of Rényi functions and multifractal spectra for Models 4, 5 and 6

Figure 3.2 shows that similar to Models 1 and 2, Models 4, 5 and 6 exhibit a parabolic-type behaviour. The spread of $T(q)$ values is wider for Model 5 when compared to Models 4 and 6 within $q \in (0, 3)$. The Rényi functions of Models 4, 5 and 6 manifest a concave down increasing and decreasing behaviour within $q \in (0, 3)$. Similar to Models 1, 2 and 3, the multifractal spectra of Models 4, 5 and 6 show a concave down increasing behaviour within $q \in (0, 3)$.

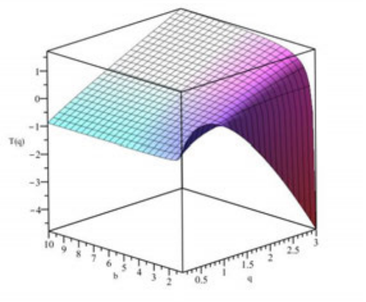
Finally, we illustrate the impact of parameter b on the Rényi function using Models 1, 2, 3, 4, 5 and 6. For Model 1, $\sigma_Y = 1$ was selected. Then, for Models 2 and 3, the parameters $\lambda = 3$ and $\beta = 2$ were used. The parameter $k = 2$ was chosen for Models 5 and 6. Figure 3.3 suggests that the Rényi functions for all the models exhibit a similar pattern. It suggests that the Rényi functions for Models 1, 3 and 4 are more concave than for other models for small values of b. The dispersion of $T(q)$ values on the interval $(0, 3)$ is smallest for Model 3 and largest for Model 5. Concavities of the Rényi functions become smaller when b increases and the other functions have almost linear behaviours for $q \in (0, 3)$ for large values of b.



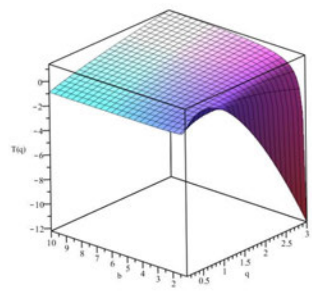
(a) Rényi functions of Model 1



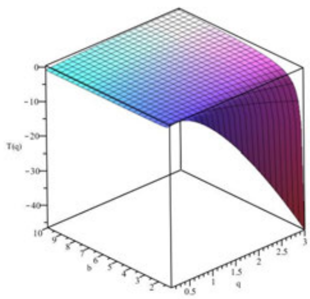
(b) Rényi functions of Model 2



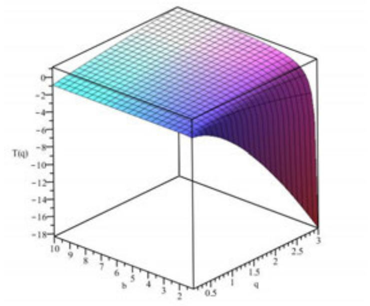
(c) Rényi functions of Model 3



(d) Rényi functions of Model 4



(e) Rényi functions of Model 5



(f) Rényi functions of Model 6

Figure 3.3: Dependence of the Rényi function on the parameter b

3.8 Numerical studies

3.8.1 Simulation methodology

There are numerous models for which explicit expressions for the Rényi function in terms of elementary functions or even series are not available. Also, in the majority of cases obtaining explicit mathematical formulae for Rényi functions is a difficult problem and rigorous results were derived only for some ranges of the parameter q. For example, for all models in Leonenko and Shieh (2013) and this chapter, $T(q)$ was derived for $q \in [1, 2]$ only. These results are also likely to be true for wider ranges of q, but showing it requires new proof strategies, see Section 3.5 and the example in Section 3.7. For such difficult cases, random field simulations can be used to obtain realizations of random fields from theoretical models and compute empirical Rényi functions. These empirical Rényi functions can be used as a substitute for exact mathematical functions when verifying whether or not real data are consistent with considered theoretical models.

The R package RandomFields provides a wide range of simulation techniques and algorithms for random fields, see Schlather et al. (2019) for more details. The function

RFSimulate simulates Gaussian random fields for a given covariance or variogram model and parameters defined by the arguments RMnugget and MTrend.

To compute a sample Rényi function the ratio $\frac{\log_2 E \sum_l \mu(S_l^{(m)})^q}{\log_2 |S_l^{(m)}|}$ from (3.6) for large values of m was used. This ratio was replaced by an empirical estimator that employs the HEALPix structure. For the HEALPix resolution $n = 1024$ with 12582912 pixels (in the following denoted by i), 196608 sets $S_l^{(m)}$ with 64 pixels per set were used to estimate the ratio. As CMB measurements $M(i)$ can take negative values, they were transformed to non-negative ones by subtracting their minimum value: $\tilde{M}(i) = M(i) - \min(M(i))$. Then, the terms $E(\mu(S_l^{(m)}))^q$ were estimated by $\hat{\mu}_l{}^q = \left(\sum_{i \in S_l^{(m)}} \tilde{M}(i) \,/ \sum \tilde{M}(i)\right)^q$. Finally, the empirical Rényi function was computed as

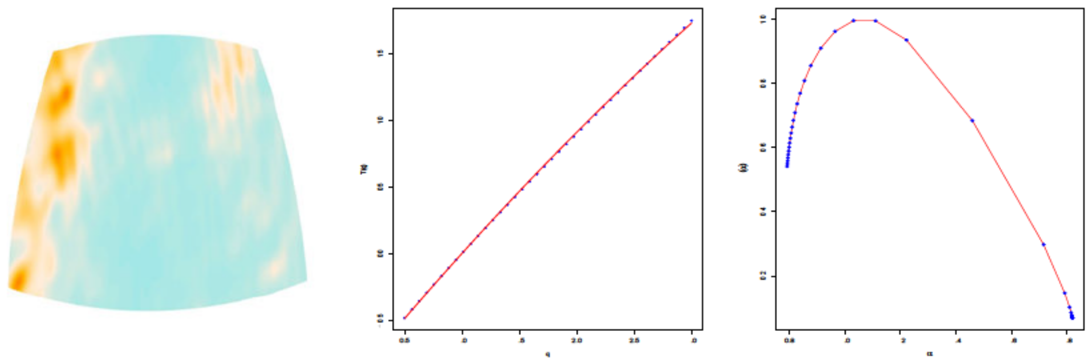
$$\hat{T}(q) = \frac{\log_2(\sum_l \hat{\mu}_l{}^q)}{\log_2 |S_l^{(m)}|}.$$

The empirical multifractal spectrum was estimated by

$$\hat{f}(\alpha) = q \cdot \hat{\alpha}(q) - \hat{T}(q), \qquad \hat{\alpha}(q) = \frac{\sum_l ((\hat{\mu}_l{}^q \,/ \sum \hat{\mu}_l{}^q) \cdot \ln(\hat{\mu}_l))}{\log_2 |S_l^{(m)}|}.$$

For the selected large number of pixels, $\hat{T}(q)$ and $\hat{f}(q)$ provide reliable estimations and can be used for wider intervals of q values than $[1, 2]$. In this chapter we considered intervals $(0.5, 3)$ and $(-10, 10)$ when it was required.

Figure 3.4a shows a realization of a multifractal random field in a large spherical window. The field was obtained from a Gaussian mother random field $Y(x)$ with the exponential covariance model and its variance equals 2. This covariance function has an exponential form and obviously satisfies the inequality in (3.4). As an approximation of the limit field, a finite product field $\Lambda_{40}(x)$ with $b = 3$ was used.



(a) Realization of a multifractal random field

(b) Sample Rényi function with the fitted Log-Normal model

(c) $f(\alpha)$ versus α

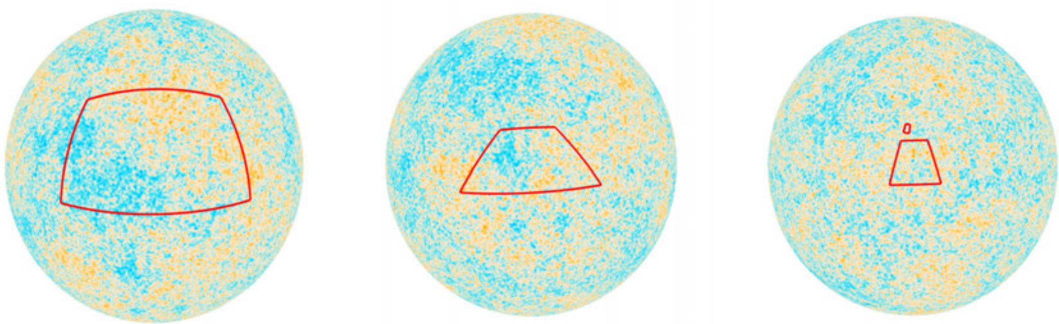
Figure 3.4: Analysis of a simulated multifractal random field

First 40 realizations $Y_i(b^i x),\ i = 1, \ldots, 40$, were simulated on a sphere by the package RandomFields. Then the finite-product field $\Lambda_{40}(x)$ was computed by transforming the simulated values according to the formula $\exp(\sum_{i=1}^{40} Y_i(b^i x) - 40)$. The dot plot of the empirical Rényi function is shown in Figure 3.4b. The solid straight line is used as a reference to see departures from the fitted Model 1. It is clear that the empirical Rényi function and the theoretical one from (3.8) are very close on an interval that is wider than $[1, 2]$. Figure 3.4c shows the spread of the multifractal spectrum.

The simulation studies suggest that the theoretical results from previous sections also hold for intervals wider than in the theorems.

3.8.2 Computing the Rényi function for CMB data

In this section, empirical Rényi functions were calculated for real cosmological data obtained from the NASA/IPAC Infrared Science Archive (IRSA (2021)). Figure 3.5 gives examples of sky windows CMB data from which were used to get empirical Rényi functions in the following examples.

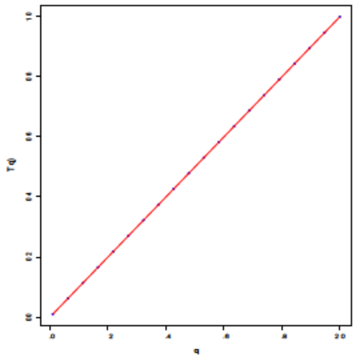


(a) Large window (b) Medium window (c) Small and very small windows

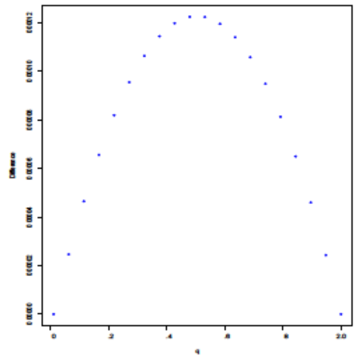
Figure 3.5: Different sky windows of CMB data

Extensive numerical studies were conducted for different windows in various sky locations. As in all cases we obtained rather similar results, we restrict our presentation only to few typical examples. The R package rcosmo was used for computations and visualizations, see Fryer et al. (2020) and Fryer et al. (2019) for more details. For small windows, the function fRen was slightly modified to change the support of the measure μ from the whole sky to the selected window.

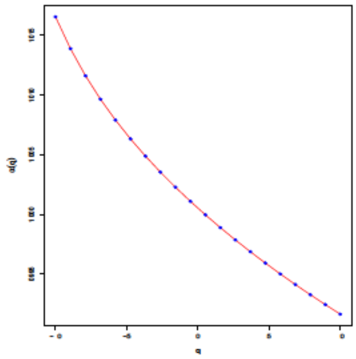
First the Rényi function was computed for the whole sky. The obtained sample Rényi function is shown in Figure 3.6a by dots. The straight line in the Figure 3.6a was drawn to assess departures of the sample Rényi function from a linear behaviour. The difference of the sample Rényi function and the linear function that connects the points (1,0) and



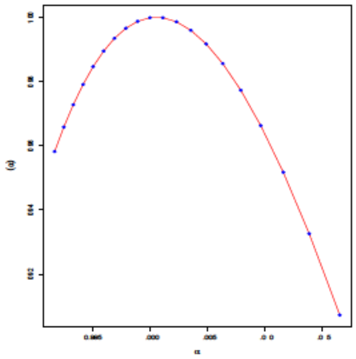
(a) Sample Rényi function versus linear function



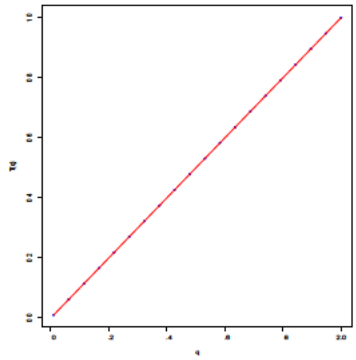
(b) Difference of sample Rényi function and linear function



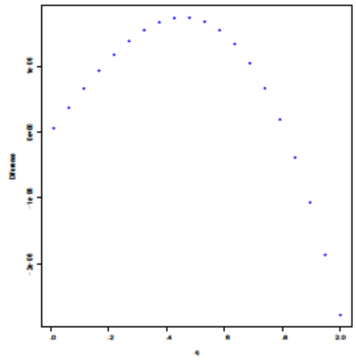
(c) $\alpha(q)$ versus q



(d) $f(\alpha)$ versus α



(e) Sample Rényi function with the fitted Log-Normal model



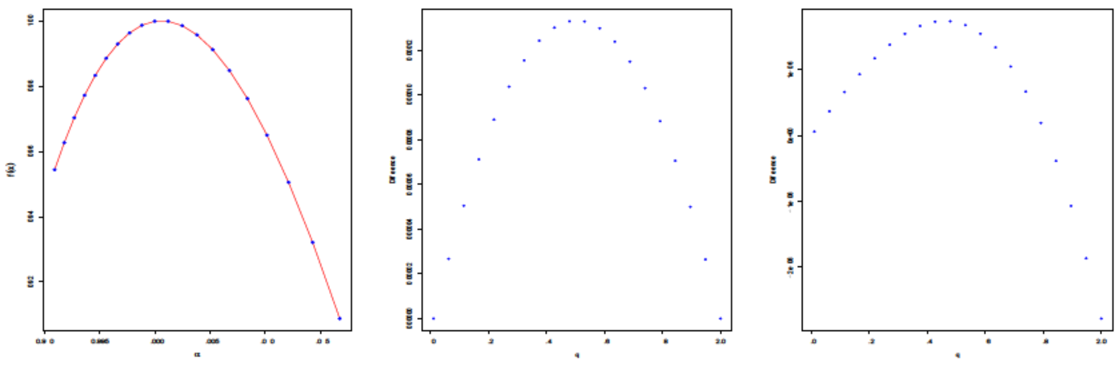
(f) Difference between sample Rényi function and the fitted Log-Normal model

Figure 3.6: Whole sky data analysis

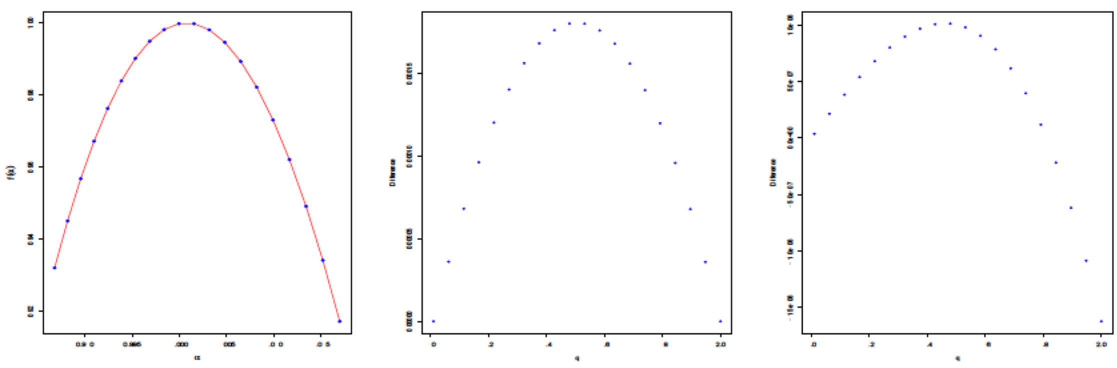
(2,1) is shown in Figure 3.6b. It is clear that the departure from a linear behaviour is not substantial. Figure 3.6c shows the function $\alpha(q)$ and Figure 3.6d plots the function $f(\alpha(q))$ versus $\alpha(q)$. As it was discussed in Section 3.8.1 to compute $\alpha(q)$ and $f(\alpha(q))$, we used the formula for Rényi functions for the range $(-10, 10)$ as simulation studies and analysis in Grahovac and Leonenko (2014) suggest the same analytical form of the Rényi function as for the range $[1, 2]$. All these plots confirm only very small multifractality of the CMB data. Similar results were also obtained for different sky windows, see, for example, Figure 3.7a, Figure 3.7b, Figure 3.7d and Figure 3.7e.

The Rényi functions, multifractal spectra, similar analysis and plots were produced for different window sizes of the CMB unit sphere. Large, medium, small and very small window sizes with areas 1.231, 0.4056, 0.0596 and 0.0017 were selected, see Figure 3.5. The Rényi function was computed for small windows located at different places of the sky sphere such as near the pole, near the equator and other places of the sphere. Although different window sizes of the sphere were investigated, there's not that much of evidence to suggest that we have substantial multifractality. The ranges of y scale in Figure 3.6b, Figure 3.7b and Figure 3.7e suggest that this multifractality is very small. These results and the variations of the values of $\hat{T}(q)$ between windows suggest that collecting data at very fine scales and further tests of hypothesis are required. We plan to develop tests of hypothesis about Rényi functions in future publications and use them for new high resolution data that will be available from the next generation CMB experiments CMB-S4, see Abazajian et al. (2019) for more details. As the obtained plots are rather similar, we present only two of them for large and small windows in Figure 3.7.

Then, all models from Sections 3.6 and 3.7 were used to fit the empirical Rényi function. For the log-normal model, we present the results for all windows. For other models, only results for CMB data in a large window are given. Similar results were also obtained for other windows. To fit models to empirical Rényi functions, several methods were employed. For the log-normal model, the simple linear regression approach was used whereas for the other models, the non-linear regression approach was applied.



(a) $f(\alpha)$ versus α for large window (b) Difference with linear function for large window (c) Difference with Model 1 for large window



(d) $f(\alpha)$ versus α for small window (e) Difference with linear function for small window (f) Difference with Model 1 for small window

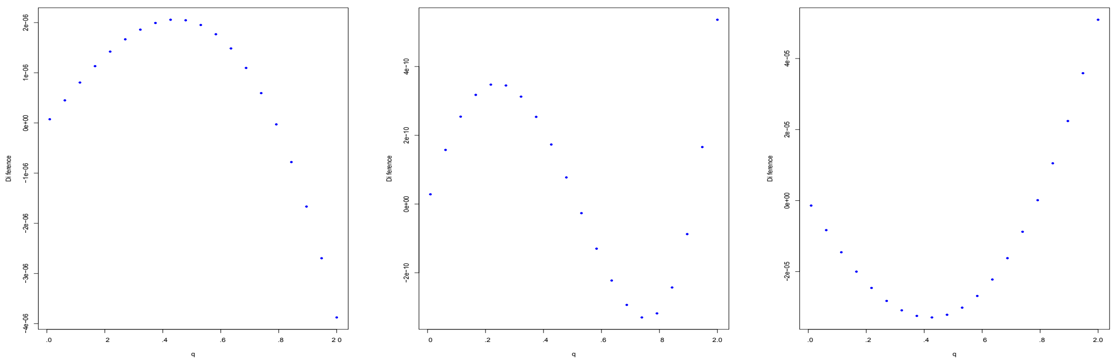
Figure 3.7: Analysis of large and small sky windows data

As the Rényi function of the log-normal model is specified by (3.8), substituting $a = \frac{\sigma_Y^2}{4 \ln b}$, results in the form $T(q) = a(-q^2 + q) + q - 1$. Then the R function "lm" was used for a simple linear regression fit with the intercept 0 to $T(q) - q + 1$. The values of the parameter a and the root mean square error for deviations of Model 1 from the empirical Rényi function are given in Table 3.1.

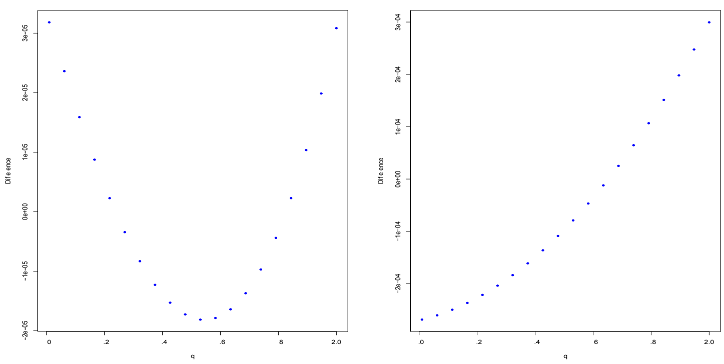
Observation window	$[\alpha_{\min}, \alpha_{\max}]$	$\alpha_{\max}$ - $\alpha_{\min}$	**a**	**RMSE**
Whole Sky	[0.9916, 1.0165]	0.024917	0.000513	$1.3602 \cdot 10^{-6}$
Large	[0.9908, 1.0167]	0.025846	0.000555	$1.3590 \cdot 10^{-6}$
Medium	[0.9893, 1.0159]	0.026620	0.000629	$1.1033 \cdot 10^{-6}$
Small	[0.9867, 1.0170]	0.030219	0.000745	$7.9095 \cdot 10^{-7}$
Very Small	[0.9842, 1.0543]	0.070150	0.001500	$1.3949 \cdot 10^{-5}$

Table 3.1: Analysis of different sky windows data with Model 1

Figure 3.6e demonstrates the fit of the log-normal model (shown in the red colour) to the empirical Rényi function. As this plot is rather similar for all other models and windows, we present only the plots of residuals in Figure 3.6f, Figure 3.7c, Figure 3.7f and Figure 3.8.



(a) Difference with Model 2 (b) Difference with Model 3 (c) Difference with Model 4



(d) Difference with Model 5 (e) Difference with Model 6

Figure 3.8: Differences between the sample Rényi function and the fitted model

As the estimated value of a is close to zero, the fit of Model 1 gives an almost degenerated case, when either σ_Y^2 is very small or b is very large, which is consistent with the plot in Figure 3.3. The results in Table 3.1 also confirm that multifractality is very small as for all observation windows a is almost zero and $\alpha_{\max} - \alpha_{\min}$ is very small.

Next, for the log-gamma model specified by (3.11), we used the reparameterisation $A = \frac{2}{\beta}\ln(b)$, $B = \lambda^{-1}$ and considered the non-linear model $T(q) - q + 1 = A^{-1}(\ln(1 - Bx) - x\ln(1 - B))$. The command "nlsLM" from the R package minpack.lm with appropriate initial values was used to fit the model to the sample values of $\hat{T}(q) - q + 1$. The values of estimated parameters were $\hat{A} = 0.029407$ and $\hat{B} = 0.005469$ with $RMSE = 1.7198 \times 10^{-6}$. The corresponding values of b, λ and β satisfy the assumptions of Theorem 3.5.

For the log-negative-inverse-gamma model given by (3.12), the reparameterisation $A = 2\ln(b)$, $B = \beta$, $C = \sqrt{\lambda}$ and application of "nlsLM" resulted in $\hat{A} = 0.254719$, $\hat{B} = 5.695755$, $\hat{C} = 0.386207$ and $RMSE = 2.6201 \times 10^{-10}$. Note that the obtained $\hat{C}$ corresponds to λ that is outside of $L_{\beta,\lambda}$ in Theorem 3.6. For parameters satisfying the conditions of Theorem 3.6, RMSE is substantially larger. As the results in Theorem 3.6 might be also true for other parameters (see the discussion in Section 3.8.1), we used the obtained value of $\hat{C}$.

Model 4 was fitted by using a linear regression model with the parameter $A = -\frac{1}{2\log_2 b}$. The estimated parameter $\hat{A}$ was -0.000667 and $RMSE = 2.56533 \times 10^{-5}$. For Models 5 and 6 the non-linear regression approach and the R function "nls" were used. For Model 5 the estimates were found as $\hat{A} = -0.000762$, $\hat{k} \approx 1$ and $RMSE = 1.6347 \times 10^{-5}$. Finally, for Model 6 the estimated parameters were $\hat{A} = 0.000269$, $\hat{k} = 1$ and $RMSE = 1.8393 \times 10^{-4}$.

Figure 3.7c and Figure 3.8 demonstrate that departures of the fitted models from the empirical Rényi function are very small, but have different patterns. The numerical studies suggested that Models 2 and 3 are more flexible than the other models. However, to fit these models one has to very carefully choose initial values of the parameters for the nls estimation. Different initial values can lead to different results which can be a potential issue for data which, similarly to CMB, show minor multifractality. Also, nls method's rates of convergence for Models 2 and 3 are very slow. Models 1, 4, 5 and 6 have less parameters and are less flexible than Models 2 and 3. However, in many cases they give a reasonable fit very quickly, are robust to the choice of initial values and are more computationally efficient.

All models gave a good fit to the empirical Rényi functions. The analysis in this section suggests no significant or very small multifractality for the currently available resolution of CMB measurements.

3.9 Conclusion

This chapter investigates the multifractal behaviour of spherical random fields and some applications to cosmological data from the mission Planck. The aim of this chapter is to introduce several multifractal models for random fields on a sphere and to propose simpler models where the Rényi function can be computed explicitly. All Rényi functions for the specified models exhibit either parabolic or approximately linear behaviours. We present the Rényi function computations for different CMB sky windows located at different places of the sphere. Finally, we fit the specified models to actual CMB data. All models fit to the data. The analysis suggests that there may exist a very minor multifractality of the CMB data for the currently available resolution.

Some related problems and extensions of the current research that would be interesting for future studies:

- Develop statistical tests for different types of Rényi functions;
- Prove that the theoretical results and the formulae for the Rényi functions are also valid for the values of q outside the interval $[1, 2]$, see Denisov and Leonenko (2016);
- Study other models based on vector random fields (similar to Model 6), where the Rényi functions can be computed explicitly;
- Develop some approaches to study rates of convergence for the obtained asymptotics, that would serve as analogous of classical convergence rates in central and non-central limit theorems, see Anh et al. (2015) and Anh et al. (2019a);
- Investigate changes of the Rényi functions depending on evolutions of random fields driven by SPDEs on the sphere, see Anh et al. (2018), Broadbridge et al. (2019) and Broadbridge et al. (2020);
- Apply the developed models and methodology to other spherical data, in particular, to new high-resolution CMB data from future CMB-S4 surveys (Abazajian et al. (2019)) which will be collecting $3D$ observations.

This chapter studies data that are modelled as restrictions of $3D$ random fields to the unit sphere. Compared to the available literature, this approach is more consistent with real CMB observations that exist in $\mathbb{R}^3$, but are measured only on $s_2(1)$, see the discussions in Anh et al. (2018), Broadbridge et al. (2019) and Broadbridge et al. (2020). For other applications it would be important to develop similar results for the case of intrinsic spherical random fields, i.e., random fields directly defined on $s_2(1)$, see the discussion about differences of covariance models of random fields with supports in $\mathbb{R}^3$ and $s_2(1)$ in Gneiting (2013).

3.10 Proofs

This section gives the proofs in this chapter. It specifically gives the proofs of Theorem 3.3, Theorem 3.8, Example 3.1, Theorem 3.9 and Theorem 3.10.

Proof of Theorem 3.3. By Remark 3.1, from the weak convergence of the measures μ_k to μ and the assumption of exponential boundedness of the covariance function of the mother field, it follows that

$$E\mu_k^q(B^3) \to E\mu^q(B^3), \quad k \to \infty.$$

By the Lyapunov's inequality, see (Loéve, 1977, p.162),

$$E\mu_k^q(B^3) \leq (E\mu_k^2(B^3))^{q/2}, \text{ for } q \in [1, 2].$$

Therefore, to guarantee $E\mu^q(B^3) < +\infty$, $q \in [1, 2]$, it is sufficient to provide such b and σ_Λ^2 that

$$\sup_{k\in\mathbb{N}} E\mu_k^2(B^3) < +\infty.$$

By (3.4), the non-negativity of $\Lambda_k(y)$ and independence of $\Lambda^{(i)}$ it holds

$$E\mu_k^2(B^3) = E\int_{B^3}\int_{B^3} \Lambda_k(y)\Lambda_k(\tilde{y})d\tilde{y}dy = \int_{B^3}\int_{B^3} E[\Lambda_k(y)\Lambda_k(\tilde{y})]d\tilde{y}dy$$

$$= \int_{B^3}\int_{B^3} E\prod_{i=0}^{k} \Lambda^{(i)}(yb^i)\Lambda^{(i)}(\tilde{y}b^i)d\tilde{y}dy = \int_{B^3}\int_{B^3} \prod_{i=0}^{k} E\Lambda^{(i)}(yb^i)\Lambda^{(i)}(\tilde{y}b^i)d\tilde{y}dy$$

$$= \int_{B^3}\int_{B^3} \prod_{i=0}^{k}\Big(E(\Lambda^{(i)}(yb^i) - 1)(\Lambda^{(i)}(\tilde{y}b^i) - 1) + E\Lambda^{(i)}(yb^i) + E\Lambda^{(i)}(\tilde{y}b^i) - 1\Big)d\tilde{y}dy$$

$$= \int_{B^3} \int_{B^3} \prod_{i=0}^{k} (Cov(\Lambda(yb^i), \Lambda(\tilde{y}b^i)) + 1) d\tilde{y}dy = \int_{B^3} \int_{B^3} \prod_{i=0}^{k} \left(1 + \sigma_\Lambda^2 \rho_\Lambda(\|y - \tilde{y}\|b^i)\right) d\tilde{y}dy$$

$$\leq \int_{B^3} \int_{B^3} \prod_{i=0}^{k} (1 + \sigma_\Lambda^2 C e^{-\gamma\|y-\tilde{y}\|b^i}) d\tilde{y}dy \leq \int_{B^3} \int_{B^3} \prod_{i=0}^{\infty} (1 + \sigma_\Lambda^2 C e^{-\gamma\|y-\tilde{y}\|b^i}) d\tilde{y}dy.$$

From the inequality $1 + a \leq e^a$, it follows that

$$E\mu_k^2(B^3) \leq \int_{B^3} \int_{B^3} \prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-\gamma\|y-\tilde{y}\|b^i}} d\tilde{y}dy.$$

Introducing the new variables $z = y$, $\tilde{z} = y - \tilde{y}$, one obtains

$$E\mu_k^2(B^3) \leq \int_{B^3} dz \int_{B^3 - B^3} \prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-\gamma\|\tilde{z}\|b^i}} d\tilde{z},$$

where $B^3 - B^3 = \{\tilde{z} : \tilde{z} = y - \tilde{y},\ y,\ \tilde{y} \in B^3\}$.

Hence, by using the spherical change of variables,

$$E\mu_k^2(B^3) \leq |B^3| \int_0^{diam(B^3)} r^2 \prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-\gamma r b^i}} dr$$

$$= \frac{|B^3|}{\gamma^3} \int_0^{\gamma diam(B^3)} r^2 \prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-r b^i}} d\tau.$$

As the exponent $e^{\sigma_\Lambda^2 C e^{-rb^i}}$ is a decreasing function of r, selecting $n(r) = \max(0, -[\log_b(r)])$, $r > 0$, we obtain

$$\prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-rb^i}} \leq \prod_{i=0}^{n(r)-1} e^{\sigma_\Lambda^2 C e^{-rb^i}} \prod_{i=n(r)}^{\infty} e^{\sigma_\Lambda^2 C e^{-rb^i}}$$

$$\leq e^{\sigma_\Lambda^2 C n(r)} \prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-rb^{i+n(r)}}} \leq e^{\sigma_\Lambda^2 C n(r)} \prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-b^i}}.$$

Notice that

$$\prod_{i=0}^{\infty} e^{\sigma_\Lambda^2 C e^{-b^i}} = e^{\sigma_\Lambda^2 C \sum_{i=0}^{\infty} e^{-b^i}} \leq e^{\sigma_\Lambda^2 C \sum_{i=0}^{\infty} e^{-(1+(b-1)^i)}}$$

$$= e^{\frac{\sigma_\Lambda^2 C}{e} \sum_{i=0}^{\infty} e^{-(b-1)^i}} = e^{\frac{\sigma_\Lambda^2 C}{e} \frac{1}{1-e^{-(b-1)}}} < +\infty.$$

Therefore,

$$E\mu_k^2(B^3) \leq \frac{|B^3|}{\gamma^3} e^{\frac{\sigma_\Lambda^2 C}{e(1-e^{-(b-1)})}} \int_0^{\gamma diam(B^3)} z^2 e^{\sigma_\Lambda^2 C n(z)} dz$$

$$= \frac{|B^3|}{\gamma^3} e^{\frac{\sigma_\Lambda^2 C}{e(1-e^{-(b-1)})}} \int_0^{\gamma diam(B^3)} z^2 \max\left(1,\ z^{-\frac{\sigma_\Lambda^2}{\ln(b)}}\right) dz.$$

The integral is finite if $2 - \frac{\sigma_\Lambda^2 C}{\ln(b)} > -1$, i.e. $b > e^{\frac{\sigma_\Lambda^2 C}{3}}$. □

Proof of Theorem 3.8. By the definition of Model 4 it follows that

$$E\Lambda(x) = E(Y^2(x)) = \rho_Y(0) = 1, \quad \sigma_\Lambda^2 = Var\Lambda(x) = E(Y^4(x)) - 1 = 2,$$

$$Cov(\Lambda(x), \Lambda(y)) = E(Y^2(x) - 1)(Y^2(y) - 1) = 2\rho_Y^2(\|x - y\|).$$

To compute the covariance, we used the property

$$E(H_k(Y(x))H_l(Y(y))) = \delta_k^l k! \rho_Y^k(\|x - y\|), \quad x, y \in \mathbb{R}^3, \tag{3.24}$$

where $H_k(u)$, $k \geq 0$, $u \in \mathbb{R}$, are the Hermite polynomials, see Peccati and Taqqu (2011). For $k = 2$, the Hermite polynomial of order 2 is $H_2(u) = u^2 - 1$.

Thus, Model 4 satisfies Conditions 3.1 and 3.2.

Note that the condition $|\rho_\Lambda(r)| \leq Ce^{-\gamma r}$, $r > 0$, $\gamma > 0$, is equivalent to

$$|\rho_Y(r)| \leq C^{'} e^{-\gamma^{'} r}, \ r > 0, \ \gamma^{'} > 0. \tag{3.25}$$

So, if (3.25) is satisfied, then one can apply Theorems 3.1 and 3.2 and the Rényi function of the limit measure equals to

$$T(q) = q - 1 - \frac{1}{2}\log_b EY^{2q}(x).$$

Finally, noting that for $p > -1$ and $Z \sim N(\mu, \sigma^2)$

$$E|Z - \mu|^p = \sigma^p \frac{2^{p/2}\Gamma(\frac{p+1}{2})}{\sqrt{\pi}} \tag{3.26}$$

finalises the proof. □

Proof of Example 3.1. By Remark 3.3, it is enough to check that

$$\sup_{k \in N} E\mu_k^4(B^3) = \sup_{k \in N} \int_{B^3}\int_{B^3}\int_{B^3}\int_{B^3} \prod_{i=0}^{k} E\left(\prod_{j=1}^{4} Y^2(y_j b^i)\right) \prod_{j=1}^{4} dy_j < +\infty.$$

Notice, that by Wick's theorem

$$E\left(\prod_{j=1}^{4} Y^2(y_j b^i)\right) = \sum_{p\in P_4^2} \prod_{(j,\tilde{j})\in p} Cov(Y(y_j b^i), Y(y_{\tilde{j}} b^i)), \tag{3.27}$$

where the sum is over all parings p of $\{1,1,2,2,3,3,4,4\}$, which are distinct ways of partitioning $\{1,1,2,2,3,3,4,4\}$ into pairs (i,j). The product in (3.27) is over all pairs contained in p, see Janson (1997).

Notice that for the pairing $p^* = \{(1,1),(2,2),(3,3),(4,4)\}$.

$$\prod_{(j,\tilde{j})\in p^*} Cov(Y(y_j b^i), Y(y_{\tilde{j}} b^i)) = \prod_{j=1}^{4} EY^2(y_j b^i) = 1.$$

In all other cases of pairing, there is at least one pair $(j,\tilde{j})$ such that $j \neq \tilde{j}$. Therefore, the expectation $E(\prod_{j=1}^{4} Y^2(y_j b^i))$ equals

$$1 + \sum_{\substack{p\in P_4^2\\ p\neq p^*}} \prod_{(j,\tilde{j})\in p} Cov(Y(y_j b^i), Y(y_{\tilde{j}} b^i)).$$

As, $1 + a < e^a$, it can be estimated by

$$\exp\left(\sum_{\substack{p\in P_4^2\\ p\neq p^*}} \prod_{(j,\tilde{j})\in p} Cov(Y(y_j b^i), Y(y_{\tilde{j}} b^i))\right).$$

As at least for one pairing $(j,\tilde{j}) \in p \neq p^*$ it holds that $j \neq \tilde{j}$, then one can use the upper bound

$$|Cov(Y(y_j b^i), Y(y_{\tilde{j}} b^i))| \leq \sigma_Y^2 C e^{-\gamma \|y_j - y_{\tilde{j}}\| b^i},$$

and the approach from the proof of Theorem 3.3.

Namely,

$$E\left(\prod_{j=1}^{4} Y^2(y_j b^i)\right) \leq e^{\sum_{p\in p_4^2} \prod_{(j,\tilde{j})\in p} \sigma_Y^2 C e^{-\gamma \|y_j - y_{\tilde{j}}\| b^i}}$$

$$\leq e^{(\max(\sigma_\Lambda^2 C, 1))^4 \sum_{1\leq j\leq \tilde{j}\leq 4} e^{-\gamma \|y_j - y_{\tilde{j}}\| b^i}}.$$

Hence,

$$\sup_{k\in N} \int_{B^3}\int_{B^3}\int_{B^3}\int_{B^3} \prod_{i=0}^{k} E\left(\prod_{j=1}^{4} Y^2(y_j b^i)\right) \prod_{j=1}^{4} dy_j$$

$$\leq \int_{B^3}\int_{B^3}\int_{B^3}\int_{B^3}\prod_{i=0}^{\infty} e^{(\max(\sigma_\Lambda^2 C,1))^4 \sum_{1\leq j\leq \tilde{j}\leq 4} e^{-\gamma\|y_j - y_{\tilde{j}}\| b^i}}$$

$$\leq \left(\int_{B^3}\int_{B^3}\int_{B^3}\int_{B^3}\prod_{i=0}^{\infty} e^{6(\max(\sigma_\Lambda^2 C,1))^4 \sum_{1\leq j\leq \tilde{j}\leq 4} e^{-\gamma\|y_j - y_{\tilde{j}}\| b^i}} \prod_{j=1}^{4} dy_j\right),$$

where the last inequality follows from the generalized Hölder's inequality

$$\left\|\prod_{k=1}^{K} f_k\right\|_1 \leq \prod_{k=1}^{K} \|f_k\|_{p_k},$$

with $\sum_{k=1}^{K} {p_k}^{-1} = 1$. In our case $K = 6$ is the number of different j and $\tilde{j}$ satisfying $1 \leq j \leq \tilde{j} \leq 4$.

Finally, similar to the proof of Theorem 3.3, from equation (3.18) we obtain the condition $b > e^{\frac{6(\max(\sigma_\Lambda C,1))^4}{3}}$. □

Proof of Theorem 3.9. It follows from (3.26) that

$$E\Lambda(x) = EY^{2k}(x) = \sigma^{2k}\frac{2^k\Gamma(k+\frac{1}{2})}{\sqrt{\pi}} = 1,$$

$$\sigma_\Lambda^2 = Var\Lambda(x) = E(Y^{4k}) - 1 = \left(\frac{\sqrt{\pi}}{2^k\Gamma(k+\frac{1}{2})}\right)^2 \frac{2^{2k}\Gamma(2k+\frac{1}{2})}{\pi} - 1$$

$$= \frac{\sqrt{\pi}\Gamma(2k+\frac{1}{2})}{\Gamma^2(k+\frac{1}{2})} - 1 < +\infty.$$

To compute the covariance function we use (3.24) and the following Hermite expansion (Abramowitz and Stegun, 1964, page 775)

$$z^{2k} = (2k)!\sum_{i=0}^{k}\frac{H_{2k-2i}(z)}{2^i i!(2k-2i)!}.$$

Therefore,

$$Cov(\Lambda(x),\Lambda(y)) = E(Y^{2k}(x)-1)(Y^{2k}(y)-1) = \frac{\pi E(\tilde{Y}^{2k}(x)\tilde{Y}^{2k}(y))}{2^{2k}\Gamma^2(k+\frac{1}{2})} - 1$$

$$= ((2k)!)^2\frac{\pi}{2^{2k}\Gamma^2(k+\frac{1}{2})}\sum_{i=0}^{k}\frac{E[H_{2k-2i}(\tilde{Y}(x))H_{2k-2i}(\tilde{Y}(y))]}{2^{2i}(i!)^2((2k-2i)!)^2} - 1$$

$$= ((2k)!)^2\frac{\pi}{2^{2k}\Gamma^2(k+\frac{1}{2})}\sum_{i=0}^{k}\frac{\tilde{\rho}^{2k-2i}(\|x-y\|)}{2^{2i}(i!)^2(2k-2i)!} - 1, \qquad (3.28)$$

where $\tilde{Y}(x) = Y(x)/\left(\frac{\sqrt{\pi}}{2^k\Gamma(k+\frac{1}{2})}\right)^{1/2k}$ is a zero-mean unit variance Gaussian HIRF with the covariance function

$$\tilde{\rho}(\|x-y\|) = \left(\frac{2^k\Gamma(k+\frac{1}{2})}{\sqrt{\pi}}\right)^{1/k} \rho_Y(\|x-y\|).$$

Notice, that for $i=k$ in (3.28) by the Legendre duplication formula

$$\frac{((2k)!)^2\pi}{2^{2k}\Gamma^2(k+\frac{1}{2})2^{2k}(k!)^2} = \frac{\Gamma^2(2k+1)\pi}{2^{4k}\Gamma^2(k+\frac{1}{2})k^2\Gamma^2(k)} = \frac{(2k)^2\Gamma^2(2k)\pi}{2^{4k}k^2 2^{2-4k}\pi\Gamma^2(2k)} = 1.$$

Hence,

$$Cov(\Lambda(x),\Lambda(y)) = \frac{((2k)!)^2\pi}{2^{2k}\Gamma^2(k+\frac{1}{2})} \sum_{i=0}^{k-1} \frac{(2^k\Gamma(k+\frac{1}{2}))^{2+\frac{2i}{k}}}{2^{2i}(i!)^2(2k-2i)!\pi^{1/2k}} \tilde{\rho}^{2k-2i}(\|x-y\|).$$

Therefore, if $|\tilde{\rho}(r)| \le C^{'}e^{-\gamma^{'}r}$, $r>0$, $\gamma^{'}>0$, then the covariance function of Model 5 satisfies the condition $|\rho_\Lambda(r)| \le Ce^{-\gamma r}, r>0$, $\gamma>0$, and the Rényi function equals

$$T(q) = q-1-\frac{1}{2}\log_b EY^{2kq}(x) = q-1-\frac{1}{2}\log_b\left(\frac{2^{kq}\Gamma(kq+\frac{1}{2})}{\sqrt{\pi}}\right). \qquad \square$$

Proof of Theorem 3.10. By properties of the chi-square distribution, it follows that

$$E\Lambda(x) = \frac{2}{k}EY(x) = 1, \quad Var\Lambda(x) = \frac{4}{k^2}VarY(x) = \frac{2}{k} < +\infty,$$

$$Cov(\Lambda(x),\Lambda(y)) = \frac{4}{k^2}\rho_Y(\|x-y\|).$$

Notice that if $Y(x) = \frac{1}{2}(Z_1^2(x)+...+Z_k^2(x))$, $x\in\mathbb{R}^3$, where $Z_i(x)$, $i=1,...,k$, are independent zero-mean unit variance components of k-dimensional vector Gaussian HIRF with a covariance function $\rho_Z(r)$, $r\ge 0$ of each component, then

$$Cov(\Lambda(x),\Lambda(y)) = \frac{4}{k^2}\cdot\frac{k}{2}\rho_Z^2(\|x-y\|) = \frac{2}{k}\rho_Z^2(\|x-y\|).$$

Therefore, Model 6 satisfies Conditions 3.1 and 3.2 and $|\rho_\Lambda(r)| \le Ce^{-\gamma r}$, $r>0$, $\gamma>0$, if $|\rho_Y(r)| \le C'e^{-\gamma^{'}r}$ or $|\rho_Z(r)| \le C'e^{-\gamma^{'}r}$, $r\ge 0, \gamma^{'}>0$.

Then, the corresponding Rényi function is given by (3.20). $\square$

Chapter 4

On multifractionality of spherical random fields with cosmological applications

This chapter is based on the article, Broadbridge, P., Nanayakkara, R., and Olenko, A. On multifractionality of spherical random fields with cosmological applications, which has been submitted for publication.

Due to artistic reasons, the format of this paper was changed in accordance with the style of the thesis. This did not change the main contents of the paper, but gave rise to slight changes in the paper layout.

4.1 Introduction

The notion of fractional Brownian motion (FBM) was introduced by B. Mandelbrot and Van Ness in 1968. The FBM depends on the Hurst parameter H, where $H \in (0, 1)$. The Hurst parameter can be used to define the Hölder regularity of FBM. The multifractional Brownian motion (MBM) was first considered by Péltier and Lévy Véhel in 1995 extending the FBM, see Ayache and Véhel (2004). The concept of multifractionality induces from fractionality allowing local properties to depend on space-time locations. The Hurst parameter H of FBM is replaced by $H(t)$ in MBM. The MBM was proposed to model data that cannot be described by standard processes with stationary increments since their pointwise smoothness changes from point to point.

Multifractional processes were used to study complex stochastic systems which exhibit nonlinear behaviour in space and time. Multifractional behaviour of data has been found in many applications such as, image processing, stock price movements, signal processing, see Ayache and Véhel (2004), Bianchi and Pianese (2007) and Sheng et al. (2011). Multifractional processes are more flexible in comparison to FBM and can be non-stationary. Multifractional Gaussian processes were studied in Benassi et al. (1998a) where a method to evaluate the multifractionality using discrete observations of a process's single sample path was proposed.

The generalized multifractional Brownian motion (GMBM) is a continuous Gaussian process that was introduced by generalizing the traditional FBM and MBM, see Ayache and Véhel (2004). Comparing to MBM, the Hölder regularity of GMBM can substantially vary. For example, GMBM can allow discontinuous Hölder exponents. This has been an advantage to applications, specifically, medical image modelling, telecommunication, turbulence and finance where the pointwise Hölder exponent can change rapidly. A Fourier spectrum's low frequencies control the long-range dependence of a stochastic process while the higher frequencies control the Hölder regularity. Therefore, GMBM can be used to model processes that exhibit erratic behaviour of the local Hölder exponent and long-range dependence, see Ayache and Véhel (2004).

This chapter deals with cosmological applications. The universe originated about 14 billion years ago and had extremely high temperature. The atoms were broken down into electrons and protons. The universe started to cool down and hydrogen atoms were formed 380,000 years after the Big Bang. As a result, photons were emitted and started moving without any restriction. This utmost ancient glow of light which is the leftover radiation after the Big Bang is called the cosmic microwave background radiation (CMB), see Planck Satellite (2021). The CMB which dates back to nearly 400,000 years from the Big Bang was first discovered by Arno Pensiaz and Robert Wilson in 1964. The CMB is an electromagnetic radiation caused by the thermal movement of particles left in the universe. In the microwave region, the CMB spectrum closely follows that of a black body at equilibrium temperature 2.735K, tracing back to a plasma temperature of around 4000K at a time corresponding to redshift z=1500 at 50% atomic combination.

Although the equilibrium spectrum is important, there are important departures from equilibrium that give information on the state of the early universe, see for example Pietroni (2009). Relative anisotropic variations of spectral intensity from that of a black body are of the order of 10^{-4}. Calculations by Khatri and Sunyaev (Khatri and Sunyaev (2012)) showed that outside of a relatively small range of redshifts, external energy inputs from sources such as massive particle decay, would dissipate by Compton and double Compton scattering and other relaxation processes to affect the signal by several lower orders of magnitude. The primary sources of anisotropy were large-scale acoustic waves whose compressions in the plasma universe were associated with raised temperatures. Using the current angular widths of anisotropies in the CMB, the current standard model ΛCDM (cold dark matter plus dark energy) affords an estimate of the Hubble constant at $H_0 = 67.4 \pm 1.4$ km/s/MPc (Aghanim et al. (2020)). This agrees well with data from the POLARBEAR Antarctica telescope that give $H_0 = 67.2 \pm 0.57$ km/s/MPc (Adachi et al. (2020)). However estimates from more recent emissions from closer galaxies using both cepheid variables and type Ia supervovae as distance markers, give $H_0 = 74.03 \pm 1.42$ km/s/MPc (Riess et al. (2019)). This unexplained discrepancy will eventually be resolved by newly found errors in the methodology of one or both of the competing large-z and small-z measurements, or in new physical processes that are currently unidentified.

Within a turbulent plasma, there are electrodynamical processes that are far more complicated than the large-scale acoustic waves. When radiation by plasma waves is taken into account, useful kinetic equations and spectral functions can no longer be constructed by Bogoliubov's approach of closing the moment equations for electron distribution functions, see Chapter 5 in Klimontovich (1967). Even in controlled tokomak devices, the dynamical description of magnetic field lines has fractal attracting sets (Viana et al. (2011)) and charged particle trajectories may have fractal attractors under the influence of multiple magnetic drift waves (Mathias et al. (2017)). At CMB frequencies below 3 GHz (i.e. wavelengths larger than 10 cm), there have been indications of spectral intensities much higher than that of a 2.7 K black body (Baiesi et al. (2020)). Although there is a high level of confidence in measuring the universe's expansion factor from CMB since the decoupling of photons from charged particles, the level of complexity of magnetohydrodynamics in plasma suggests that this subject might not be a closed book. Multi-fractal analysis is a tool that might contribute to understanding the multi-scale data that are becoming successively more fine-grained with each generation of radio telescope.

The space missions that have studied the CMB so far are Cosmic Background Explorer, Wilkinson Microwave Anisotropy Probe and Planck. The Planck mission was launched in 2009 to measure the CMB with an extraordinary accuracy over a wide spectrum of infrared wavelengths. The Planck mission traced the CMB anisotropies at narrow angles with a high resolution and sensitivity. The measured temperature intensities from the Planck mission together with the polarisation data can be used to check for the existence of anomalies within the CMB data. The CMB data can be utilized to understand how the early universe originated and to find out the key parameters of the Big Bang model, see European Space Agency (2021b). One of the key assumptions of the modern cosmology is that the universe looks the same in any direction. It has been in debate for several years by using the CMB data. Numerous research suggested that the CMB data are either non-Gaussian or cannot be accurately described by statistical or mathematical models with few constant parameters, see Ade et al. (2016), Hill (2018), Leonenko et al. (2021), Marinucci (2004), Minkov et al. (2019) and Starck et al. (2004). The classical book by Weinberg (Weinberg (2008)) explained that this anisotropy in the plasma universe was significant enough to produce anisotropy in current galaxy distributions. For some recent results and discussion of fundamental cosmological models of the universe, see Broadbridge and Deutscher (2020). To detect departures from the isotropic model in actual CMB data several statistical approaches can be employed, see, for example, Leonenko et al. (2021) and Hamann et al. (2021). Different approaches can result in different results and suggest to cosmologists sky regions for further investigations. The motivation of this chapter is to check for multifractionality of the CMB data from the Planck mission.

Theoretical multifractional space-time models which differ from the standard cosmological model have been studied by Calcagni (2019) and Calcagni et al. (2016). They proposed several theoretical advancements using multifractional space-time that change its properties from place to place. It was suggested that the universe is not expanding monotonically which produces multifractional behaviour. In Calcagni et al. (2016), the CMB data from Planck mission and Far Infrared Absolute Spectrophotometer were used to establish speculative constraints on multifractional space-time expansion scenarios. Further, fractional SPDEs were employed to model the CMB data from Planck mission and study their changes, see Anh et al. (2018). The considered fractional SPDE models exhibited long-range dependence.

This chapter uses the theory of multifractional random fields and develops methodology to investigate fractional properties of random fields on the unit sphere. The presented detailed analysis of actual CMB data suggests the presence of multifractionality.

The developed methodology was also used to detect anomalies in CMB maps. The obtained results were compared with a different method from Hamann et al. (2021). It was demonstrated that the both methods can find same anomalies, but each method also can detect its own CMB regions of unusual behaviour. It was shown that applications of the developed methodology resulted in spatial clusters with high values of the proposed discrepancy statistics. The clusters matched very well with the TMASK of unreliable CMB intensities.

The structure of the chapter is as follows. Section 4.2 provides main notations and definitions related to the theory of random fields. Section 4.3 introduces the concept of multifractionality and discusses the GMBM. Section 4.4 presents results on the estimation of the pointwise Hölder exponent by using quadratic variations of random fields. Section 4.5 discusses the suggested estimation methodology. Numerical studies including computing the estimates of pointwise Hölder exponents for different one- and two-dimensional regions of the CMB sky sphere are given in Section 4.6. This section also demonstrates an application of the developed methodology to detect regions with anomalies in the cleaned CMB maps. Finally, the conclusions and some future research directions are presented in Section 4.7.

All numerical studies were carried out by using the software Python version 3.9.4 and R version 4.0.3, specifically, the R package RCOSMO (Fryer et al. (2020) , Fryer et al. (2019)). A reproducible version of the code in this chapter is available in Appendix B.

4.2 Main notations and definitions

This section presents background material in the theory of random fields, fractional spherical fields and fractional processes. Most of the material included in this section are based on Ayache (2018), García-Ancona et al. (2020), Herbin (2006), Lang and Schwab (2015), Malyarenko (2012) and Marinucci and Peccati (2011).

Let $\mathbb{R}^3$ be the real 3-dimensional Euclidean space and $s_2(1)$ be the unit sphere defined in $\mathbb{R}^3$. That is, $s_2(1) = \{x \in \mathbb{R}^3, \|x\| = 1\}$ where $\| \cdot \|$ represents the Euclidean distance in $\mathbb{R}^3$. Let $SO(3)$ denote the group of rotations on $\mathbb{R}^3$.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. The symbol $\stackrel{d}{=}$ denotes the equality in the sense of the finite-dimensional distributions.

Definition 4.1. A function $T(\omega, x) : \Omega \times s_2(1) \to \mathbb{R}$ is called a real-valued random field defined on the unit sphere. For simplicity, it will also be denoted by $T(x)$, $x \in s_2(1)$.

Definition 4.2. The random field $T(x)$ is called strongly isotropic if for all $k \in \mathbb{N}$, $x_1, \ldots, x_k \in s_2(1)$ and $g \in SO(3)$, the joint distributions of the random variables $T(x_1)$, $\ldots$, $T(x_k)$ and $T(gx_1), \ldots, T(gx_k)$ have the same law.

It is called 2-weakly isotropic (in the following, it will be just called isotropic) if the second moment of $T(x)$ is finite, i.e. if $E\Big(|T(x)|^2\Big) < +\infty$ for all $x \in s_2(1)$ and if for all pairs of points $x_1, x_2 \in s_2(1)$, and for any rotation, $g \in SO(3)$, it holds

$$E\Big(T(x)\Big) = E\Big(T(gx)\Big), \quad E\Big(T(x_1)\, T(x_2)\Big) = E\Big(T(gx_1)\, T(gx_2)\Big).$$

Definition 4.3. $T(x)$ is called Gaussian if for all $k \in \mathbb{N}$ and $x_1, \ldots, x_k \in s_2(1)$ the random variables $T(x_1), \ldots, T(x_k)$ are multivariate Gaussian distributed; that is, $\sum_{i=1}^{k} a_i T(x_i)$ is a normally distributed random variable for all $a_i \in \mathbb{R}$, $i = 1, \ldots, k$, such that $\sum_{i=1}^{k} a_i^2 \neq 0$.

Let $T = \{T(r, \theta, \varphi) : 0 \leq \theta \leq \pi, 0 \leq \varphi < 2\pi, r > 0\}$ be a spherical random field that has zero-mean, finite variance and is mean-square continuous. Let the corresponding Lebesgue measure on the unit sphere be $\sigma_1(du) = \sigma_1(d\theta \cdot d\varphi) = \sin\theta d\theta d\varphi$, $u = (\theta, \varphi) \in s_2(1)$. For two points on $s_2(1)$, we use Θ to denote the angle formed between two rays originating at the origin and pointing at these two points. Θ is called the angular distance between these two points. To emphasize that a random field depends on Euclidean coordinates, the notation $\tilde{T}(x) = T(r, \theta, \varphi)$, $x \in \mathbb{R}^3$, will be used.

Remark 4.1. A real-valued second order random field $\tilde{T}(x)$, $x \in s_2(1)$, with $E\Big(\tilde{T}(x)\Big) = 0$ is isotropic if $E\Big(\tilde{T}(x_1)\tilde{T}(x_2)\Big) = B(\cos\Theta)$, $x_1, x_2 \in s_2(1)$, depends only on the angular distance Θ between x_1 and x_2.

The spherical harmonics are defined by

$$Y_l^m(\theta,\varphi) = c_l^m \exp(im\varphi) P_l^m(\cos\theta), \quad l = 0, 1, ..., \ m = 0, \pm 1, ..., \pm l,$$

with
$$c_l^m = (-1)^m \left(\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right)^{1/2},$$

and the Legendre polynomials $P_l^m(\cos\theta)$ having degree l and order m.

Then the following spectral representation of spherical random fields holds in the mean-square sense:

$$T(r,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_l^m(\theta,\varphi) a_l^m(r),$$

where $a_l^m(r)$ is a set of random coefficients defined by

$$a_l^m(r) = \int_0^{\pi}\int_0^{2\pi} T(r,\theta,\varphi)\overline{Y_l^m(\theta,\varphi)} r^2 \sin\theta d\theta d\varphi = \int_{s_2(1)} \tilde{T}(ru)\overline{Y_l^m(u)}\sigma_1(du),$$

where $u = \frac{x}{\|x\|} \in s_2(1)$, $r = \|x\|$.

Definition 4.4. A real-valued random field $\tilde{T}(x)$, $x \in \mathbb{R}^3$, is with stationary increments if the equality

$$\tilde{T}(x + x') - \tilde{T}(x') \stackrel{d}{=} \tilde{T}(x) - \tilde{T}(0), \ x \in \mathbb{R}^3,$$

holds, for all $x' \in \mathbb{R}^3$.

Remark 4.2. When $\tilde{T}(x)$, $x \in \mathbb{R}^3$, is a second order random field with stationary increments, then one has,

$$E\Big(\tilde{T}(x + x') - \tilde{T}(x')\Big)^2 = \mathcal{V}_{\tilde{T}}(x), \quad \text{for every } (x, x') \in \mathbb{R}^3 \times \mathbb{R}^3,$$

where $\mathcal{V}_{\tilde{T}}$ is called the variogram of the field $\tilde{T}$.

Definition 4.5. A real-valued random field $\tilde{T}(x)$, $x \in \mathbb{R}^3$, is said to be globally self-similar, if for some fixed positive real number H and for each positive real number a, it satisfies

$$a^{-H}\tilde{T}(ax) \stackrel{d}{=} \tilde{T}(x), \ x \in \mathbb{R}^3. \tag{4.1}$$

Remark 4.3. Beside the degenerate case, the scale invariance property (4.1) holds only for a unique H which we declare as the global self-similarity exponent.

Definition 4.6. (Ayache (2018)) For each fixed $H \in (0,1)$, there exists a real-valued globally H-self-similar isotropic centered Gaussian field with stationary increments. Up to a multiplicative constant, this field is unique in distribution. It is called fractional Brownian field (FBF) of Hurst parameter H, and denoted by $B_H(t)$, $t \in \mathbb{R}^3$. The corresponding covariance function, is given, for all $(t', t'') \in \mathbb{R}^3 \times \mathbb{R}^3$, by

$$E\Big(B_H(t')\, B_H(t'')\Big) = 2^{-1}\,\mathrm{Var}\left(B_H(e_0)\right)\Big(\|t'\|^{2H} + \|t''\|^{2H} - \|t' - t''\|^{2H}\Big),$$

where e_0 denotes an arbitrary vector of the unit sphere $s_2(1)$.

Remark 4.4. In the particular case where $H = 1/2$, FBF is denoted by $B(t)$, $t \in \mathbb{R}^3$, and called Lévy Brownian Motion.

Similarly, one can introduce a H-self-similar process in the one-dimensional case. We also denote it by $B_H(t)$, $t \geq 0$. It will be called the fractional Brownian motion (FBM).

Definition 4.7. (Peltier and Véhel (1995)) The FBM with Hurst index $H(0 < H < 1)$ is defined as the stochastic integral

$$B_H(t) = \frac{1}{\Gamma(H+1/2)}\left\{\int_{-\infty}^{0}\Big((t-s)^{H-1/2} - (-s)^{H-1/2}\Big)\mathrm{d}W(s) + \int_0^t (t-s)^{H-1/2}\mathrm{d}W(s)\right\},$$

where $t \geq 0$ and $W(\cdot)$ denotes a Wiener process on $(-\infty, \infty)$.

The Hurst index H is also known as the Hurst parameter which specifies the degree of self-similarity. When $H = 0.5$, FBM reduces to the standard Brownian motion. In contrast to the Brownian motion, the increments of FBM are correlated. The FBM process $B_H(t)$ has the covariance function

$$Cov\Big(B_H(s), B_H(t)\Big) = \frac{\sigma^2}{2}\Big(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\Big).$$

The mean value of FBM is $E\Big(B_H(t)\Big) = 0$ and the variance function of FBM is $Var\Big(B_H(t)\Big) = \sigma^2|t|^{2H}/2$. The FBM has the following properties,

(i) Stationary increments: $B_H(t) - B_H(s) \stackrel{d}{=} B_H(t-s)$.

(ii) Long-range dependence of increments: $\sum_{n=1}^{\infty} E\Big(B_H(1)(B_H(n+1) - B_H(n))\Big)=\infty$, $H > 1/2$.

(iii) Self-similarity: $B_H(at) \stackrel{d}{=} |a|^H B_H(t)$.

4.3 Multifractional processes

This section provides definitions and theorems related to multifractional processes. Most of the material presented in this section is based on Benassi et al. (1998a), Ayache (2018), Peltier and Véhel (1995), Ayache (2013) and Benassi et al. (1997).

Let C^1 be the class of continuously differentiable functions and C^2 be the class of functions where both first and second derivatives exist and are continuous.

First, we introduce multifractional processes in the one-dimensional case. They will be used to analyse CMB data using the ring ordering HEALPix scheme.

Definition 4.8. (Benassi et al. (1998a)) Multifractional Gaussian processes (MGP) $X(t)$, $t \in [0,1]$, are real Gaussian processes whose covariance function $C(t,s)$ is of the form

$$C(t,s) = \int_{\mathbb{R}} f(t,\lambda)\overline{f(s,\lambda)}\mathrm{d}\lambda,$$

where

$$f(t,\lambda) = \frac{\left(\mathrm{e}^{it\lambda} - 1\right) a(t,\lambda)}{|\lambda|^{1/2+\alpha(t)}}.$$

The smoothness of the process is determined by the function $\alpha(\cdot)$ which is from C^1 with $0 < \alpha(t) < 1$, $t \in [0,1]$. The modulation of the process is determined by the function $a(t,\lambda)$ which is defined on $[0,1] \times \mathbb{R}$ and satisfies $a(t,\lambda) = a_\infty(t) + R(t,\lambda)$, where $a_\infty(\cdot)$ is $C^1([0,1])$ with, $a_\infty(t) \neq 0$ for all $t \in [0,1]$, and $R(\cdot,\cdot) \in C^{1,2}([0,1] \times \mathbb{R})$ is such that there exists some $\eta > 0$ that for $i = 0,1$ and $j = 0,2$ it holds

$$\left| \frac{\partial^{i+j}}{\partial t^i \partial \lambda^j} R(t,\lambda) \right| \leqslant \frac{C}{|\lambda|^{\eta+j}}.$$

Definition 4.9. (Peltier and Véhel (1995)) The multifractional Brownian motion (MBM) is given by

$$B_{H(t)}(t) = \frac{\sigma}{\Gamma(H(t)+1/2)} \left\{ \int_{-\infty}^{0} \left((t-s)^{H(t)-1/2} - (-s)^{H(t)-1/2}\right) \mathrm{d}B(s) \right.$$
$$\left. + \int_0^t (t-s)^{H(t)-1/2} \mathrm{d}B(s) \right\},$$

where $B(s)$ is the standard Brownian motion and $\sigma^2 = Var\Big(B_{H(t)}(t)\Big)|_{t=1}$.

For the MBM, $E\Big(B_{H(t)}(t)\Big) = 0$ and $Var\Big(B_{H(t)}(t)\Big) = \sigma^2 |t|^{2H(t)}/2$. The FBM is a special case of the MBM where the local Hölder exponent $H(t)$ is a constant, namely, $H(t) = H$. The MBM which is a non-stationary Gaussian process does not have independent stationary increments in contrast to the FBM.

Definition 4.10. A function $H(\cdot):\mathbb{R}\to\mathbb{R}$ is a (β,c)-Hölder function, $\beta>0$ and $c>0$, if $|H(t_1)-H(t_2)|\leqslant c\,|t_1-t_2|^{\beta}$, for all t_1,t_2 satisfying $|t_1-t_2|<1$.

The MBM admits the following harmonizable representation, see Benassi et al. (1997). If $H(\cdot):\mathbb{R}\to[a,b]\subset(0,1)$ is a β-Hölder function satisfying the assumption $\sup H(t)<\beta$, then the MBM with functional parameter $H(\cdot)$ can be written as $\operatorname{Re}\left(\int_{\mathbb{R}}\frac{\left(\mathrm{e}^{it\xi}-1\right)}{\|\xi\|^{H(t)+1/2}}\mathrm{d}\tilde{W}(\xi)\right)$, where $\tilde{W}(\cdot)$ is the complex isotropic random measure that satisfies $\mathrm{d}\tilde{W}(\cdot)=\mathrm{d}W_1(\cdot)+i\mathrm{d}W_2(\cdot)$. Here, $W_1(\cdot)$ and $W_2(\cdot)$ are independent real-valued Brownian measures.

Now we introduce the concept of the generalized multifractional Brownian motion (GMBM). The GMBM is an extension of the FBM and MBM. The GMBM was introduced to overcome the limitations existed in applying the MBM to model data whose pointwise Hölder exponent has an irregular behaviour.

The following definitions will be used to analyse the CMB data using the ring and nested ordering HEALPix schemes for $d=1,2$ respectively.

Definition 4.11. (Ayache and Véhel (2004)) Let $[a,b]\subset(0,1)$ be an arbitrary fixed interval. *An* admissible sequence $(H_n(\cdot))_{n\in\mathbb{N}}$ is *a* sequence of Lipschitz functions defined on $[0,1]$ and taking values in $[a,b]$ with Lipschitz constants δ_n such that $\delta_n\leqslant c_1 2^{n\alpha}$, for all $n\in\mathbb{N}$, where $c_1>0$ and $\alpha\in(0,a)$ are constants.

Definition 4.12. (Ayache and Véhel (2004)) Let $(H_n(\cdot))_{n\in\mathbb{N}}$ be an admissible sequence. The generalized multifractional field (GMF) with the parameter sequence $(H_n(\cdot))_{n\in\mathbb{N}}$ is the continuous Gaussian field $Y(x,y),\ (x,y)\in[0,1]^d\times[0,1]^d$ defined for all (x,y) as

$$Y(x,y)=\operatorname{Re}\left(\int_{\mathbb{R}^d}\left(\sum_{n=0}^{\infty}\frac{\left(\mathrm{c}^{ix\xi}-1\right)}{\|\xi\|^{H_n(y)+1/2}}\hat{f}_{n-1}(\xi)\right)\mathrm{d}\tilde{W}(\xi)\right),$$

where $\tilde{W}(\cdot)$ is the stochastic measure defined previously.

The GMBM with the parameter sequence $(H_n(\cdot))_{n\in\mathbb{N}}$ is the continuous Gaussian process $X(t),\ t\in[0,1]^d$ defined as the restriction of $Y(x,y),\ (x,y)\in[0,1]^d\times[0,1]^d$ to the diagonal: $X(t)=Y(t,t)$.

Compared to the FBM and MBM, one of the major advantages of the GMBM is that its pointwise Hölder exponent can be defined through the parameter $(H_n(\cdot))_{n\in\mathbb{N}}$. For every $t\in\mathbb{R}^2$, almost surely,

$$\alpha_X(t)=H(t)=\liminf_{n\to\infty}H_n(t).$$

4.4 The Hölder exponent

This section presents basic notations, definitions and theorems associated with the pointwise Hölder exponent, see Ayache and Véhel (2004), Benassi et al. (1998a), and Istas and Lang (1997a) for additional details. The pointwise Hölder exponent determines the regularity of a stochastic process. It describes local scaling properties of random fields and can be used to detect multifractionality.

Definition 4.13. The pointwise Hölder exponent of a stochastic process $X(t)$, $t \in \mathbb{R}$, whose trajectories are continuous, is the stochastic process $\alpha_X(t)$, $t \in \mathbb{R}$, defined by

$$\alpha_X(t) = \sup\left\{\alpha : \limsup_{h\to 0} \frac{|X(t+h) - X(t)|}{|h|^\alpha} = 0\right\}.$$

The Hölder regularity of FBM can be specified at any given point t, almost surely and $\alpha_{B_H}(t) = H$ is constant for FBM. The pointwise Hölder regularity of MBM can be determined by its functional parameter similar to FBM where $\alpha_X(t)$ is the pointwise Hölder exponent. Particularly, for every $t \in \mathbb{R}$, almost surely, $\alpha_X(t) = H(t)$.

In the literature, the method of quadratic variations is a frequently used technique to estimate the Hölder exponent, see Benassi et al. (1998a) and Istas and Lang (1997a). The following definition is used to compute the total increment in the one-dimensional case and will be applied for the ring ordering scheme of HEALPix points.

Definition 4.14. Let $t \in [0,1]$. For every integer $N \geq 2$, the generalized quadratic variation $V_N^{(1)}(t)$ around t is defined by

$$V_N^{(1)}(t) = \sum_{p\in v_N(t)} \left(\sum_{k\in F} e_k X\left(\frac{p+k}{N}\right)\right)^2, \tag{4.2}$$

where $F = \{0,1,2\}$, $e_0 = 1$, $e_1 = -2$ and $e_2 = 1$ and $v_N(t) = \{p \in \mathbb{N} : 0 \leqslant p \leqslant N-2$ and $|t - p/N| \leqslant N^{-\gamma}\}$.

The following definition is used to compute the total increment in the two-dimensional case and will be used for the nested ordering scheme of HEALPix points.

Definition 4.15. Let $t = (t_1, t_2) \in [0,1]^2$. For every integer $N \geq 2$, the generalized quadratic variation $V_N^{(2)}(t)$ around t is defined by

$$V_N^{(2)}(t) = \sum_{p\in v_N(t)} \left(\sum_{k\in F} d_k X\left(\frac{p+k}{N}\right)\right)^2, \tag{4.3}$$

where $p = (p_1, p_2)$, $\varepsilon = (\varepsilon_1, \varepsilon_2)$ and $(p+\varepsilon)/N = ((p_1+\varepsilon_1)/N, (p_2+\varepsilon_2)/N)$, $F = \{0, 1, 2\}^2$ and for all $k = (k_1, k_2) \in F$, $d_k = \prod_{l=1}^{2} e_{k_l}$ with $e_0 = 1$, $e_1 = -2$ and $e_2 = 1$. Here, $v_N(t) = v_N^1(t_1) \times v_N^2(t_2)$ and for all $i = 1, 2$, $v_N^i(t_i) = \{p_i \in \mathbb{N} : 0 \leqslant p_i \leqslant N-2$ and $|t_i - p_i/N| \leqslant N^{-\gamma}\}$.

The pointwise Hölder exponents are estimated for the one-dimensional ring ordering and two-dimensional nested ordering of HEALPix points by considering sufficiently large N and $d = 1, 2$ respectively in the following Theorem which is a specification of Theorem 2.2 in Ayache and Véhel (2004) with $\delta = 1$.

Theorem 4.1. *(Ayache and Véhel (2004)) Let $X(t)$, $t \in [0,1]^d$, be a GMBM with an admissible sequence $(H_n(\cdot))_{n\in\mathbb{N}}$ ranging in $[a,b] \subset (0, 1-1/2d)$. Then, for a fixed $\gamma \in (b, 1-1/2d)$ and the sequence $(H_n(t))_{n\in\mathbb{N}}$ convergent to $H(t)$, it almost surely holds*

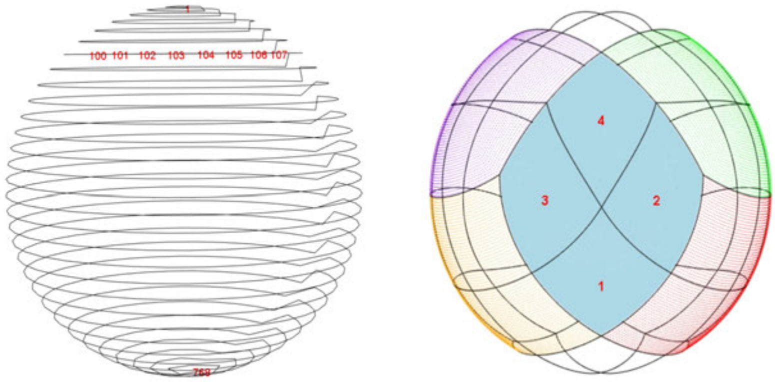
$$H(t) = \lim_{N\to\infty} \frac{1}{2}\left(d(1-\gamma) - \frac{\log\left(V_N^{(d)}(t)\right)}{\log(N)}\right). \tag{4.4}$$

4.5 Methodology

This section describes the suggested estimation methodology to study multifractionality of the CMB data that is based on theoretical results from Section 4.4. This and the next section also provide a detailed justification of this methodology and its assumptions and required modifications of the formulas for the spherical case and CMB data.

For multifractional data, $H(t)$ changes from location to location and $H(t) \not\equiv \text{const}$, where $t \in s_2(1)$. Several methods to estimate the local Hölder exponent are available in the literature. Different methods often give different results, see, for example, discussions in Bianchi (2005) and Struzik (2000) regarding inconsistent estimation results of the Hölder exponent. Inconsistent results by different techniques are due to their different assumptions, see Bianchi (2005). We propose an estimation method based on the generalized quadratic variations given by (4.2) and (4.3) and their asymptotic behaviour in (4.4). The results of this method are also compared with another conventional method that uses the rescale range (R/S) to estimate the Hölder exponent. This method is realized in the R package FRACTAL(Constantine and Percival (2017)).

Estimates of pointwise Hölder exponent values were computed using one- and two-dimensional regions of the the CMB data and the HEALPix ring and nested orderings (Gorski et al. (2005)). These HEALPix ordering schemes are shown in Figure 4.1. For the both cases, the highest available resolution, $N_{side} = 2048$ was used.



(a) HEALPix ring ordering (b) HEALPix nested ordering

Figure 4.1: HEALPix ordering schemes

The CMB data exhibit variations of the temperature intensities at very small scales ($\pm$ 1.8557 $\times 10^{-3}$). To get reliable estimates of $H(t)$, a large amount of observations in neighbourhoods of each t is required. Thus, in this publication, we do not discuss the preciseness of the local estimators of $H(t)$, but only pay attention to differences in the estimated values at different locations.

For computing purposes, the temperature intensities were scaled as

$$\text{Scaled Intensity}(t) = \frac{\text{Intensity}(t)}{\max_{s \in s_2(1)} |\text{Intensity}(s)|}.$$

It is clear from Definition 4.13 that this scaling does not change the values of $\alpha_X(t)$. Also, by (4.2) and (4.3) the generalized quadratic variation of the scaled process $cX(t)$ is $c^2 V_N^{(d)}(t)$, $d = 1, 2$. By (4.4),

$$\lim_{N\to\infty} \frac{\log\left(c^2 V_N^{(d)}(t)\right)}{\log(N)} = \lim_{N\to\infty} \left(\frac{\log(c^2)}{\log(N)} + \frac{\log\left(V_N^{(d)}(t)\right)}{\log(N)} \right) = \lim_{N\to\infty} \frac{\log\left(V_N^{(d)}(t)\right)}{\log(N)}, \tag{4.5}$$

which means that this scaling also does not affect $H(t)$.

As it was mentioned before, for small values of $\log(N)$ the estimates of $H(t)$ can be biased, which is now evident by the term $\frac{\log(c^2)}{\log(N)}$ in (4.5). However, this bias is due to the scaling effect only and is exactly the same for all values of t. Even if it might result in some errors in estimates $\hat{H}(t)$, it will not effect the analysis of differences in $H(t)$ values for different locations, which is the main aim of this analysis.

4.6 Numerical studies

This section presents numerical studies and applications of the estimation methodology from Section 4.5 to CMB data. The pointwise Hölder exponent estimates $\hat{H}(t)$ are computed and analysed for one- and two-dimensional regions of CMB data acquired from the NASA/IPAC Infrared Science Archive (IRSA (2021)). The estimated Hölder exponents are used to quantify roughness of the CMB data. The developed methodology is also applied to detect possible anomalies in the CMB data.

4.6.1 Estimates of Hölder exponent for one-dimensional CMB regions

For the one-dimensional case, the HEALPix ring ordered CMB temperature intensities were modelled by a stochastic process $X(t)$. Their Hölder exponents $H(t)$ were estimated by using the expression from the equation (4.4) for the given large N with $d = 1$, where $V_N^{(1)}(t)$ was computed using the equation (4.2), which can be explicitly written as

$$V_N^{(1)}(t) = \sum_{p=0}^{N-2} \left(X\left(\frac{p}{N}\right) - 2X\left(\frac{p+1}{N}\right) + X\left(\frac{p+2}{N}\right)\right)^2 .$$

As pixels on relatively small ring segments can be considered lying on approximately straight lines, the results from the case $d = 1$ can be used. The parameter N was chosen to give approximately the number of pixels within a half ring of the CMB sky sphere. The parameter r is the distance from a HEALPix point t that is the center of an interval in which we compute the total increment $V_N^{(1)}(t)$. By the expression of $v_N(t)$ in Definition 4.14, the parameter γ was computed according to the formula, $\gamma = -\left(\log(r)/\log(N)\right)$ for selected values of N and r. Then, it was used in the equation (4.4) to compute the estimated pointwise Hölder exponent values.

According to the HEALPix structure of the CMB data with the resolution $N_{side} = 2048$, there are 50331648 pixels on the CMB sky sphere. The HEALPix ring ordering scheme results in $4 \times N_{side} - 1$ rings, see Hivon (2021). That is, for $N_{side} = 2048$, the CMB sky sphere consists of 8191 rings. Based on the HEALPix geometry, the number of pixels in the upper part rings increase with the ring number, Ring $= 1, \ldots, 2047$, as $(4 \times \text{Ring})$. The $(2N_{side} + 1) = 4097$ set of rings in the middle part of the CMB sky sphere have equal number of pixels which is $4 \times N_{side}$. The number of pixels in each of the final $(N_{side} - 1) = 2047$ rings in the lower part decreases according to the formula $(4 \times (8191 - \text{Ring} + 1))$.

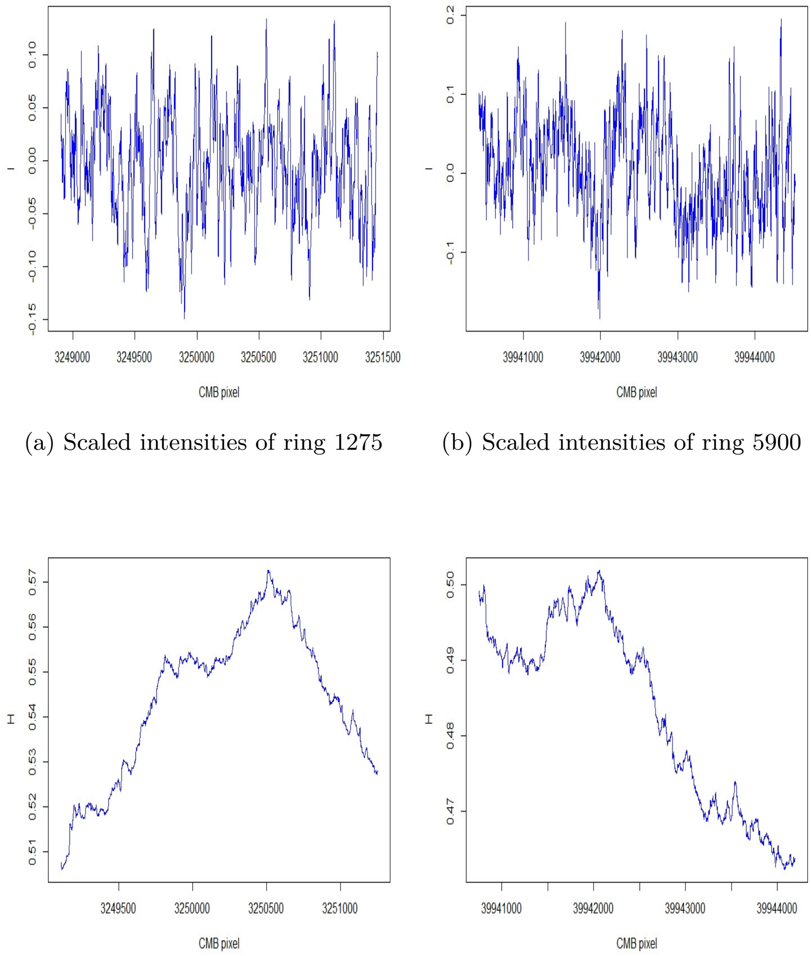
For the one-dimensional case, the estimated pointwise Hölder exponent values $\hat{H}(t)$ were computed as follows. First, a random CMB pixel was selected and its ring was determined. Then pixels belonging to the half of that particular ring were selected. Then, for each CMB pixel in this rim segment, the quadratic variation was computed by $V_N^{(1)}(t)$ given in equation (4.2). When computing the generalized quadratic variation for a CMB pixel, the pixels within a distance $r = 0.08$ from it were considered. For these pixels, the squared increments were computed and used to obtain the total of increments. Finally, the Hölder exponents were estimated by substituting the total of increments and the other parameters in the equation (4.4).

First, three CMB pixels '552300', '1533000', '3253800' located in the corresponding upper part rings 525, 875 and 1275 were chosen. Then for each CMB pixel in these half rings, their corresponding estimated Hölder exponents $\hat{H}(t)$ were computed. Next, another three pixels '10047488', '32575488', '39948288' were chosen in the middle part of the CMB sky sphere. Their ring numbers were 2250, 5000 and 5900 respectively. Finally, three CMB pixels '47656664', '48651704', '49375304' belonging to the lower part rings, 7035, 7275 and 7500 were selected and the pointwise Hölder exponents of pixels in their rim segments were estimated.

For example, Figure 4.2 shows the plots of the scaled intensities and the estimated pointwise Hölder exponents of the rim segments of rings 1275 and 5900, which belong to the upper and middle parts of the CMB sky sphere respectively. It can be seen from Figures 4.2a and 4.2b that the majority of scaled intensities fall into the range $[-0.2, 0.2]$ and their fluctuations are random. Figures 4.2c and 4.2d exhibit that the $\hat{H}(t)$ values in both rim sections are changing and the dispersion range for ring 1275 is wider than that of ring 5900. Similar plots and results were also obtained for other rings.

Part of CMB sky	**Case**	**Ring number**	γ	$[\hat{H}(t)_{\min}, \hat{H}(t)_{\max}]$	$\hat{H}(t)_{\max}$-$\hat{H}(t)_{\min}$	**Mean** $\hat{H}(t)$
Upper part	1	525	0.3631	[0.5681, 0.6215]	0.0534	0.5960
	2	875	0.3382	[0.5443, 0.5782]	0.0339	0.5605
	3	1275	0.3220	[0.5059, 0.5727]	0.0668	0.5439
Middle part	4	2250	0.3037	[0.4824, 0.5479]	0.0655	0.5137
	5	5000	0.3037	[0.4372, 0.4847]	0.0475	0.4626
	6	5900	0.3037	[0.4622, 0.5019]	0.0397	0.4835
Lower part	7	7035	0.3260	[0.5067, 0.5384]	0.0317	0.5234
	8	7275	0.3361	[0.5256, 0.5553]	0.0297	0.5410
	9	7500	0.3492	[0.5548, 0.5896]	0.0348	0.5701

Table 4.1: Summary of $\hat{H}(t)$ values for pixels in different rings of the CMB sky sphere



(a) Scaled intensities of ring 1275 (b) Scaled intensities of ring 5900

(c) $\hat{H}(t)$ values of ring 1275 (d) $\hat{H}(t)$ values of ring 5900

Figure 4.2: Examples of scaled intensities and $\hat{H}(t)$ values for one-dimensional CMB regions

The summary of the estimated pointwise Hölder exponent values obtained by the discussed methodology is shown in Table 4.1. It is clear that the dispersion range of the $\hat{H}(t)$ values and the mean $\hat{H}(t)$ value change with ring numbers. These results suggest that the pointwise Hölder exponent values change from location to location. The summary of the estimated pointwise Hölder exponent values obtained by the conventional (R/S) method using the command "RoverS" from the R package FRACTAL is given in Table 4.2. It can be seen that the dispersion range and the mean $\hat{H}(t)$ value change with the spiraling ring number. Similar results were also obtained for other available estimators of the Hölder exponent. Although these numerical values are inconsistent between different methods, they all suggest that the pointwise Hölder exponent values change from location to location.

It is expected that temperature intensities are positively dependent/correlated in close

Part of CMB sky	Case	Ring number	$[\hat{H}(t)_{\min}, \hat{H}(t)_{\max}]$	$\hat{H}(t)_{\max}$-$\hat{H}(t)_{\min}$	Mean $\hat{H}(t)$
Upper part	1	525	[0.8106, 0.9035]	0.0929	0.8758
	2	875	[0.8527, 0.9146]	0.0619	0.8867
	3	1275	[0.8577, 0.9088]	0.0511	0.8898
Middle part	4	2250	[0.8757, 0.9148]	0.0391	0.8975
	5	5000	[0.8656, 0.9079]	0.0423	0.8883
	6	5900	[0.8702, 0.9072]	0.0370	0.8926
Lower part	7	7035	[0.8617, 0.9081]	0.0464	0.8889
	8	7275	[0.8599, 0.9098]	0.0499	0.8897
	9	7500	[0.8348, 0.9004]	0.0656	0.8714

Table 4.2: Summary of $\hat{H}(t)$ values for pixels in different rings of the CMB sky sphere using the R/S method

regions, see the covariance analysis in Broadbridge et al. (2019). Therefore, running standard equality of means tests under independence assumptions will provide even more significant results if the hypothesis of equal means is rejected.

To prove that distributions of $\hat{H}(t)$ are statistically different between different sky regions, we carried out several equality of means tests. Before that, the Shapiro test was used to ensure that the $\hat{H}(t)$ values satisfy the normality assumption. For all the considered cases in Table 4.1, their $\hat{H}(t)$ values failed the normality assumption. Since the CMB pixels close to each other can be dependent, to get more reliable results we chose CMB pixels at distance 50 apart on a ring. The Shapiro test confirmed that in all the considered upper and lower part cases in Table 4.1, $\hat{H}(t)$ values at step 50 satisfied the normality assumption, whereas the $\hat{H}(t)$ values in the middle part failed the normality assumption.

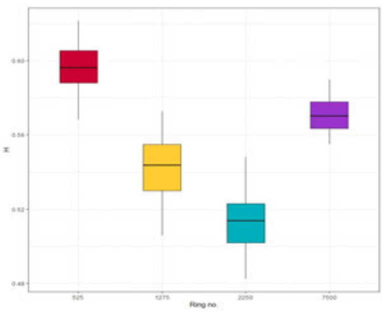


Figure 4.3: The distribution of $\hat{H}(t)$ values of four rim segments

1275	3.048×10^{-15}		
2250	7.939×10^{-11}	3.606×10^{-12}	
7500	1.533×10^{-8}	4.605×10^{-10}	3.717×10^{-13}
	525	1275	2250

Table 4.3: p-values for Wilcoxon tests between different rings

Let μ_1 and μ_2 be the mean($\hat{H}(t)$) values of the rim segments of rings 525 and 1275 respectively. To test the hypothesis $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$ we carried out the Wilcoxon test. The obtained p-value (3.048×10^{-15}) is significantly less than 0.05 and suggests that the means are different at 5% level of significance. Similar results were obtained for the Wilcoxon tests between all pairs of the cases in Table 4.1. For example, Table 4.3 shows Wilcoxon test results for selected four rings, two in the upper part, and the other two correspondingly in the middle and lower parts of the CMB sky sphere. Figure 4.3 shows the distribution box plots of the $\hat{H}(t)$ values in the rim segments of rings 525, 1275, 2250 and 7500. It is clear from Figure 4.3 that the mean($\hat{H}(t)$) values are different from each other in these cases.

Analogously to Table 4.3, for all carried out Wilcoxon tests between the rim sectors in the upper, middle and lower parts, their p-values < 0.05. Therefore, there is enough statistical evidence to suggest that the pointwise Hölder exponents change from location to location. While we compared Hölder exponents for different rings, from Figure 4.2 it is clear that $\hat{H}(t)$ is also changing for pixels within same rings.

4.6.2 Estimates of Hölder exponent for two-dimensional CMB regions

For two-dimensional sky regions, pointwise Hölder exponent values $H(t)$ were estimated according to the equation (4.4) with $d = 2$, where $V_N^{(2)}(t)$ was computed using the equation (4.3). The equation (4.3) in Definition 4.15 can be written in the following explicit form

$$
\begin{aligned}
V_N^{(2)}(t) &= \sum_{p\in v_N(t)} \left(\sum_{k_1\in\{0,1,2\}} \sum_{k_2\in\{0,1,2\}} e_{k_1} e_{k_2} X\left(\frac{p_1+k_1}{N}, \frac{p_2+k_2}{N}\right)\right)^2 \\
&= \sum_{p\in v_N(t)} \left(X\left(\frac{p_1}{N}, \frac{p_2}{N}\right) - 2X\left(\frac{p_1}{N}, \frac{p_2+1}{N}\right) - 2X\left(\frac{p_1+1}{N}, \frac{p_2}{N}\right)\right. \\
&+ X\left(\frac{p_1}{N}, \frac{p_2+2}{N}\right) + X\left(\frac{p_1+2}{N}, \frac{p_2}{N}\right) + 4X\left(\frac{p_1+1}{N}, \frac{p_2+1}{N}\right) \\
&\left. - 2X\left(\frac{p_1+1}{N}, \frac{p_2+2}{N}\right) - 2X\left(\frac{p_1+2}{N}, \frac{p_2+1}{N}\right) + X\left(\frac{p_1+2}{N}, \frac{p_2+2}{N}\right)\right)^2.
\end{aligned}
$$

To compute quadratic increments of spherical random fields, relatively small parts of the sphere can be approximately considered as regions of the plane and the above formula can be applied. Note that the internal summation set $\left\{\left(\frac{p_1+k_1}{N}, \frac{p_2+k_2}{N}\right) : k_1, k_2 \in \{0,1,2\}\right\}$ can be very efficiently represented by the HEALPix nested structure. Indeed, all pixels have either 7 or 8 neighbours, see Figure 4.4. The 3×3 configuration with 8 neighbours perfectly matches the internal summation set and can be directly used in computations of $V_N^2(t)$. For the case of 7 neighbours, an additional 8^{th} neighbour which intensity equals to the one of its adjusted pixel was imputed. For the resolution $N_{side} = 2048$ only 24 out of 50331648 pixels have 7 neighbours. For such small number of pixels the imputation has a negligible impact on the results.

Circular regions with radius $R = 0.23$ were used in this section computations. Let N denote the number of pixels within such circular regions. Then, $N \approx 662700$ pixels. To

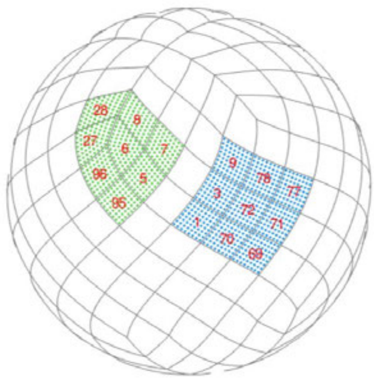
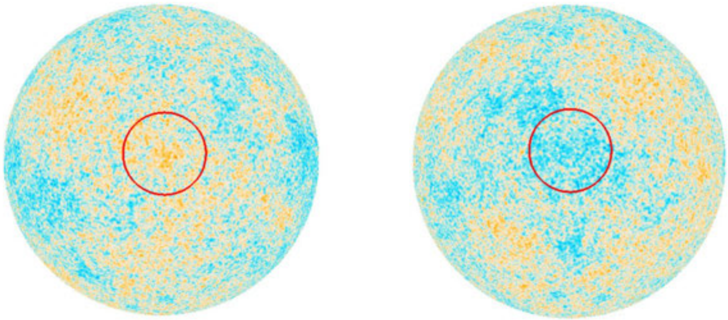


Figure 4.4: Examples of pixels with 7 and 8 neighbours for $N_{side} = 4$

reduce the computation time, we chose a grid of 1000 CMB pixels with the step $662 = [662700/1000]$, where $[\cdot]$ denotes the integer part, over the total number of pixels. To compute local estimators $\hat{H}(t)$, for each chosen CMB pixel, a circular window with radius $r = 0.01$ was selected. The value of γ was computed as $\gamma = -\left(\log(\sqrt{\pi}r/2)/\log(\sqrt{N})\right)$ for given values of N and r. The factor $\sqrt{\pi}/2$ appeared as the number of pixels is proportional to a window area. To match the number of pixels in circular window regions that were used in computations and square regions used for summation in $V_N^{(2)}(t)$, the length $2d_0$ of the side of squares should satisfy the equation $(2d_0)^2 = \pi r^2$. The obtained γ was substituted in the equation (4.4) to compute the estimated pointwise Hölder exponent values. For $r = 0.01$, there are approximately 2836 pixels in each specified window. For each of these pixels, the squared increment was computed and the total of increments was obtained by the expression for $V_N^{(2)}(t)$.

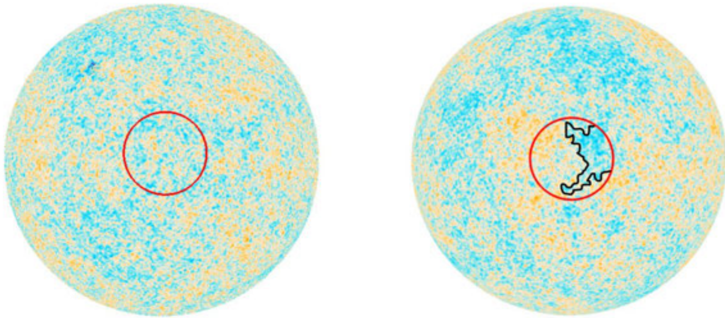
First, a circular CMB sky window of radius $R = 0.23$ from a warm area with the majority of high temperature intensities was selected. The mean temperature intensity in the selected CMB sky region covering the warm area was 5.97861×10^{-5}. The window is shown in Figure 4.5a. The number of pixels in that specific window was 662685. Then, different circular CMB sky windows having a radius of $R = 0.23$ covering cold, mixture of warm and cold regions and having a borderline of warm and cold regions shown in Figures 4.5b, 4.5c, and 4.5d were chosen. In each of the cold, mixture of warm and cold and a borderline having warm and cold regions, the number of pixels were 662697, 662706 and 662725 respectively. The value of γ was computed as $\gamma = 0.705$ for each CMB sky region. The corresponding mean temperature intensities were -8.34055×10^{-5}, -1.74035×10^{-5} and 7.59851×10^{-6}.

The plots of the estimated pointwise Hölder exponent values for each case are displayed in Figures 4.6a, 4.6b, 4.6c and 4.6d. These $\hat{H}(t)$ values are mostly dispersed in the interval



(a) A sky window from the warm region

(b) A sky window from the cold region



(c) A sky window with a mixture of temperatures

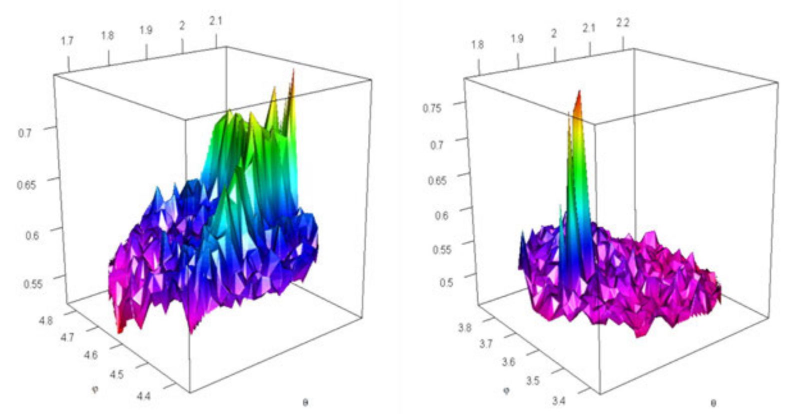
(d) A sky window with the borderline region

Figure 4.5: Sky windows used for computations

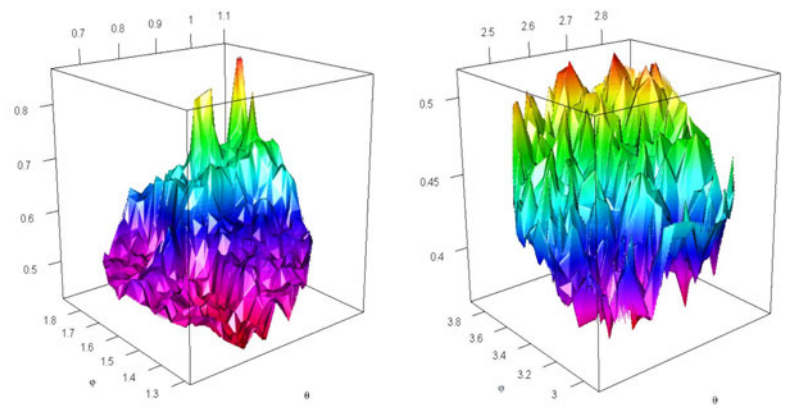
[0.36, 0.86]. Figures 4.6a, 4.6b, 4.6c and 4.6d show an erratic and an irregular behaviour in the distribution of $\hat{H}(t)$ values. It can be noticed that the estimates in Figures 4.6a and 4.6d with substantial warm temperatures have larger $\hat{H}(t)$ fluctuations than the $\hat{H}(t)$ values for cold regions.

The summary of the estimated pointwise Hölder exponents for each selected region is given in Table 4.4. It shows the mean CMB temperature intensities of each circular window. Table 4.4 also presents the estimated minimum, maximum and mean $\hat{H}(t)$ values computed by using the selected 1000 CMB pixels. It is clear from Table 4.4 that the mean $\hat{H}(t)$ value from the warm region is the highest and it is the lowest for the borderline region. The mean $\hat{H}(t)$ values of the cold region and mixture case lie in between them. It is apparent from Table 4.4 that the range of the estimated pointwise Hölder exponent values change with respect to the temperature of the chosen regions of the CMB sky sphere.

To further investigate the estimated pointwise Hölder exponents, they were computed for 100 random CMB pixels in each of the considered regions. It was apparent that even if



(a) $\hat{H}(t)$ values from the warm region (b) $\hat{H}(t)$ values from the cold region



(c) $\hat{H}(t)$ values from the region with mixture of temperatures (d) $\hat{H}(t)$ values from the borderline region

Figure 4.6: Local estimates $\hat{H}(t)$ for two-dimensional regions

one accounts for variation by considering these 100 CMB pixels, the $\hat{H}(t)$ values between different regions are different. The analyses suggested that all $\hat{H}(t)$ values for 100 and 1000 CMB pixels are consistent. Therefore, the results suggest that the estimated pointwise Hölder exponent values change from place to place.

Inspection Window	**Mean Intensity**	$[\hat{H}(t)_{\min}, \hat{H}(t)_{\max}]$	$\hat{H}(t)_{\min}$-$\hat{H}(t)_{\max}$	**Mean** $\hat{H}(t)$
Warm region	$5.97861 \cdot 10^{-5}$	[0.5217, 0.7484]	0.2267	0.5994
Cold region	$-8.34055 \cdot 10^{-5}$	[0.4534, 0.7806]	0.3272	0.5151
Mixture case	$-1.74035 \cdot 10^{-5}$	[0.4302, 0.8592]	0.4290	0.5563
Borderline case	$7.59851 \cdot 10^{-6}$	[0.3629, 0.5158]	0.1529	0.4407

Table 4.4: Analysis of CMB sky windows with different temperatures

To prove that $\hat{H}(t)$ is statistically different between different sky windows, we carried out several equality of means tests. Initially, we carried out the Shapiro test to ensure that the $\hat{H}(t)$ values satisfy the normality assumption. However, for all the considered cases in Table 4.4, their $\hat{H}(t)$ values failed the normality assumption. Figure 4.7 displays the distribution box plots of the $\hat{H}(t)$ values in the CMB sky windows with warm, cold, mixture of temperatures and having a borderline region. It can be noticed from Figure 4.7 that the $\hat{H}(t)$ distributions have extreme values in all the four cases. Thus, we present only results from the Wilcoxon test as it is reliable amidst the non-normality of data and in the presence of outliers.

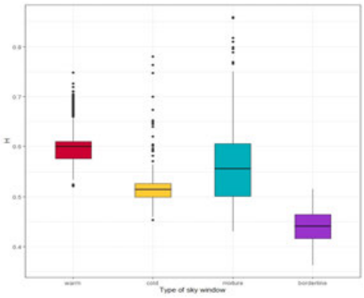


Figure 4.7: The distribution of $\hat{H}(t)$ values for chosen sky windows

	warm	cold	mixture
cold	$< 2.2 \times 10^{-16}$		
mixture	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$	
borderline	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$

Table 4.5: p-values for Wilcoxon tests between chosen sky windows

Let μ_1 and μ_2 be the mean($\hat{H}(t)$) values in the sky windows with warm and cold regions respectively. Testing the hypothesis $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$ by carrying out the Wilcoxon test, we obtained a p-value ($< 2.2 \times 10^{-16}$) that is significantly less than 0.05. It suggests that the means are different at 5% level of significance. Similar results were obtained for the Wilcoxon tests between all pairs of the cases and the corresponding p-values are shown in Table 4.5. It suggests that the mean($\hat{H}(t)$) values are different from each other in all the cases. Apart from variations between cases, it can be observed from Figure 4.6 and Table 4.4 that the estimated Hölder exponents do change within individual sky windows as well.

Therefore, there is enough statistical evidence to suggest that the pointwise Hölder exponents change from location to location of the CMB sky sphere.

4.6.3 Analysis of CMB temperature anomalies

As previously discussed in Section 4.1, several missions have measured the CMB temperature anisotropies gradually increasing their precision by using advanced radio telescopes. This section discusses applications of the multifractional methodology to detect regions of CMB maps with "anomalies". In particular, it can help in evaluating various reconstruction methods for blocked regions with unavailable or too noisy data.

It is well known that the CMB maps are affected by the interference coming from the Milky Way and radio signals emitting from our galaxy are much noisy than the CMB. Thus, the Milky Way blocks the CMB near the galactic plane. However, the smooth and predictable nature of Milky Way's radiation spectrum has enabled to disclose the cosmological attributes by subtracting the spectrum from the initially observed intensities (Castelvecchi (2019)). From Planck 2015 results, the CMB maps have been cleaned and reconstructed using different techniques namely, COMMANDER, NILC, SEVEM, SMICA see Ade et al. (2016) and Adam et al. (2016a) for more information. We are using the CMB map produced from the SMICA method (IRSA (2021)) with $N_{side} = 2048$.

To examine the random behaviour of isotropic Gaussian fields on the sphere, a direction dependent novel mathematical tool has been proposed in Hamann et al. (2021). They have applied their probe to investigate the CMB maps from Planck PR2 2015 and PR3 2018 with specific consideration to cosmological data from the inpainted maps. To detect departures from the traditional statistical model of the CMB data, they have utilized the auto-correlation of the sequence of full-sky Fourier coefficients and have proposed an "AC discrepancy" function on the sphere. For the inpainted Planck 2015 COMMANDER map, Hamann et al. (2021) shows the maximum "AC discrepancy" for the galactic coordinates[1], $(l, b) = (353.54, 1.79)$. Similarly, for the inpainted Planck COMMANDER 2018, NILC 2018, SEVEM 2018 and SMICA 2018 with $N_{side} = 1024$, there are significant departures at the galactic coordinates (12.57, 0.11), (61.17,-30.73), (261.25,-2.99) and (261.34,-2.99) respectively. A majority of these locations are the masked regions of the galactic plane. The galactic coordinates corresponding to the largest deviations are different for each map depicting the discrepancies in the underlying inpainting techniques.

The approach in Hamann et al. (2021) used directional dependencies in CMB data on the unit sphere. The results below are based on a different approach that uses the local roughness properties of these data. Therefore, the detected regions of high anomalies can be different for these two methods as they reflect different physical anisotropic properties of CMB, see, for example, Figure 4.10. The estimated local Hölder exponents on one-dimensional rings can be considered as directional local probes of CMB anisotropy. However, the estimates for two-dimensional regions are more complex and aggregate local information about roughness in different directions.

In the following analysis, we use estimated values of the Hölder exponent to detect regions of possible anomalies in CMB maps. Figure 4.8 shows the plots of scaled intensities and estimated Hölder exponent values $\hat{H}(t)$ in one- and two-dimensional CMB regions of

[1]The galactic coordinate system with Sun as the center is used in astronomy to locate the relative positions of objects and motions within the Milky Way Galaxy. It consists of galactic longitude $l, 0 \leq l < 2\pi$ and galactic latitude $b, -\pi/2 \leq b \leq \pi/2$. They are related to the spherical coordinates by $l = \phi$ and $b = (\pi/2 - \theta)$.

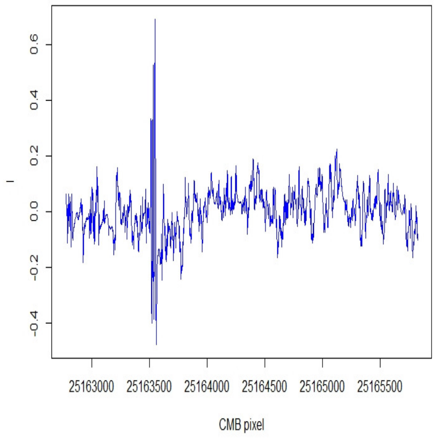
the great circle. It can be noticed from Figure 4.8a that there is an increase in the fluctuations of the scaled intensity values between the HEALPix range $[25163000, 25164000]$ of the great circle ring. A low plateau of estimated $\hat{H}(t)$ values in Figure 4.8b corresponds to that range of HEALPix values. The equator rim segment with the unusual plateau of $\hat{H}(t)$ values has CMB pixel numbers ranging from 25163208 to 25163852. Their corresponding galactic coordinates were found to be between, $(65.02, 0.01)$ and $(93.32, 0.01)$.

Similarly, this unusual behaviour of $\hat{H}(t)$ values was observed in the two-dimensional CMB regions near the galactic plane. Figure 4.8c shows the plot of scaled intensities in the two-dimensional space and a spike in intensities can be observed near the specified range of HEALPix values. The corresponding lower valley of $\hat{H}(t)$ values can be seen in Figure 4.8d. The four corners of the spherical region having unusual $\hat{H}(t)$ values have HEALPix values 23404309, 23391936, 23564929 and 24158424. Their galactic coordinates were found as $(85.91, -1.66)$, $(76.82, -1.66)$, $(76.82, 4.05)$ and $(85.91, 4.05)$ respectively.

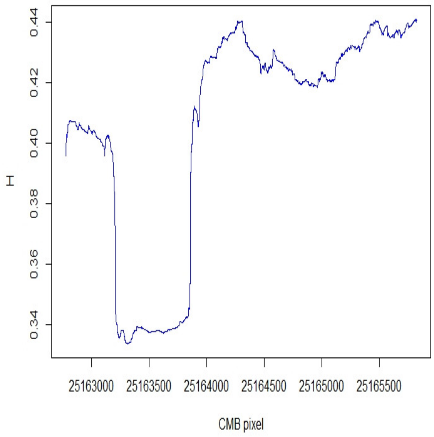
Table 4.6 shows the summary of CMB intensities at these one- and two-dimensional equatorial regions. The two-dimensional region around the unusual values was extracted as a rectangular spherical region from the circular CMB window using the previously identified galactic coordinates to split them as the unusual and the remaining regions. It is clear that the range of temperature intensities is wider in the one- and two-dimensional regions around the unusual values than in the regions excluding them. Further, the variances of intensities in the anomalous regions are larger than in the remaining regions. Moreover, Table 4.6 confirms that the mean $\hat{H}(t)$ values in the anomalous regions are lower than in the remaining regions.

Inspection Window	$[I_{\min}, I_{\max}]$ **(in** 10^{-3}**)**	$I_{\max} - I_{\min}$ **(in** 10^{-3}**)**	**Mean** I **(in** 10^{-5}**)**	**Variance** I **(in** 10^{-8}**)**	$[\hat{H}(t)_{\min}, \hat{H}(t)_{\max}]$	$\hat{H}(t)_{\min}$-$\hat{H}(t)_{\max}$	**Mean** $\hat{H}(t)$
One-dimensional region excluding the region of unusual values	[-0.3688,0.7578]	1.1266	1.4846	1.5654	[0.3351, 0.4411]	0.1060	0.4168
One-dimensional region around unusual values	[-0.8865,1.2851]	2.1716	-9.2156	4.9138	[0.3336, 0.3496]	0.0160	0.3384
Two-dimensional region excluding the region of unusual values	[-0.3935,0.2721]	0.6656	-3.6390	1.1433	[0.4097, 0.7448]	0.3351	0.5489
Two-dimensional region around unusual values	[-0.7310,0.2751]	1.0061	-6.3779	3.6978	[0.3015, 0.5975]	0.2960	0.4398

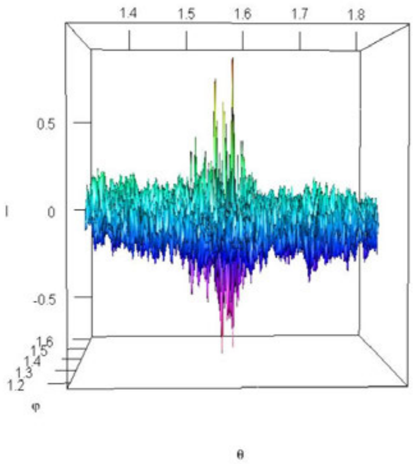
Table 4.6: Analysis of CMB intensities near the equatorial region



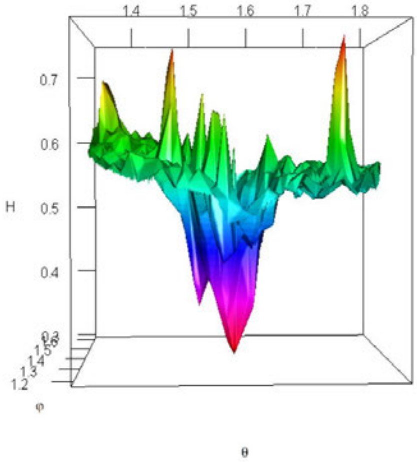
(a) Scaled intensities of great circle/ring 4096



(b) $\hat{H}(t)$ values of great circle/ring 4096

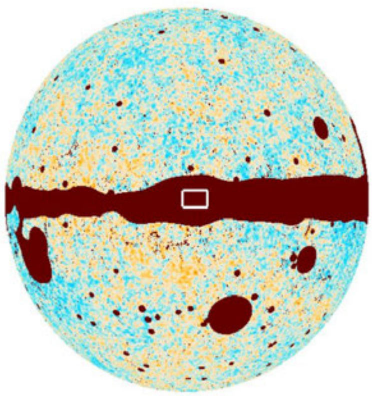


(c) Scaled intensities of equator region

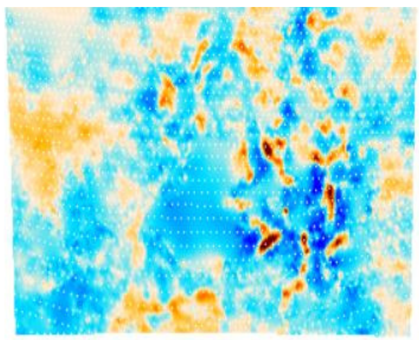


(d) $\hat{H}(t)$ values of equator region

Figure 4.8: Scaled intensities and estimated $\hat{H}(t)$ values in one- and two-dimensional regions of the great circle



(a) The anomalous sky window

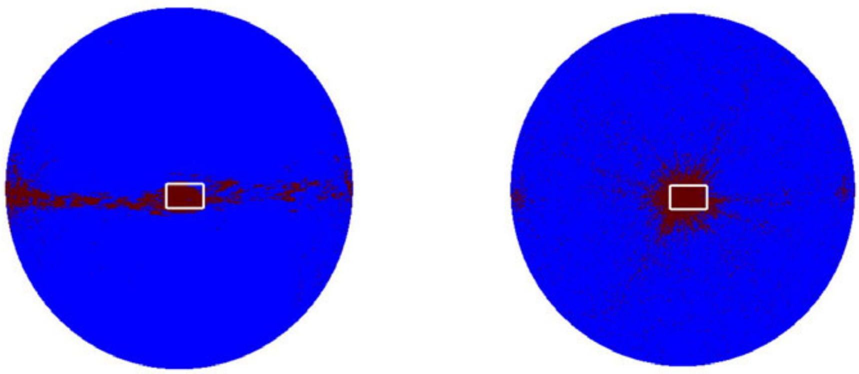


(b) The enlarged anomalous sky window

Figure 4.9: SMICA 2015 map with TMASK and the region of anomalies

Figure 4.9 shows the Planck 2015 map with blocked non-reliable CMB values. The region where TMASK applied by the SMICA reconstruction technique is removed in Figure 4.9. The TMASK of the CMB intensities utilized by the SMICA method determines the region where the inpainted CMB intensities in the galactic plane are considered to be reliable. The rectangular window shows a possible region of anomalies detected by the developed multifractional methodology.

Now we apply this approach and investigate $\hat{H}(t)$ for all $t \in s_2(1)$. First, the one-dimensional methodology was used. $\hat{H}(t)$ was estimated using the CMB intensities on rims, similar to the analysis in Figure 4.8a and 4.8b. The moving windows with 4096 consecutive pixels, which is approximately a half of a full ring, were used to obtain values of $\hat{H}(t)$. To clearly show local behaviours, after several trials, sets $v_N(t)$ with 61 HEALPix points, i.e. with the radius equals 30 pixels, were selected. The obtained results are shown in Figure 4.10a. To compare them with the AC discrepancy approach in Hamann et al. (2021), Figure 4.10b shows the corresponding map obtained by applying the direction-dependent probe. The code from Wang (2021) was used to compute values of AC discrepancies for SMICA 2015 CMB intensities. The first map highlights $\hat{H}(t)$ values below the 5^{th} percentile. AC discrepancy values above the 95^{th} percentile were used for the second map. The both approaches detected the region of anomalies in Figure 4.9. However, from locations of other discrepancy values, it is clear that these approaches detect different CMB anomalies.



(a) Hölder exponent approach (b) AC discrepancy approach

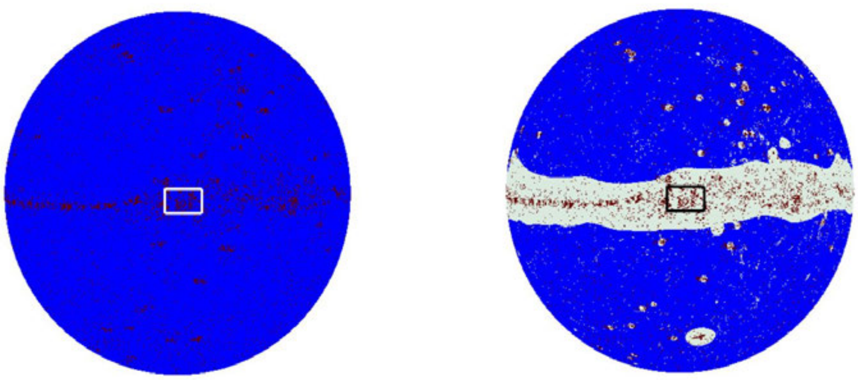
Figure 4.10: Discrepancy maps for CMB intensities from SMICA 2015

Very sharp changes in $\hat{H}(t)$ values in Figure 4.8b motivated the second method to detect anomalies, which is based on increments of $\hat{H}(t)$ values. Figure 4.8b demonstrated substantial changes of $\hat{H}(t)$ for nearby t locations. These changes are permanent as $\hat{H}(t)$

exhibits stable behaviour after a rapid "jump". Such changes are different from noise or outliers, when values in random distinct locations lay at an abnormal distance from other values in their surrounding points.

To detect such rapid changes, we used the statistics $\hat{H}_{\Delta}(t) = \min_{t_1 \in \Delta(t)} |\hat{H}(t) - \hat{H}(t_1)|$. where t and t_1 are indices of ring-ordered pixels and the set $\Delta(t) = \{t+10, ..., t+20\}$. The delay of 10 was selected to detect jumps that occur over short distances. The minimum over the set of consecutive points $\Delta(t)$ was used to eliminate outliers or noise that can result in distinct large differences $|\hat{H}(t) - \hat{H}(t_1)|$.

Figure 4.11a shows the computed $\hat{H}_{\Delta}(t)$ values for SMICA 2015 CMB intensities. $\hat{H}_{\Delta}(t)$ values above the 95^{th} percentile are plotted.



(a) $\hat{H}_{\Delta}$ discrepancy map (b) $\hat{H}_{\Delta}$ discrepancies over TMASK

Figure 4.11: $\hat{H}_{\Delta}$ discrepancy maps for CMB intensities from SMICA 2015

In Figure 4.11b, 5% of largest $\hat{H}_{\Delta}(t)$ values are shown on the TMASK map. It can be seen that in most cases, clusters of largest $\hat{H}_{\Delta}(t)$ values are within the TMASK. It seems that $\hat{H}_{\Delta}$ statistics rather accurately detected many regions with unreliable CMB values. Analysis of other CMB maps gave similar results.

Summarising, the implemented methodology to investigate multifractionality in CMB data, could also serve as a mechanism to detect regions of anomalies in CMB maps.

4.7 Conclusion

This chapter examined multifractional spherical random fields and their applications to analysis of cosmological data from the Planck mission. It estimated pointwise Hölder exponent values for the actual CMB data and checked for the presence of multifractionality. The estimators of pointwise Hölder exponents for one- and two-dimensional regions were obtained by using the ring and nested orderings of the HEALPix visualization structure.

The carried out analysis conveyed some multifractionality in the CMB data since the computed pointwise Hölder exponent values do change from place to place in the CMB sky sphere. The proposed approach was also applied to introduce statistics that were used for detecting regions with anomalies in CMB data. The developed methodology can be used for other spherical data.

Some numerical approaches that were used to speed up computations for big CMB data sets will be reported in detail in future publications. In future studies, it would be also interesting to:

- Develop the distribution theory for the estimators of $H(t)$;
- Develop hypothesis tests of equality of the local Hölder exponents taking into account the dependence structure of random fields;
- Investigate reliability and accuracy of various estimators of the Hölder exponent for CMB data;
- Study rates of convergence in Theorem 4.1;
- Investigate changes of the Hölder exponents depending on evolutions of random fields driven by SPDEs on the sphere, see Anh et al. (2018), Broadbridge et al. (2019), Broadbridge et al. (2020) and Restrepo et al. (2021);
- Study directional changes of the Hölder exponent by extending the obtained results for the conventional ring ordering to rings with arbitrary orientations;
- Apply the developed methodology to other spherical data, in particular, to new high-resolution CMB data from future CMB-S4 surveys (Abazajian et al. (2019));
- Explore relations between the locations of the detected CMB anomalies and other cosmic objects.

Chapter 5

Asymptotic normality of simultaneous estimators of cyclic long-memory processes

This chapter is based on the article, Ayache, A., Fradon, M., Nanayakkara, R., and Olenko, A. Asymptotic normality of simultaneous estimators of cyclic long-memory processes, which will appear in *Electronic Journal of Statistics*.

Due to artistic reasons, the format of this paper was changed in accordance with the style of the thesis. This did not change the main contents of the paper, but gave rise to slight changes in the paper layout.

5.1 Introduction

Time series with cyclic long-memory behaviours attracted increasing attention in recent years, see Alomari et al. (2020), Arteche (2020), Arteche and Robinson (1999), Arteche and Robinson (2000), del Barrio Castro and Rachinger (2021) and the references therein. It was due to importance of such time series in finance, hydrology, cosmology, internet modelling, and other applications to data with non-seasonal cyclicities, see Arteche (2020), Arteche and Robinson (1999), Artiach and Arteche (2011), Boubaker and Sghaier (2015), Ferrara and Guégan (2001) and Whitcher (2004). At the same time, various statistics of cyclic long-memory processes have complex asymptotic behaviour that has not yet been fully understood and investigated, see Hosoya (1997), Ivanov et al. (2013), Klykavka et al. (2012) and Olenko (2013).

To link characterizations of the long-memory phenomena in temporal and spectral domains researchers usually employ Abelian and Tauberian theorems. These results establish connections between asymptotics of covariance functions at the infinity and singularities of the corresponding spectral densities, see Klykavka et al. (2012) and Leonenko and Olenko (2013). The most frequent definition of long-memory in the literature is a hyperbolic-type decay of a non-integrable covariance function. While this classical long-memory dependence is often related to unboundedness of spectral densities at the origin, spectral singularities at nonzero frequencies can also result in hyperbolic-type oscillating non-integrable covariance functions. Such spectral representations can be used to simultaneously model cyclicity and long-memory.

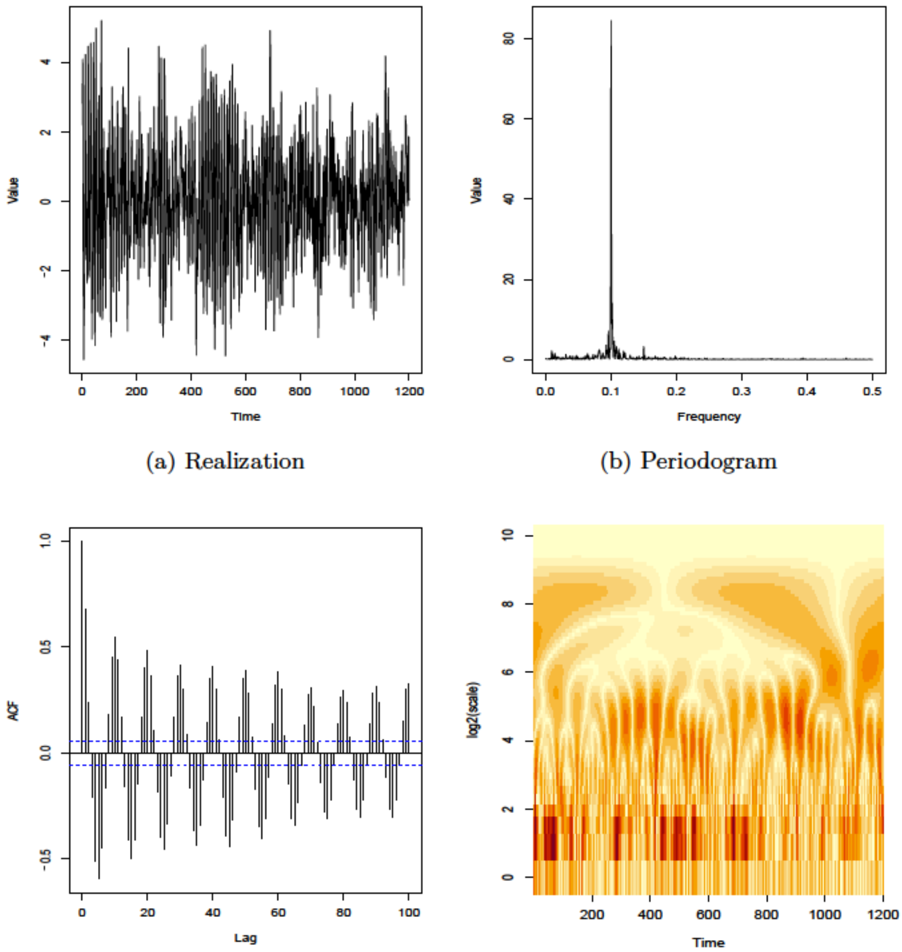
Cyclical long-memory time series are much more difficult to investigate and there were relatively few publications on this topic compared to classical models with the only singularity at the origin. Several least squares and likelihood-based approaches have been proposed to estimate parameters of singularity poles, see Arteche (2020), Arteche and Robinson (1999), Arteche and Robinson (2000), Barboza and Viens (2017), Beran et al. (2009), Espejo et al. (2015), Giraitis et al. (2001), Hidalgo (2005) and Tsai et al. (2015). Unfortunately, for the majority of these approaches incorrect specifications of a statistical model can result in inconsistent estimates of the parameters. The empirical studies in Beaumont and Smallwood (2019) and Whitcher (2004) demonstrated various issues of the traditional estimators and that wavelet-based approach can give results that are equivalent to ordinary least squares and maximum likelihood estimates under the assumption of knowing the explicit form of the spectrum. However, for the cases when the model is not fully specified, wavelets can provide better estimates.

To avoid repetitions, we refer the readers to very detailed motivation, discussion and various examples in Alomari et al. (2020).

This chapter investigates time series for which the spectral density $f(\cdot)$ has the following semiparametric form

$$f(\lambda) = \frac{h(\lambda)}{|\lambda^2 - s_0^2|^{2\alpha}}, \quad \lambda \in \mathbb{R}.$$

The parameter s_0 determines cyclic behaviour while α is a long-memory parameter. For example, the Gegenbauer model in Espejo et al. (2015) has a spectral density of this form.



(a) Realization

(b) Periodogram

(c) Sample covariance function

(d) Wavelet coefficients

Figure 5.1: Cyclic long-memory time series

Figure 5.1 shows a realization of such time series together with its estimated spectral density and covariance function. In this example a spectral density with a sharp spike at its singularity location was chosen. It clearly demonstrates that the spectral density has a singularity at a non-zero frequency and the corresponding covariance function indicates some cyclic behaviour. The wavelet coefficients of this time series are shown in the fourth subplot. Unfortunately, contrary to perfect cyclic signals or spectral densities with singularity at the origin, it is more difficult to use the wavelet approach for estimating cyclicity and long-memory parameters simultaneously. An even more challenging problem is a development of statistical inference for these parameters.

The publication Alomari et al. (2020) proposed a new methodology for simultaneous estimation of cyclic and long-memory parameters. It used filter transformations of functional time series. The approach included wavelet transformations as a particular case. The strong consistency of the proposed estimators was proved.

This chapter further develops the approach from Alomari et al. (2020). Now we obtain asymptotic normality of the proposed estimators. It requires very careful investigations of quadratic functionals of filter coefficients and their increments. Obtaining asymptotic properties of wavelet-based statistics is a difficult problem and there are only few general results about their asymptotic normality. The developed methodology and the obtained results can also find applications for other wavelet-based statistics.

In addition, for the case when empirical values of the statistics are outside the feasible region, we propose new adjusted estimators and investigate their properties. It is shown that these estimators have same asymptotic distributions as the corresponding ones in Alomari et al. (2020), but are computationally simpler.

The chapter is organized as follows. Section 5.2 gives basic definitions and introduces a semi-parametric model and filter transforms studied in this chapter. Various asymptotic properties of quadratic functionals of filter transforms are derived in Section 5.3. Section 5.4 proves asymptotic normality of two auxiliary statistics of the semiparametric model, which are based on quadratic functionals of filter transforms and their increments. Section 5.5 proposes and investigates adjusted simultaneous estimators of the location and long-memory parameters. Numerical studies to support the theoretical findings are presented in Section 5.6.

All computations, plotting and simulations in this chapter were performed using the software R version 4.0.3 and Maple 17, Maplesoft. In particular, the R packages WAVESLIM (Whitcher (2020)) and MassSpecWavelet (Du et al. (2006)) were used to simulate realizations of cyclic long-memory processes and compute their wavelet transforms in the numerical examples. A reproducible version of the code in this chapter is available in Appendix C.

5.2 Definitions and assumptions

This section introduces classes of functional time series and their filter transforms that are used in the chapter. The notations are consistent with ones in Alomari et al. (2020), where the authors proposed simultaneous filter estimators of parameters of cyclic long-memory processes.

In the following $\{a_j\}_{j\in\mathbb{N}}$ denotes an arbitrary unboundedly strictly monotone increasing sequence of positive real numbers. $\{m_j\}_{j\in\mathbb{N}}$ is a sequence of positive integers such that $\lim_{j\to+\infty} m_j = +\infty$. $\{b_{jk}\}_{(j,k)\in\mathbb{N}\times\mathbb{Z}}$ stands for an infinite array of real numbers.

The symbols $\xrightarrow{a.s.}$ and $\xrightarrow{d}$ will be used for almost sure convergence and convergence in distribution respectively.

Let $X(t)$, $t \in \mathbb{R}$, be a measurable mean-square continuous real-valued stationary zero-mean Gaussian stochastic process on a probability space $(\Omega, \mathcal{F}, P)$, with the covariance function

$$B(r) := \mathrm{Cov}(X(t), X(t')) = \int_{\mathbb{R}} \mathrm{e}^{iu(t-t')} F(du), \quad t, t' \in \mathbb{R},$$

where $r = t - t'$ and $F(\cdot)$ is a non-negative finite measure on $\mathbb{R}$.

Definition 5.1. The random process $X(t)$, $t \in \mathbb{R}$, possesses an absolutely continuous spectrum if there exists a non-negative function $f(\cdot) \in L_1(\mathbb{R})$ such that

$$F(u) = \int_{-\infty}^{u} f(\lambda) d\lambda, \quad u \in \mathbb{R}.$$

The function $f(\cdot)$ is called the spectral density of the process $X(t)$.

The process $X(t), t \in \mathbb{R}$, with an absolutely continuous spectrum has the following isonormal spectral representation

$$X(t) = \int_{\mathbb{R}} \mathrm{e}^{it\lambda} \sqrt{f(\lambda)} dW(\lambda),$$

where $W(\cdot)$ is a complex-valued Gaussian orthogonal random measure on $\mathbb{R}$.

For a real-valued stochastic process $X(t)$ the function $f(\cdot)$ is even and the random measure $W(\cdot)$ satisfies the condition $W([\lambda_1, \lambda_2]) = W([-\lambda_2, -\lambda_1])$ for any $\lambda_2 > \lambda_1 > 0$, see (Taqqu, 1979, §6).

The following assumption in the spectral domain introduces the semiparametric model investigated in this chapter.

Assumption 5.1. *Let the spectral density $f(\cdot)$ of $X(t)$ admit the following representation*

$$f(\lambda) = \frac{h(\lambda)}{|\lambda^2 - s_0^2|^{2\alpha}}, \quad \lambda \in \mathbb{R},$$

where $s_0 > 1$, $\alpha \in (0, 1/2)$ and $h(\cdot)$ is an even non-negative bounded function that is four times continuously differentiable. Its derivatives of order i satisfy $h^{(i)}(0) = 0$, $i = 1, 2, 3, 4$. Also, $h(0) = 1$, $h(\cdot) > 0$ in some neighborhood of $\lambda = \pm s_0$, and for all $\varepsilon > 0$ it holds

$$\int_{\mathbb{R}} \frac{h(\lambda)}{(1 + |\lambda|)^{\varepsilon}} d\lambda < \infty.$$

Stochastic processes with spectral densities satisfying Assumption 5.1 exhibit cyclic long memory. The boundedness of $h(\cdot)$ guarantees that their spectral densities have singularities only at the locations $\pm s_0$. Covariance functions of such processes are unintegrable and have hyperbolically decaying oscillations when $\alpha \in (0, 1/2)$, see Arteche and Robinson (1999). For example, the Gegenbauer random processes satisfy Assumption 5.1, see Espejo et al. (2015).

Real-valued functions $\psi(t) \in L_1(\mathbb{R})$, $t \in \mathbb{R}$, are used to introduce filter transforms of the process $X(t)$. The Fourier transform $\widehat{\psi}$ is defined, for each $\lambda \in \mathbb{R}$, as $\widehat{\psi}(\lambda) = \int_{\mathbb{R}} e^{-i\lambda t}\psi(t)dt$. It follows from properties of $\psi(\cdot)$ that $\widehat{\psi}(\cdot)$ is a bounded even function.

Assumption 5.2. *Let* $\operatorname{supp}\widehat{\psi} \subset [-A, A]$, $A > 0$, *and $\widehat{\psi}(\cdot)$ is of bounded variation on* $[-A, A]$.

This assumption is technical and can be replaced by a sufficiently fast decay rate of $\widehat{\psi}(\cdot)$ at infinity.

Definition 5.2. The filter transform of the process $X(t)$ is the array of centred real-valued Gaussian random variables $\{\delta_{jk}\}_{(j,k)\in\mathbb{N}\times\mathbb{Z}}$ defined as

$$\delta_{jk} := \frac{1}{\sqrt{a_j}} \int_{\mathbb{R}} \psi\left(\frac{t - b_{jk}}{a_j}\right) X(t)dt = \sqrt{a_j} \int_{\mathbb{R}} e^{ib_{jk}\xi} \frac{\overline{\widehat{\psi}(a_j\xi)}\sqrt{h(\xi)}}{|\xi^2 - s_0^2|^{\alpha}} dW(\xi). \tag{5.1}$$

Definition 5.2 provides equivalent expressions of the filter transform in the spectral and time domains.

It is easy to see that

$$\operatorname{Var}(\delta_{jk}) = a_j \int_{\mathbb{R}} \frac{|\widehat{\psi}(a_j\xi)|^2 h(\xi)}{|\xi^2 - s_0^2|^{2\alpha}} d\xi. \tag{5.2}$$

To guarantee that at each level $j \in \mathbb{N}$ the sequence $\{b_{jk}\}_{k\in\mathbb{Z}}$ does not have concentration points and covers all spectral range the following assumption is rather standard in the literature.

Assumption 5.3. *For all $j \in \mathbb{N}$ and for every $(k,l) \in \mathbb{Z}^2$ it holds*

$$|b_{jk} - b_{jl}| \geq \gamma_j |k - l|, \tag{5.3}$$

where $\{\gamma_j\}_{j\in\mathbb{N}}$ is a sequence of positive real numbers.

To get exact asymptotic behaviours of the considered statistics few versions of this assumption will be more precisely specified later.

A very detailed motivation, discussion, and various particular examples, that include wavelet transforms and Gegenbauer processes as special important cases, can be found in Alomari et al. (2020).

5.3 Preliminary results

This section derives some properties of the filter transforms and their variances that will be used in the following sections to obtain the CLT for simultaneous estimators of cyclic long-memory parameters.

Let

$$\delta_j^{(2,m_j)} := \sum_{k=1}^{m_j} \delta_{jk}^2, \quad j \in \mathbb{N}. \tag{5.4}$$

Theorem 5.1. *Assume that*

$$\lim_{j\to+\infty} \frac{a_j \ln(m_j)}{\gamma_j\, m_j^{1/2}} = 0. \tag{5.5}$$

Then, when $j \to +\infty$, the random variables

$$Y_j := \frac{\delta_j^{(2,m_j)} - \mathbb{E}(\delta_j^{(2,m_j)})}{\sqrt{\mathrm{Var}(\delta_j^{(2,m_j)})}} \tag{5.6}$$

converge in distribution to a standard Gaussian random variable.

To derive Theorem 5.1 we will use the following three lemmas. The first lemma is obtained by applying the Taylor-Lagrange formula, the second one is a rather known result and the third statement was proved in Alomari et al. (2020).

Let the function $\mathcal{I}_\zeta(\cdot)$, $\zeta \in \mathbb{R}$, be defined for $x \in \big[-(2A)^{-1}, (2A)^{-1}\big]$ as

$$\mathcal{I}_\zeta(x) := \int_{\mathbb{R}} e^{i\zeta\eta} \frac{|\widehat{\psi}(\eta)|^2 h(x\eta)}{\left(s_0^2 - x^2\eta^2\right)^{2\alpha}}\, d\eta. \tag{5.7}$$

Lemma 5.1. *If Assumptions 5.1 and 5.2 hold true, then $\mathcal{I}_\zeta(x)$ is four times continuously differentiable with respect to x, and there is a finite constant $c_1 > 0$ (not depending on ζ and x) such that, for all $\zeta \in \mathbb{R}$ and $|x| \le (2A)^{-1}$, it holds*

$$\left|\mathcal{I}_\zeta(x) - s_0^{-4\alpha}\int_{\mathbb{R}} e^{i\zeta\eta}|\widehat{\psi}(\eta)|^2\, d\eta - 2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}} e^{i\zeta\eta}\eta^2|\widehat{\psi}(\eta)|^2\, d\eta \cdot x^2\right| \le c_1\, x^4. \tag{5.8}$$

Proof of Lemma 5.1. Note that $\mathcal{I}_\zeta(\cdot)$ is a real-valued function since $\widehat{\psi}(\cdot)$ and $h(\cdot)$ are even real-valued functions. It follows from (5.7), Assumptions 5.1 and 5.2 that

$$\mathcal{I}_\zeta(x) = \int_{-A}^{A} e^{i\zeta\eta} \frac{|\widehat{\psi}(\eta)|^2 h(x\eta)}{\left(s_0^2 - x^2\eta^2\right)^{2\alpha}}\, d\eta = \int_{-A}^{A} e^{i\zeta\eta}|\widehat{\psi}(\eta)|^2 f(\eta x)\, d\eta.$$

To use the Taylor formula for $\mathcal{I}_\zeta(x)$ when $x \in \left[-(2A)^{-1}, (2A)^{-1}\right]$ one notes that $x \in \left[-(2A)^{-1}, (2A)^{-1}\right]$ and $\eta \in [-A, A]$ imply $|\eta x| \le 1/2$ and $s_0^2 - \eta^2 x^2 > 3/4$ since $s_0 > 1$. As by Assumption 5.1 the function $h(\cdot)$ is four times continuously differentiable, hence $f(\cdot)$ has four continuous derivatives with respect to x on $\left[-(2A)^{-1}, (2A)^{-1}\right]$ for any fixed η in $[-A, A]$. To prove that $I_\zeta(\cdot)$ is four times continuously differentiable, it is enough to show that the corresponding integrand and its first four derivatives with respect to x are dominated by integrable functions that do not depend on x.

First, for the integrand in (5.7) we get

$$\left|e^{i\zeta\eta}|\widehat{\psi}(\eta)|^2 f(\eta x)\right| \le \left(\frac{4}{3}\right)^{2\alpha} |\widehat{\psi}(\eta)|^2 \sup_{y\in[-1/2,1/2]} |h(y)|,$$

where the right hand side is bounded and therefore integrable on $[-A, A]$.

The n^{th} derivative of the function $f(\eta x)$ with respect to x satisfies

$$\left|\frac{\partial^n}{\partial x^n} f(\eta x)\right| = \left|\sum_{k=0}^{n} \binom{n}{k}\, \eta^{n-k}\, h^{(n-k)}(x\eta)\, \frac{\partial^k}{\partial x^k}\left((s_0^2 - \eta^2 x^2)^{-2\alpha}\right)\right|$$

$$\le \sum_{k=0}^{n} \binom{n}{k}\, A^{n-k} \sup_{y\in[-1/2,1/2]} \left|h^{(n-k)}(y)\right| \left|\frac{\partial^k}{\partial x^k}\left((s_0^2 - \eta^2 x^2)^{-2\alpha}\right)\right|.$$

For k in $\{1, 2, 3, 4\}$ we provide very simple convenient bounds for the derivatives in the last expression, which will be useful later:

$$\left|\frac{\partial}{\partial x}\left((s_0^2-\eta^2x^2)^{-2\alpha}\right)\right| = \left|\frac{4\alpha\eta^2 x}{(s_0^2-\eta^2x^2)^{2\alpha+1}}\right| \leq 4\alpha\frac{A}{2}\left(\frac{4}{3}\right)^{2\alpha+1} \leq 2A, \tag{5.9}$$

$$\left|\frac{\partial^2}{\partial x^2}\left((s_0^2-\eta^2x^2)^{-2\alpha}\right)\right| = \left|4\alpha\eta^2\ \frac{(4\alpha+1)\eta^2x^2+s_0^2}{(s_0^2-\eta^2x^2)^{2\alpha+2}}\right|$$

$$\leq 4\alpha A^2\left(\frac{4}{3}\right)^{2\alpha+2}\left(\frac{4\alpha+1}{4}+s_0^2\right) \leq 10A^2s_0^2, \tag{5.10}$$

$$\left|\frac{\partial^3}{\partial x^3}\left((s_0^2-\eta^2x^2)^{-2\alpha}\right)\right| = \left|8\alpha(2\alpha+1)\eta^4 x\ \frac{(4\alpha+1)\eta^2x^2+3s_0^2}{(s_0^2-\eta^2x^2)^{2\alpha+3}}\right| \leq 4\alpha(2\alpha+1)A^3$$

$$\times\left(\frac{4}{3}\right)^{2\alpha+3}\left(\frac{4\alpha+1}{4}+3s_0^2\right) \leq 4A^3\left(\frac{4}{3}\right)^4\left(\frac{3}{4}+3s_0^2\right) \leq 48A^3s_0^2, \tag{5.11}$$

$$\left|\frac{\partial^4}{\partial x^4}\left((s_0^2-\eta^2x^2)^{-2\alpha}\right)\right| = \left|\frac{(16\alpha(\alpha+1)+3)\eta^4x^4+6(4\alpha+3)s_0^2\eta^2x^2+3s_0^4}{(s_0^2-\eta^2x^2)^{2\alpha+4}}\right.$$

$$\left.\times 8\alpha(2\alpha+1)\eta^4\right| \leq 8\alpha(2\alpha+1)A^4\left(\frac{4}{3}\right)^{2\alpha+4}\left(\frac{16\alpha(\alpha+1)+3}{16}+\frac{6(4\alpha+3)}{4}s_0^2+3s_0^4\right)$$

$$\leq 8A^4\left(\frac{4}{3}\right)^5\left(\frac{15}{16}+\frac{15}{2}s_0^2+3s_0^4\right) \leq 400A^4s_0^4. \tag{5.12}$$

Therefore the function in the integral defining $\mathcal{I}_\zeta(\cdot)$ and its first four derivatives are dominated by an integrable function ($|\widehat{\psi}|^2$ multiplied by a large enough constant). Thus $\mathcal{I}_\zeta(\cdot)$ is $\mathcal{C}^4\left([-(2A)^{-1},(2A)^{-1}]\right)$ and its derivatives can be computed by differentiation under the integral sign. For n in $\{1,2,3,4\}$ it holds

$$\frac{d^n}{dx^n}\mathcal{I}_\zeta(x) = \sum_{k=0}^{n}\binom{n}{k}\int_{-A}^{A} e^{i\zeta\eta}|\widehat{\psi}(\eta)|^2\ \eta^{n-k}\ h^{(n-k)}(x\eta)\ \frac{\partial^k}{\partial x^k}\left((s_0^2-\eta^2x^2)^{-2\alpha}\right)\,d\eta \tag{5.13}$$

and the Taylor-Lagrange expansion provides

$$\left|\mathcal{I}_\zeta(x)-\mathcal{I}_\zeta(0)-\mathcal{I}_\zeta'(0)x-\mathcal{I}_\zeta''(0)\frac{x^2}{2!}-\mathcal{I}_\zeta^{(3)}(0)\frac{x^3}{3!}\right| \leq \sup_{y\in[-(2A)^{-1},(2A)^{-1}]}|\mathcal{I}_\zeta^{(4)}(y)|\ \frac{x^4}{4!}, \tag{5.14}$$

where

$$\mathcal{I}_\zeta(0) = \frac{1}{s_0^{4\alpha}}\int_{-A}^{A} e^{i\zeta\eta}|\widehat{\psi}(\eta)|^2\,d\eta,$$

since $h(0)=1$.

By Assumptions 5.1 the derivatives $h^{(l)}(0)=0$ for $l\in\{1,2,3,4\}$, thus

$$\frac{d^n}{dx^n}\mathcal{I}_\zeta(0) = \int_{-A}^{A} e^{i\zeta\eta}|\widehat{\psi}(\eta)|^2\ \frac{\partial^n}{\partial x^n}\left((s_0^2-\eta^2x^2)^{-2\alpha}\right)\Big|_{x=0}\,d\eta.$$

By (5.9) and (5.11) for $n = 1$ and $n = 3$ the derivatives $\frac{\partial^n}{\partial x^n}\left((s_0^2 - \eta^2 x^2)^{-2\alpha}\right)$ vanish at $x = 0$. Moreover, the expression for the second derivative in the estimate (5.10) gives

$$\frac{d^2}{dx^2}\mathcal{I}_\zeta(0) = \frac{4\alpha}{s_0^{4\alpha+2}} \int_{-A}^{A} e^{i\zeta\eta} |\widehat{\psi}(\eta)|^2 \eta^2 \, d\eta.$$

It follows from the estimates (5.9)-(5.12) that for each $k = 0, ..., 4$ the derivative $|\frac{\partial^k}{\partial x^k}(s_0^2 - \eta^2 x^2)^{-2\alpha}|$ is bounded by $400A^k s_0^4$. Hence, by (5.13), for all $x \in [-(2A)^{-1}, (2A)^{-1}]$

$$\left|\frac{d^4}{dx^4}\mathcal{I}_\zeta(x)\right| \le \sup_{\substack{y\in[-1/2,1/2]\\ n\in\{0,\dots,4\}}} |h^{(n)}(y)| \sum_{k=0}^{4} \binom{4}{k} \int_{-A}^{A} |\widehat{\psi}(\eta)|^2 A^{n-k} (400 A^k s_0^4) \, d\eta$$

$$\le 6400 \, s_0^4 \, A^4 \sup_{\substack{y\in[-1/2,1/2]\\ n\in\{0,\dots,4\}}} |h^{(n)}(y)| \int_{-A}^{A} |\widehat{\psi}(\eta)|^2 \, d\eta =: c_2.$$

Finally, the estimate (5.14) becomes

$$\left|\mathcal{I}_\zeta(x) - \frac{1}{s_0^{4\alpha}} \int_{-A}^{A} e^{i\zeta\eta} |\widehat{\psi}(\eta)|^2 \, d\eta - \frac{4\alpha}{s_0^{4\alpha+2}} \int_{-A}^{A} e^{i\zeta\eta} |\widehat{\psi}(\eta)|^2 \; \eta^2 \, d\eta \cdot \frac{x^2}{2!}\right| \le \frac{c_2}{4!} \cdot x^4,$$

which completes the proof. □

The following lemma is an immediate corollary of the Gershgorin circle theorem (Li and Zhang (2019)).

Lemma 5.2. *Let $U = (u_{ij})_{1\le i,j\le n}$ be a square matrix of order n with complex elements. If $\rho(U)$ is the spectral radius of U, that is*

$$\rho(U) := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } U\},$$

then

$$\rho(U) \le \min\Big\{\max_{1\le i\le n} \sum_{j=1}^{n} |u_{ij}|, \max_{1\le j\le n} \sum_{i=1}^{n} |u_{ij}|\Big\}.$$

Lemma 5.3. (Alomari et al. (2020)) *Let Assumptions 5.1 hold true. Then there exists a finite constant c_3 such that, for every $j \in \mathbb{N}$ such that $a_j \ge 2A$ and for all $(k, l) \in \mathbb{Z}^2$:*

$$\big|\mathrm{Cov}(\delta_{jk}, \delta_{jl})\big| \le c_3\Big(\mathbb{1}_{\{k=l\}} + \mathbb{1}_{\{k\ne l\}} \, a_j |b_{jk} - b_{jl}|^{-1}\Big). \tag{5.15}$$

Proof of Theorem 5.1. Note that $\delta_j^{(2,m_j)}$ is the squared Euclidian norm of the centred Gaussian vector $\vec{\delta}_j^{(m_j)} := (\delta_{j1}, \dots, \delta_{jm_j})$. Therefore, $\delta_j^{(2,m_j)}$ has the same distribution as

$\sum_{k=1}^{m_j} \lambda_{jk}\, \varepsilon_{jk}^2$, where $\lambda_{j1}, \ldots, \lambda_{jm_j}$ are the non-negative eigenvalues of the covariance matrix of $\vec{\delta}_j^{(m_j)}$ and $\varepsilon_{j1}, \ldots, \varepsilon_{jm_j}$ are independent standard Gaussian random variables. Thus, using a version of the Lindeberg condition (see for instance Csörgo and Révész (1981) or Lemma 2 in Istas and Lang (1997b)), it turns out that for proving the theorem it is enough to show that

$$\lim_{j\to+\infty} \frac{\max_{1\le k\le m_j} \lambda_{jk}}{\sqrt{\operatorname{Var}\left(\delta_j^{(2,m_j)}\right)}} = 0. \tag{5.16}$$

To derive (5.16) let us first prove that there is a positive constant c_4 (not depending on j), such that for all large enough j,

$$\operatorname{Var}(\delta_j^{(2,m_j)}) \ge c_4 m_j\,. \tag{5.17}$$

Using (5.4), (5.2) and the change of variable $\eta = a_j\xi$, one gets

$$\operatorname{Var}(\delta_j^{(2,m_j)}) = \sum_{k=1}^{m_j}\sum_{l=1}^{m_j} \operatorname{Cov}(\delta_{jk}^2, \delta_{jl}^2) = 2\sum_{k=1}^{m_j}\sum_{l=1}^{m_j} \operatorname{Cov}^2(\delta_{jk}, \delta_{jl}) \ge 2\sum_{k=1}^{m_j} \operatorname{Var}^2(\delta_{jk}) \tag{5.18}$$

$$= 2m_j\Big(a_j \int_{\mathbb{R}} \frac{|\widehat{\psi}(a_j\xi)|^2 h(\xi)}{|\xi^2 - s_0^2|^{2\alpha}}\, d\xi\Big)^2 = 2m_j\Big(\int_{\mathbb{R}} \frac{|\widehat{\psi}(\eta)|^2 h(a_j^{-1}\eta)}{|a_j^{-2}\eta^2 - s_0^2|^{2\alpha}}\, d\eta\Big)^2. \tag{5.19}$$

Moreover, it follows from (5.8) that

$$\lim_{j\to+\infty} \int_{\mathbb{R}} \frac{|\widehat{\psi}(\eta)|^2 h(a_j^{-1}\eta)}{|a_j^{-2}\eta^2 - s_0^2|^{2\alpha}}\, d\eta = s_0^{-4\alpha} \int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2\, d\eta > 0. \tag{5.20}$$

Then, (5.17) results from (5.19) and (5.20).

Next, by Lemma 5.2 for all $j \in \mathbb{N}$ it holds

$$\max_{1\le k\le m_j} \lambda_{jk} \le \max_{1\le k\le m_j} \sum_{l=1}^{m_j} |\operatorname{Cov}(\delta_{jk}, \delta_{jl})|. \tag{5.21}$$

Moreover, by (5.3) and (5.15), for each fixed large enough j and for every $k \in \{1, \ldots, m_j\}$, one has

$$\begin{aligned}
\sum_{l=1}^{m_j} |\operatorname{Cov}(\delta_{jk}, \delta_{jl})| &\le c_3\Big(1 + a_j \sum_{l=1,\, l\ne k}^{m_j} |b_{jk} - b_{jl}|^{-1}\Big) \\
&\le c_3\Big(1 + \frac{a_j}{\gamma_j} \sum_{l=1,\, l\ne k}^{m_j} |k-l|^{-1}\Big) \le c_3\Big(1 + \frac{2a_j}{\gamma_j} \sum_{l=1}^{m_j} l^{-1}\Big) \\
&\le c_3\Big(1 + \frac{2a_j}{\gamma_j} + \frac{2a_j}{\gamma_j}\int_1^{m_j} y^{-1}\, dy\Big) \le c_3\Big(1 + \frac{2a_j(1+\ln(m_j))}{\gamma_j}\Big).
\end{aligned} \tag{5.22}$$

Recall that the constant c_3 does not depend on (j,k,l). Finally, putting together (5.5), (5.17), (5.21), (5.22), and the fact that $\lim_{j\to+\infty} m_j = +\infty$, one gets (5.16). $\square$

To obtain the exact asymptotic variance of $\delta_j^{(2,m_j)}$ we specify asymptotic behaviours of the increments of the sequences $\{b_{jk}\}_{(j,k)\in\mathbb{N}\times\mathbb{Z}}$ in Assumption 5.3.

Assumption 3'. *For all $j\in\mathbb{N}$ and for every $(k,l)\in\mathbb{Z}^2$ it holds*

$$b_{jk} - b_{jl} = \gamma_j(k-l),$$

where $\{\gamma_j\}_{j\in\mathbb{N}}$ is a sequence of positive real numbers such that

$$\lim_{j\to+\infty}\frac{a_j}{\gamma_j} = c \in (0,+\infty) \quad \textit{and} \quad \lim_{j\to+\infty} m_j^2\left(\frac{\gamma_j}{a_j} - \frac{1}{c}\right) = 0.$$

Remark 5.1. For example, Assumption 3' is satisfied for the sequence $\{\gamma_j\}_{j\in\mathbb{N}}$ with $\gamma_j = a_j$ for all $j \geq j_0 \in \mathbb{N}$.

Lemma 5.4. *Let Assumption 3' hold true and*

$$\lim_{j\to+\infty} m_j a_j^{-8} = 0. \tag{5.23}$$

Then, the sequence of positive real numbers $\{\mathrm{Var}(\delta_j^{(2,m_j)})/m_j\}_{j\in\mathbb{N}}$ converges to a finite and strictly positive limit when $j\to+\infty$. More precisely,

$$\lim_{j\to+\infty}\frac{\mathrm{Var}(\delta_j^{(2,m_j)})}{m_j} = \mathcal{V}_1 := 4c\pi s_0^{-8\alpha}\int_{-c\pi}^{c\pi}\Big|\sum_{n\in\mathbb{Z}}\big|\widehat{\psi}(\eta+2nc\pi)\big|^2\Big|^2\,d\eta. \tag{5.24}$$

Proof of Lemma 5.4. Using (5.1), (5.7), (5.18), Assumption 3', and the change of variable $\eta = a_j\xi$ one obtains

$$\frac{\mathrm{Var}(\delta_j^{(2,m_j)})}{m_j} = \frac{2}{m_j}\sum_{k=1}^{m_j}\sum_{l=1}^{m_j}\mathcal{I}^2_{\gamma_j(k-l)/a_j}(a_j^{-1}), \quad j\in\mathbb{N},$$

where $\mathcal{I}_\zeta^2(\cdot)$ is the squared function $\mathcal{I}_\zeta(\cdot)$ defined in (5.7).

Let us denote by $F_j(\cdot)$ a bounded function defined on $[-\pi a_j/\gamma_j, \pi a_j/\gamma_j]$ as

$$F_j(\eta) := \sum_{n\in\mathbb{Z}}\big|\widehat{\psi}(\eta+2n\pi a_j/\gamma_j)\big|^2.$$

Let $\{\mu_j(k)\}_{k\in\mathbb{Z}}$ be the sequence of the Fourier coefficients of F_j. These coefficients are real-valued since $\widehat{\psi}(\cdot)$ is even. Using the fact that $\eta \mapsto e^{i\gamma_j k\eta/a_j}$ is, for each fixed $k\in\mathbb{Z}$, a $2\pi a_j/\gamma_j$-periodic function of η and the dominated convergence theorem, one gets

$$\mu_j(k) := \int_{-\pi a_j/\gamma_j}^{\pi a_j/\gamma_j} e^{i\gamma_j k\eta/a_j}\Big(\sum_{n\in\mathbb{Z}} \big|\widehat{\psi}(\eta+2n\pi a_j/\gamma_j)\big|^2\Big)\,d\eta = \int_{\mathbb{R}} e^{i\gamma_j k\eta/a_j}\big|\widehat{\psi}(\eta)\big|^2\,d\eta. \quad (5.25)$$

Now, let us show that there is a finite constant c_4 such that, for all j large enough, one has

$$m_j^{-1/2}\Big|\Big(\sum_{k=1}^{m_j}\sum_{l=1}^{m_j}\mathcal{I}^2_{\gamma_j(k-l)/a_j}(a_j^{-1})\Big)^{1/2} - \Big(\sum_{k=1}^{m_j}\sum_{l=1}^{m_j} s_0^{-8\alpha}\mu_j^2(k-l)\Big)^{1/2}\Big|$$
$$\leq c_4\Big(m_j a_j^{-8} + a_j^{-4}\Big)^{1/2}. \quad (5.26)$$

By the triangle inequality it holds

$$\Big|\Big(\sum_{k=1}^{m_j}\sum_{l=1}^{m_j}\mathcal{I}^2_{\gamma_j(k-l)/a_j}(a_j^{-1})\Big)^{1/2} - \Big(\sum_{k=1}^{m_j}\sum_{l=1}^{m_j} s_0^{-8\alpha}\mu_j^2(k-l)\Big)^{1/2}\Big|$$
$$\leq \Big(\sum_{k=1}^{m_j}\sum_{l=1}^{m_j}\big|\mathcal{I}_{\gamma_j(k-l)/a_j}(a_j^{-1}) - s_0^{-4\alpha}\mu_j(k-l)\big|^2\Big)^{1/2}. \quad (5.27)$$

Next, observe that it follows from (5.8), (5.25) and the inequalities $0<\alpha<1/2$ and $s_0>1$, that for all j large enough and for all $(k,l)\in\mathbb{Z}^2$ it holds

$$\Big|\mathcal{I}_{\gamma_j(k-l)/a_j}(a_j^{-1}) - \frac{\mu_j(k-l)}{s_0^{4\alpha}}\Big|^2 \leq \Bigg(\Big|\mathcal{I}_{\gamma_j(k-l)/a_j}(a_j^{-1}) - \frac{\int_{\mathbb{R}} e^{i\gamma_j(k-l)\eta/a_j}|\widehat{\psi}(\eta)|^2\,d\eta}{s_0^{4\alpha}}$$
$$-\frac{2\alpha}{s_0^{4\alpha-2}}\int_{\mathbb{R}} e^{i\gamma_j(k-l)\eta/a_j}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\cdot a_j^{-2}\Big| + \Big|\int_{\mathbb{R}} e^{i\gamma_j(k-l)\eta/a_j}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\Big| a_j^{-2}\Bigg)^2$$
$$\leq 2\Big|\mathcal{I}_{\gamma_j(k-l)/a_j}(a_j^{-1}) - s_0^{-4\alpha}\int_{\mathbb{R}} e^{i\gamma_j(k-l)\eta/a_j}|\widehat{\psi}(\eta)|^2\,d\eta - 2\alpha s_0^{-4\alpha-2}a_j^{-2}$$
$$\times\int_{\mathbb{R}} e^{i\gamma_j(k-l)\eta/a_j}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\Big|^2 + 2\Big|\int_{\mathbb{R}} e^{i\gamma_j(k-l)\eta/a_j}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\Big|^2 a_j^{-4}$$
$$\leq 2c_1^2 a_j^{-8} + 2\Big|\int_{\mathbb{R}} e^{i\gamma_j(k-l)\eta/a_j}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\Big|^2 a_j^{-4}, \quad (5.28)$$

where c_1 is the constant from (5.8).

By (5.27) and (5.28) to derive (5.26) it is sufficient to show that

$$\sum_{k\in\mathbb{Z}}\Big|\int_{\mathbb{R}} e^{i\gamma_j k\eta/a_j}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\Big|^2 = \sum_{k\in\mathbb{Z}}\Big|\int_{-\pi a_j/\gamma_j}^{\pi a_j/\gamma_j} e^{i\gamma_j k\eta/a_j}\sum_{n\in\mathbb{Z}}(\eta+2n\pi a_j/\gamma_j)^2$$

$$\times \left|\widehat{\psi}(\eta + 2n\pi a_j/\gamma_j)\right|^2 d\eta\Big|^2 < +\infty.$$

This inequality holds by Plancherel's identity as $\left\{ \int_{\mathbb{R}} e^{i\gamma_j k\eta/a_j}\eta^2|\widehat{\psi}(\eta)|^2 \, d\eta \right\}_{k\in\mathbb{Z}}$ is the sequence of the Fourier coefficients of the bounded on $[-\pi a_j/\gamma_j, \pi a_j/\gamma_j]$ function $\sum_{n\in\mathbb{Z}}(\eta + 2n\pi a_j/\gamma_j)^2\left|\widehat{\psi}(\eta + 2n\pi a_j/\gamma_j)\right|^2$.

Next, let us define $F_0(\cdot)$ as

$$F_0(\eta) := \sum_{n\in\mathbb{Z}} \left|\widehat{\psi}(\eta + 2nc\pi)\right|^2, \quad \eta \in [-c\pi, c\pi], \tag{5.29}$$

where c is the same positive constant as in Assumption 3'. $F_0(\cdot)$ is a bounded function on $[-c\pi, c\pi]$.

Let us now show that

$$\lim_{j\to+\infty} \frac{1}{m_j} \sum_{k=1}^{m_j}\sum_{l=1}^{m_j} \mu_j^2(k-l) = 2c\pi \int_{-c\pi}^{c\pi} |F_0(\eta)|^2 \, d\eta. \tag{5.30}$$

Note that

$$\frac{1}{m_j} \sum_{k=1}^{m_j}\sum_{l=1}^{m_j} \mu_j^2(k-l) = \frac{1}{m_j} \sum_{k=1}^{m_j}\sum_{q=k-m_j}^{k-1} \mu_j^2(q)$$

and for the sequence $\{\mu_0(k)\}_{k\in\mathbb{Z}}$ of the Fourier coefficients of F_0 it holds

$$\frac{1}{m_j}\left| \sum_{k=1}^{m_j}\sum_{q=k-m_j}^{k-1} \mu_j^2(q) - \sum_{k=1}^{m_j}\sum_{q=k-m_j}^{k-1} \mu_0^2(q) \right| \le \frac{C}{m_j} \sum_{k=1}^{m_j}\sum_{q=k-m_j}^{k-1} |\mu_j(q) - \mu_0(q)| \tag{5.31}$$

as $\mu_j(q)$ and $\mu_0(q)$ are bounded by $\int_{\mathbb{R}} \left|\widehat{\psi}(\eta)\right|^2 d\eta$.

Using the expressions for Fourier coefficients and Assumption 5.2, we get that for $k = 1, ..., m_j$

$$\sum_{q=k-m_j}^{k-1} |\mu_j(q) - \mu_0(q)| \le \sum_{q=k-m_j}^{k-1} \int_{-A}^{A} \left| e^{i\frac{\gamma_j q\eta}{a_j}} - e^{i\frac{q\eta}{c}} \right| \left|\widehat{\psi}(\eta)\right|^2 d\eta$$

$$\le C' \sum_{q=k-m_j}^{k-1} \int_{-A}^{A} \left| \sin\left(\frac{q\eta}{2}\left(\frac{\gamma_j}{a_j} - \frac{1}{c}\right)\right)\right| d\eta$$

$$\le C' \sum_{q=-m_j}^{m_j} \int_{-A}^{A} \left| \sin\left(\frac{q\eta}{2}\left(\frac{\gamma_j}{a_j} - \frac{1}{c}\right)\right)\right| d\eta.$$

Hence, it follows from the inequality $|\sin(x)| \le |x|$ and Assumption 3' that

$$\sum_{q=k-m_j}^{k-1} |\mu_j(q) - \mu_0(q)| \le C'' m_j^2 \left| \frac{\gamma_j}{a_j} - \frac{1}{c} \right| \to 0, \; j \to +\infty. \tag{5.32}$$

Thus, by (5.31), (5.32) and the Cesàro mean convergence theorem one gets

$$\frac{1}{m_j}\left|\sum_{k=1}^{m_j}\sum_{q=k-m_j}^{k-1}\mu_j^2(q)-\sum_{k=1}^{m_j}\sum_{q=k-m_j}^{k-1}\mu_0^2(q)\right|\to 0,\ j\to 0. \tag{5.33}$$

Now, by Plancherel's identity

$$\frac{1}{m_j}\sum_{k=1}^{m_j}\sum_{q=k-m_j}^{k-1}\mu_0^2(q)=\sum_{q=-\infty}^{+\infty}\mu_0^2(q)-\frac{1}{m_j}\sum_{k=1}^{m_j}\sum_{q=k}^{+\infty}\mu_0^2(q)-\frac{1}{m_j}\sum_{k=1}^{m_j}\sum_{q=-\infty}^{k-m_j-1}\mu_0^2(q)$$

$$=2c\pi\int_{-c\pi}^{c\pi}|F_0(\eta)|^2\,d\eta-\frac{1}{m_j}\sum_{k=1}^{m_j}\sum_{q=k}^{+\infty}\mu_0^2(q)-\frac{1}{m_j}\sum_{k'=1}^{m_j}\sum_{q=-\infty}^{-k'}\mu_0^2(q). \tag{5.34}$$

Next, observe that the sequence $\left\{\sum_{q=k}^{+\infty}\mu_0^2(q)\right\}_{k\in\mathbb{N}}$ converges to zero. Consequently by the Cesàro mean convergence theorem one gets

$$\lim_{j\to+\infty}\frac{1}{m_j}\sum_{k=1}^{m_j}\sum_{q=k}^{+\infty}\mu_0^2(q)=0. \tag{5.35}$$

Using the same arguments, one obtains that

$$\lim_{j\to+\infty}\frac{1}{m_j}\sum_{k'=1}^{m_j}\sum_{q=-\infty}^{-k'}\mu_0^2(q)=0. \tag{5.36}$$

Putting together (5.33), (5.34), (5.35) and (5.36) it follows that (5.30) holds true.

Finally, combining (5.30) with (5.23), (5.26) and (5.29) one obtains (5.24). □

5.4 Asymptotic normality of two auxiliary statistics

This section proves asymptotic normality of two auxiliary statistics of the semiparametric model defined by Assumption 5.1. They are two functions of the parameters s_0 and α. The results will be used in the following sections to derive and investigate simultaneous estimators of s_0 and α.

Let us set

$$\overline{\delta}_j^{(2,m_j)}:=\frac{\delta_j^{(2,m_j)}}{m_j}=\frac{1}{m_j}\sum_{k=1}^{m_j}\delta_{jk}^2,\quad j\in\mathbb{N}, \tag{5.37}$$

where δ_{jk} is given in Definition 5.2.

The following theorem introduces the first statistics and derives its asymptotic normality.

Theorem 5.2. *Let the array $\{b_{jk}\}_{(j,k)\in\mathbb{N}\times\mathbb{Z}}$ satisfy Assumption 3' and*

$$\lim_{j\to+\infty} m_j a_j^{-4} = 0. \tag{5.38}$$

Then, when j goes to $+\infty$, the random variables

$$\overline{Y}_j := \sqrt{m_j}\left(\overline{\delta}_j^{(2,m_j)} - s_0^{-4\alpha}\int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2\, d\eta\right) \tag{5.39}$$

converge in distribution to a centred Gaussian random variable $\overline{Y}$ with the variance $\operatorname{Var}(\overline{Y}) = \mathcal{V}_1$ *given by* (5.24).

Remark 5.2. If the array $\{b_{jk}\}_{(j,k)\in\mathbb{N}\times\mathbb{Z}}$ satisfies Assumption 3', then the condition (5.5) of Theorem 5.1 holds true for any $\{m_j\}_{j\in\mathbb{N}}$.

Proof of Theorem 5.2. By Theorem 5.1, when j goes to $+\infty$, the random variables $\sqrt{\mathcal{V}_1}\,Y_j$ converge in distribution to a centred Gaussian random variable $\overline{Y}$ whose variance equals $\mathcal{V}_1$. Moreover, by (5.6) and (5.37) the random variable $\sqrt{\mathcal{V}_1}\,Y_j$ equals

$$\sqrt{\mathcal{V}_1}\,Y_j = \sqrt{\mathcal{V}_1 \times \frac{m_j}{\operatorname{Var}\left(\delta_j^{(2,m_j)}\right)}}\,\sqrt{m_j}\left(\overline{\delta}_j^{(2,m_j)} - \mathbb{E}\left(\overline{\delta}_j^{(2,m_j)}\right)\right),$$

and, by Lemma 5.4, it holds

$$\lim_{j\to+\infty}\sqrt{\mathcal{V}_1 \times \frac{m_j}{\operatorname{Var}\left(\delta_j^{(2,m_j)}\right)}} = 1.$$

Thus, when j goes to $+\infty$, the random variables $\sqrt{m_j}\left(\overline{\delta}_j^{(2,m_j)} - \mathbb{E}\left(\overline{\delta}_j^{(2,m_j)}\right)\right)$ converge in distribution to $\overline{Y}$. To show that the sequence $\{\overline{Y}_j\}_{j\in\mathbb{N}}$ shares the same property, it is enough to prove that

$$\lim_{j\to+\infty}\sqrt{m_j}\left(\mathbb{E}\left(\overline{\delta}_j^{(2,m_j)}\right) - s_0^{-4\alpha}\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\, d\eta\right) = 0. \tag{5.40}$$

It follows from from (5.2), (5.7) and (5.37) that $\mathbb{E}\left(\overline{\delta}_j^{(2,m_j)}\right) = \mathcal{I}_0(a_j^{-1})$. Thus, using Lemma 5.1 and (5.38) one obtains (5.40). $\square$

Let $\{M_j\}_{j\in\mathbb{N}}$ be a sequence of positive integers defined as

$$M_j := \left[\frac{m_j}{(a_{j+1}^{-2} - a_{j+2}^{-2})^2}\right], \tag{5.41}$$

where $[\cdot]$ denotes the integer part function.

Remark 5.3. By (5.41) the sequence $\{M_j\}_{j\in\mathbb{N}}$ satisfies

$$M_j = \left[\frac{m_j(a_{j+1}a_{j+2})^4}{(a_{j+2}^2 - a_{j+1}^2)^2}\right] \geq \left[m_j a_{j+1}^4\right] \to +\infty\,.$$

Assumption 3*. *For all $j \in \mathbb{N}$ and for every $(k,l) \in \mathbb{Z}^2$ it holds*

$$b_{jk} - b_{jl} = \gamma_j(k-l),$$

where $\{\gamma_j\}_{j\in\mathbb{N}}$ is a sequence of positive real numbers such that

$$\lim_{j\to+\infty} \frac{a_j}{\gamma_j} = c \in (0,+\infty) \quad \textit{and} \quad \lim_{j\to+\infty} m_j^2 a_j^8 \left(\frac{\gamma_j}{a_j} - \frac{1}{c}\right) = 0.$$

Remark 5.4. For example, Assumption 3* is satisfied if for all $j \geq j_0 \in \mathbb{N}$ it holds $\gamma_j = a_j$.

Now we introduce the second auxiliary statistics

$$\Delta\overline{\delta}_{j+1}^{(2,M_j)} := \frac{\overline{\delta}_{j+1}^{(2,M_j)} - \overline{\delta}_{j+2}^{(2,M_j)}}{a_{j+1}^{-2} - a_{j+2}^{-2}}$$

via increments of $\overline{\delta}_j^{(2,M_j)}$ and prove its asymptotic normality.

Theorem 5.3. *Assume that the following conditions hold:*

1. *There exists $B \in (0,A)$ such that $\widehat{\psi}$ vanishes on the interval $[-B,B]$, that is*

$$\operatorname{supp}\widehat{\psi} \subseteq \{\xi \in \mathbb{R} : B \leq |\xi| \leq A\}. \tag{5.42}$$

2. *Assumption 3* holds true and for some $j_0 \in \mathbb{N}$ the sequence $\{a_j\}_{j\in\mathbb{N}}$ satisfies*

$$\frac{a_{j+1}}{a_j} \geq \frac{A}{B} > 1\,, \quad \textit{for all} \quad j \geq j_0. \tag{5.43}$$

3. *The sequence $\{m_j\}_{j\in\mathbb{N}}$ satisfies (5.38).*

Then, when j goes to $+\infty$, the random variables

$$\overline{Z}_j := \sqrt{m_j}\left(\Delta\overline{\delta}_{j+1}^{(2,M_j)} - 2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}} \eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right) \tag{5.44}$$

converge in distribution to a centred Gaussian random variable $\overline{Z}$ with the variance $\operatorname{Var}(\overline{Z}) = 2\mathcal{V}_1$.

Remark 5.5. Notice that (5.42) and (5.43) imply that $\operatorname{supp}\widehat{\psi}(a_j\cdot)\bigcap\operatorname{supp}\widehat{\psi}(a_{j+1}\cdot)$ is a Lebesgue negligible set for all sufficiently large $j\in\mathbb{N}$.

Proof of Theorem 5.3. First notice that it follows from (5.1) and Remark 5.5 that $\operatorname{Cov}(\delta_{(j+1)k},\delta_{(j+2)l})=0$ for all $(k,l)\in\{1,\dots,M_j\}^2$ and sufficiently large $j\in\mathbb{N}$, which means that the centred Gaussian vectors $\vec{\delta}_{j+1}^{(M_j)}:=(\delta_{(j+1)1},\dots,\delta_{(j+1)M_j})$ and $\vec{\delta}_{j+2}^{(M_j)}:=(\delta_{(j+2)1},\dots,\delta_{(j+2)M_j})$ are independent. Therefore, the two random variables

$$\delta_{j+1}^{(2,M_j)}:=\sum_{k=1}^{M_j}\delta_{(j+1)k}^2 \quad\text{and}\quad \delta_{j+2}^{(2,M_j)}:=\sum_{k=1}^{M_j}\delta_{(j+2)k}^2$$

are independent.

By Remark 5.3 the sequence $\{M_j\}_{j\in\mathbb{N}}$ approaches $+\infty$ when j increases. Hence, by Assumption 3* condition (5.5) is satisfied if m_j is replaced by M_{j-1} or by M_{j-2}. Therefore, by Theorem 5.1, when j goes to $+\infty$, the random variables

$$Z_{1,j}:=\frac{\delta_{j+1}^{(2,M_j)}-\mathbb{E}(\delta_{j+1}^{(2,M_j)})}{\sqrt{\operatorname{Var}\left(\delta_{j+1}^{(2,M_j)}\right)}}$$

converge in distribution to a standard Gaussian random variable, and that the random variables

$$Z_{2,j}:=\frac{\delta_{j+2}^{(2,M_j)}-\mathbb{E}(\delta_{j+2}^{(2,M_j)})}{\sqrt{\operatorname{Var}\left(\delta_{j+2}^{(2,M_j)}\right)}}$$

share the same property.

Next, using (5.38), (5.41) and (5.43), one gets that

$$\lim_{j\to+\infty}\frac{M_j}{a_{j+1}^8}=\lim_{j\to+\infty}\left(\frac{m_j}{a_{j+1}^4}\cdot\frac{(a_{j+2}/a_{j+1})^4}{\left((a_{j+2}/a_{j+1})^2-1\right)^2}\right)=0$$

as the function $\frac{x^4}{(x^2-1)^2}$ is bounded from above for $x\in[A/B,+\infty)$. The same is also true for M_j/a_{j+2}^8 since $a_{j+2}\ge a_{j+1}$.

Therefore, by Lemma 5.4

$$\lim_{j\to+\infty}\frac{\sqrt{\operatorname{Var}\left(\delta_{j+1}^{(2,M_j)}\right)}}{\sqrt{M_j}}=\sqrt{\mathcal{V}_1}\quad\text{and}\quad\lim_{j\to+\infty}\frac{\sqrt{\operatorname{Var}\left(\delta_{j+2}^{(2,M_j)}\right)}}{\sqrt{M_j}}=\sqrt{\mathcal{V}_1}.$$

Thus, when j goes to $+\infty$, the sequence

$$Z'_{1,j} := \frac{\sqrt{\mathrm{Var}\left(\delta_{j+1}^{(2,M_j)}\right)}}{\sqrt{M_j}}\, Z_{1,j} = \frac{\delta_{j+1}^{(2,M_j)} - \mathbb{E}(\delta_{j+1}^{(2,M_j)})}{\sqrt{M_j}}$$

converges in distribution to a centred Gaussian random variable with variance $\mathcal{V}_1$, and the sequence

$$Z'_{2,j} := \frac{\sqrt{\mathrm{Var}\left(\delta_{j+2}^{(2,M_j)}\right)}}{\sqrt{M_j}}\, Z_{2,j} = \frac{\delta_{j+2}^{(2,M_j)} - \mathbb{E}(\delta_{j+2}^{(2,M_j)})}{\sqrt{M_j}}$$

shares the same property. Therefore, using the fact that for sufficiently large j these two sequences are independent and the equalities $\mathbb{E}(\delta_{j+1}^{(2,M_j)}) = M_j \mathcal{I}_0(a_{j+1}^{-1})$ and $\mathbb{E}(\delta_{j+2}^{(2,M_j)}) = M_j \mathcal{I}_0(a_{j+2}^{-1})$, one gets that the random variables

$$\begin{aligned} Z'_{1,j} - Z'_{2,j} &= \frac{\delta_{j+1}^{(2,M_j)} - \delta_{j+2}^{(2,M_j)}}{\sqrt{M_j}} - \sqrt{M_j}\,(\mathcal{I}_0(a_{j+1}^{-1}) - \mathcal{I}_0(a_{j+2}^{-1})) \\ &= \sqrt{M_j}\Big(\overline{\delta}_{j+1}^{(2,M_j)} - \overline{\delta}_{j+2}^{(2,M_j)} - (\mathcal{I}_0(a_{j+1}^{-1}) - \mathcal{I}_0(a_{j+2}^{-1}))\Big) \end{aligned}$$

converge in distribution to a centred Gaussian random variable with the variance $2\mathcal{V}_1$, when $j \to +\infty$.

By (5.41) the sequence of

$$\begin{aligned} \overline{Z}'_j &:= \frac{\sqrt{m_j}\,(a_{j+1}^{-2} - a_{j+2}^{-2})^{-1}}{\sqrt{M_j}}(Z'_{1,j} - Z'_{2,j}) \\ &= \sqrt{m_j}\left(\frac{\overline{\delta}_{j+1}^{(2,M_j)} - \overline{\delta}_{j+2}^{(2,M_j)}}{a_{j+1}^{-2} - a_{j+2}^{-2}} - \frac{\mathcal{I}_0(a_{j+1}^{-1}) - \mathcal{I}_0(a_{j+2}^{-1})}{a_{j+1}^{-2} - a_{j+2}^{-2}}\right) \end{aligned}$$

shares the same property.

Thus, it turns out that for deriving the theorem it is enough to show that

$$\lim_{j\to+\infty} \sqrt{m_j}\left(\frac{\mathcal{I}_0(a_{j+1}^{-1}) - \mathcal{I}_0(a_{j+2}^{-1})}{a_{j+1}^{-2} - a_{j+2}^{-2}} - 2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2\, d\eta\right) = 0. \tag{5.45}$$

Using Lemma 5.1 one gets that

$$\left|\mathcal{I}_0(a_{j+1}^{-1}) - \mathcal{I}_0(a_{j+2}^{-1}) - \Big(2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2\, d\eta\Big)(a_{j+1}^{-2} - a_{j+2}^{-2})\right|$$

$$\leq \left|\mathcal{I}_0(a_{j+1}^{-1}) - s_0^{-4\alpha}\int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2\, d\eta - \Big(2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2\, d\eta\Big) a_{j+1}^{-2}\right|$$

$$+\left|\mathcal{I}_0(a_{j+2}^{-1}) - s_0^{-4\alpha}\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta - \left(2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right)a_{j+2}^{-2}\right|$$
$$\leq c_1(a_{j+1}^{-4}+a_{j+2}^{-4}),$$

where c_1 is the constant in (5.8). Thus,

$$\sqrt{m_j}\left|\frac{\mathcal{I}_0(a_{j+1}^{-1})-\mathcal{I}_0(a_{j+2}^{-1})}{a_{j+1}^{-2}-a_{j+2}^{-2}} - 2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right| \leq \frac{c_1\sqrt{m_j}(a_{j+1}^{-4}+a_{j+2}^{-4})}{a_{j+1}^{-2}-a_{j+2}^{-2}}. \tag{5.46}$$

Finally, combining (5.38), (5.43) and (5.46) one gets

$$\frac{\sqrt{m_j}(a_{j+1}^{-4}+a_{j+2}^{-4})}{a_{j+1}^{-2}-a_{j+2}^{-2}} = \frac{\sqrt{m_j}}{a_{j+1}^2}\cdot\frac{1+(a_{j+1}/a_{j+2})^4}{1-(a_{j+1}/a_{j+2})^2} \to 0,\ j\to+\infty,$$

which confirms (5.45) and finishes the proof. □

Remark 5.6. For example, the sequence $\{a_j\}_{j\in\mathbb{N}}$ with $a_j = a^j$, $j\in\mathbb{N}$, and $a \geq A/B$ satisfies the assumptions of Theorem 5.3.

Note that under the conditions of Theorem 5.3, for sufficiently large $j\in\mathbb{N}$, the random variable $\overline{Y}_j$ defined in (5.39) is independent of $\overline{Z}_j$ defined by (5.44). It is easy to see as the centred Gaussian random vectors $\vec{\delta}_j^{(m_j)} := (\delta_{j1},\dots,\delta_{jm_j})$, $\vec{\delta}_{j+1}^{(M_j)} := (\delta_{(j+1)1},\dots,\delta_{(j+1)M_j})$ and $\vec{\delta}_{j+2}^{(M_j)} := (\delta_{(j+2)1},\dots,\delta_{(j+2)M_j})$ are independent. Therefore, the following result follows from Theorems 5.2 and 5.3.

Corollary 1. *When j goes to $+\infty$, the random vectors $(\overline{Y}_j,\overline{Z}_j)$ converge in distribution to the random vector $(\overline{Y},\overline{Z})$ with the bivariate centred Gaussian distribution* $\mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\mathcal{V}_1 & 0\\ 0 & 2\mathcal{V}_1\end{pmatrix}\right)$.

5.5 Asymptotic normality of adjusted estimators

In this section the axillary statistics $\overline{\delta}_j^{(2,m_j)}$ and $\Delta\overline{\delta}_{j+1}^{(2,M_j)}$ are used for deriving adjusted statistics to estimate the parameters of interest. The central limit theorem is proved for the proposed adjusted statistics.

By (5.39), (5.44) and Corollary 1, under the assumptions of Theorem 5.3 one has

$$\sqrt{m_j}\begin{pmatrix}\overline{\delta}_j^{(2,m_j)} - s_0^{-4\alpha}\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta \\ \\ \Delta\overline{\delta}_{j+1}^{(2,M_j)} - 2\alpha s_0^{-4\alpha-2}\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0,\begin{pmatrix}\mathcal{V}_1 & 0\\ 0 & 2\mathcal{V}_1\end{pmatrix}\right), \tag{5.47}$$

when $j\to+\infty$.

This two-dimensional central limit theorem gives the fluctuation rate for the corresponding law of large number proven in Alomari et al. (2020)

$$\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2 \, d\eta}, \frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2 \, d\eta} \right) \xrightarrow{a.s.} \Phi(s_0, \alpha) := \left(s_0^{-4\alpha}, \alpha s_0^{-4\alpha-2} \right), \tag{5.48}$$

when $j \to +\infty$.

Let us consider the function $g : [-1, +\infty) \to [-1/e, +\infty)$ defined as $g(t) = te^t$. This is an increasing continuous one-to-one function. Its inverse function is LambertW that is continuous, defined on $[-1/e, +\infty)$ with values in $[-1, +\infty)$ and satisfies

$$\mathrm{LambertW}(y)\, e^{\mathrm{LambertW}(y)} = y \qquad \text{i.e.} \qquad e^{\mathrm{LambertW}(y)} = \frac{y}{\mathrm{LambertW}(y)},$$

with the convention that $0/0 = 1$.

As stated in Alomari et al. (2020), the vector-valued function $\Phi : (1, +\infty) \times (0, 1/2) \to \mathcal{D}$ defined in (5.48) is a continuous one-to-one function taking values in

$$\mathcal{D} = \left\{ (y_1, y_2) \in \mathbb{R}^2 \;:\; 0 < y_1 < 1 \;\text{ and } 0 < y_2 < \frac{y_1^2}{2} \right\}.$$

Its inverse function $\Phi^{-1} : \mathcal{D} \to (1, +\infty) \times (0, 1/2)$ is continuous and given by

$$\begin{aligned} \Phi^{-1}(y_1, y_2) = \Bigg(& \exp\left(\frac{1}{2} \mathrm{LambertW}\left(-\frac{y_1 \ln(y_1)}{2y_2} \right) \right), \\ & \frac{y_2}{y_1} \exp\left(\mathrm{LambertW}\left(-\frac{y_1 \ln(y_1)}{2y_2} \right) \right) \Bigg). \end{aligned}$$

Let us define the following continuous vector-valued truncating function $\mathcal{T}$ defined for $\varepsilon \in (0, 1)$, $(y_1, y_2) \in \mathbb{R}^2$, and taking values in $\mathcal{D}$

$$\mathcal{T}(y_1, y_2, \varepsilon) = \Big(\mathcal{T}_1(y_1, \varepsilon) \;,\; \mathcal{T}_2(y_1, y_2, \varepsilon) \Big) \in \mathcal{D},$$

where

$$\mathcal{T}_1(y_1, \varepsilon) := \max(\varepsilon, \min(y_1, 1 - \varepsilon)) = \begin{cases} \varepsilon, & \text{if } y_1 \le \varepsilon, \\ y_1, & \text{if } \varepsilon \le y_1 \le 1 - \varepsilon, \\ 1 - \varepsilon, & \text{if } y_1 > 1 - \varepsilon, \end{cases}$$

$$\mathcal{T}_2(y_1, y_2, \varepsilon) := \max\left(\varepsilon^2/4, \min\left(y_2, \frac{\left(\mathcal{T}_1(y_1,\varepsilon)\right)^2}{2} - \varepsilon^2/4\right)\right)$$

$$= \begin{cases} \varepsilon^2/4, & \text{if } y_2 \le \varepsilon^2/4, \\ y_2, & \text{if } \varepsilon^2/4 \le y_2 \le \frac{\left(\mathcal{T}_1(y_1,\varepsilon)\right)^2}{2} - \varepsilon^2/4, \\ \frac{\left(\mathcal{T}_1(y_1,\varepsilon)\right)^2}{2} - \varepsilon^2/4, & \text{if } y_2 > \frac{\left(\mathcal{T}_1(y_1,\varepsilon)\right)^2}{2} - \varepsilon^2/4. \end{cases}$$

For values outside the feasible region $\mathcal{D}$, some typical mappings by the truncating function $\mathcal{T}$ are sketched in Figure 5.2.

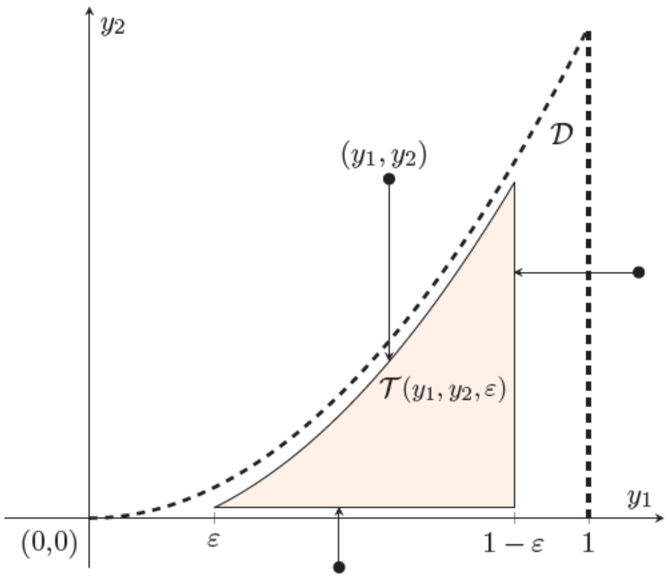


Figure 5.2: Plot of $(y_1,\, y_2)$ and the corresponding truncated values

Note that for each $(y_1, y_2) \in \mathcal{D}$ there is a small enough $\varepsilon > 0$ such that $\mathcal{T}(y_1, y_2, \varepsilon) = (y_1, y_2)$ because $\mathcal{D}$ is an open set. Assumption 5.1 on the parameters ensures that $(s_0, \alpha) \in (1, +\infty) \times (0, 1/2)$ and therefore $\Phi(s_0, \alpha) \in \mathcal{D}$.

Definition 5.3. The adjusted statistic for the parameter (s_0, α) is

$$\widehat{(s_0,\alpha)}_j := \Phi^{-1}\left(\mathcal{T}\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2\, d\eta}\,,\, \frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2\, d\eta}\,,\, \frac{1}{m_j}\right)\right).$$

Note that for some observations the values $\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2\, d\eta}, \frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2\, d\eta}\right)$ may not be in the feasible region $\mathcal{D}$. Therefore, the truncation $\mathcal{T}$ was needed to guarantee that Φ^{-1} acts only on values from $\mathcal{D}$.

Remark 5.7. As for sufficiently large j the vector $\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2\, d\eta}, \frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2\, d\eta}\right)$ falls in $\mathcal{D}$, then $\widehat{(s_0,\alpha)}_j$ and the corresponding adjusted statistic in Alomari et al. (2020) coincide almost surely. At the same time the new statistic requires only the simple truncation $\mathcal{T}$

compared to more complex reflections with respect to the boundary of $\mathcal{D}$ in Alomari et al. (2020). Therefore, for small j the adjusted statistic $\widehat{(s_0,\alpha)}_j$ is computationally simpler than the one in Alomari et al. (2020).

Now we are ready to formulate the main result.

Theorem 5.4. *Under the conditions of Theorem 5.3, the adjusted statistic $\widehat{(s_0,\alpha)}_j$ is a strongly consistent asymptotically normal estimator of the parameter (s_0,α). When j goes to $+\infty$, the random vectors $\sqrt{m_j}\left(\widehat{(s_0,\alpha)}_j-(s_0,\alpha)\right)$ have the asymptotic bivariate centred Gaussian distribution $\mathcal{N}(0,V_{s_0,\alpha})$ with the covariance matrix $V_{s_0,\alpha}$ given by*

$$V_{s_0,\alpha}:=\frac{c\pi s_0^2\int_{-c\pi}^{c\pi}\left|\sum_{n\in\mathbb{Z}}|\widehat{\psi}(\eta+2nc\pi)|^2\right|^2 d\eta}{4\alpha^2(1+2\ln s_0)^2}\begin{pmatrix}(V_{s_0,\alpha})_{11} & (V_{s_0,\alpha})_{12}\\ (V_{s_0,\alpha})_{12} & (V_{s_0,\alpha})_{22}\end{pmatrix}, \tag{5.49}$$

where

$$\begin{aligned}(V_{s_0,\alpha})_{11}&:=\frac{(1-4\alpha\ln s_0)^2}{\left(\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta\right)^2}+\frac{8s_0^4(\ln s_0)^2}{\left(\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right)^2},\\ (V_{s_0,\alpha})_{12}&:=\frac{(1-4\alpha\ln s_0)\alpha(4\alpha+2)s_0^{-1}}{\left(\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta\right)^2}-\frac{8\alpha s_0^3\ln s_0}{\left(\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right)^2},\\ (V_{s_0,\alpha})_{22}&:=\frac{\alpha^2(4\alpha+2)^2s_0^{-2}}{\left(\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta\right)^2}+\frac{8\alpha^2s_0^2}{\left(\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right)^2}.\end{aligned}$$

Proof of Theorem 5.4. The feasible region $\mathcal{D}$ is an open set. Therefore, it follows from (5.48) that, for any $\delta>0$ and for almost all $\omega\in\Omega$, there is $J(\omega,\delta)$ large enough such that for $j\geq J$ the random vector $\left(\frac{\bar{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta},\frac{\Delta\bar{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta}\right)$ belongs to the δ-neighbourhood of $\Phi(s_0,\alpha)$. Notice that $1/m_j\to 0$ when $j\to+\infty$. Hence, for almost all $\omega\in\Omega$ there is $J(\omega)$ large enough such that for $j\geq J$ the image under $\mathcal{T}(\cdot,1/m_j)$ of the vector $\left(\frac{\bar{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta},\frac{\Delta\bar{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta}\right)$ equals to the vector itself.

Thus, for $j\to+\infty$

$$\begin{aligned}\sqrt{m_j}\Bigg|\mathcal{T}\left(\frac{\bar{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta},\frac{\Delta\bar{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta},\frac{1}{m_j}\right)&\\ -\left(\frac{\bar{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta},\frac{\Delta\bar{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta}\right)\Bigg|&\xrightarrow{a.s.}0,\end{aligned} \tag{5.50}$$

where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^2$. Note that (5.50) holds for any norm and any normalising factor, not only $\sqrt{m_j}$, because the difference almost surely vanishes for j larger than some random J.

Hence, by (5.48) and (5.50)

$$\mathcal{T}\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta}\,,\,\frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta}\,,\,\frac{1}{m_j}\right)\xrightarrow{a.s.}\Phi(s_0,\alpha),\ j\to+\infty,$$

which means that the vector $\mathcal{T}\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta}\,,\,\frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta}\,,\,\frac{1}{m_j}\right)$ is a strongly consistent estimator of $\Phi(s_0,\alpha)$.

Moreover, by multivariate Slutsky's lemma (Van der Vaart, 1998, Theorem 2.7(iv)) it follows from (5.50) and the central limit theorem (5.47) that for $j\to+\infty$ it holds

$$\sqrt{m_j}\left(\mathcal{T}\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta}\,,\,\frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta}\,,\,\frac{1}{m_j}\right)-\Phi(s_0,\alpha)\right)\xrightarrow{d}\mathcal{N}(0,V_{\mathcal{V}_1}),\qquad(5.51)$$

where

$$V_{\mathcal{V}_1}:=\mathcal{V}_1\begin{pmatrix}\frac{1}{\left(\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta\right)^2} & 0\\ 0 & \frac{1}{2\left(\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right)^2}\end{pmatrix}.$$

The continuity of Φ^{-1} implies that the estimator $\widehat{(s_0,\alpha)}_j$ is strongly consistent

$$\widehat{(s_0,\alpha)}_j\xrightarrow{a.s.}(s_0,\alpha),\ j\to+\infty.$$

As the central limit theorem in (5.51) can be rewritten as

$$\sqrt{m_j}\left(\Phi(\widehat{(s_0,\alpha)}_j)-\Phi(s_0,\alpha)\right)\xrightarrow{d}\mathcal{N}(0,V_{\mathcal{V}_1}),\ j\to+\infty,$$

then to obtain the asymptotic distribution of the estimator $\widehat{(s_0,\alpha)}_j$ around the parameter of interest (s_0,α) one can use the delta method with the inverse function Φ^{-1}.

To justify it one has to check that Φ^{-1} is differentiable at the point $\Phi(s_0,\alpha)$. By the inverse function theorem, the derivative $D(\Phi^{-1})(\Phi(s_0,\alpha))$ exists if the Jacobian $D\Phi$ of the function $\Phi(\cdot,\cdot)$ at the point (s_0,α) is invertible. In this case it holds $D(\Phi^{-1})(\Phi(s_0,\alpha))=(D\Phi(s_0,\alpha))^{-1}$.

Notice that for any $(s_0,\alpha)\in(1,+\infty)\times(0,1/2)$ it holds

$$D\Phi(s_0,\alpha)=s_0^{-4\alpha-2}\begin{pmatrix}-4\alpha s_0 & -4s_0^2\ln s_0\\ \alpha(-4\alpha-2)s_0^{-1} & 1-4\alpha\ln s_0\end{pmatrix}.\qquad(5.52)$$

Thus, since $s_0 > 1$,

$$\det\left(D\Phi(s_0,\alpha)\right) = s_0^{-8\alpha-4}\left(-4\alpha s_0 - 8\alpha s_0 \ln s_0\right) = -4\alpha s_0^{-8\alpha-3}(1+2\ln s_0) \neq 0$$

and the Jacobian matrix is invertible.

Therefore, by the multivariate delta method (see, for example, (Van der Vaart, 1998, Theorem 3.1))

$$\sqrt{m_j}\left(\widehat{(s_0,\alpha)}_j) - (s_0,\alpha)\right) \xrightarrow{d} \mathcal{N}(0, V_{s_0,\alpha}),\ j \to +\infty,$$

where

$$V_{s_0,\alpha} := \left(D\Phi(s_0,\alpha)\right)^{-1} V_{\mathcal{V}_1}\left(\left(D\Phi(s_0,\alpha)\right)^{-1}\right)^T. \tag{5.53}$$

The covariance matrix given by (5.53) can be explicitly computed. It follows from (5.52) that

$$\left(D\Phi(s_0,\alpha)\right)^{-1} = -\frac{s_0^{4\alpha+1}}{4\alpha(1+2\ln s_0)}\begin{pmatrix} 1-4\alpha\ln s_0 & 4s_0^2\ln s_0 \\ \alpha(4\alpha+2)s_0^{-1} & -4\alpha s_0 \end{pmatrix}.$$

Hence,

$$V_{s_0,\alpha} = \frac{s_0^{8\alpha+2}\mathcal{V}_1}{16\alpha^2(1+2\ln s_0)^2}\begin{pmatrix} 1-4\alpha\ln s_0 & 4s_0^2\ln s_0 \\ \alpha(4\alpha+2)s_0^{-1} & -4\alpha s_0 \end{pmatrix}$$

$$\times \begin{pmatrix} \frac{1}{\left(\int_{\mathbb{R}}|\widehat{\psi}(\eta)|^2\,d\eta\right)^2} & 0 \\ 0 & \frac{1}{2\left(\int_{\mathbb{R}}\eta^2|\widehat{\psi}(\eta)|^2\,d\eta\right)^2} \end{pmatrix}\begin{pmatrix} 1-4\alpha\ln s_0 & \alpha(4\alpha+2)s_0^{-1} \\ 4s_0^2\ln s_0 & -4\alpha s_0 \end{pmatrix}.$$

The straightforward matrix multiplication and application of (5.24) give (5.49), which completes the proof. □

5.6 Numerical examples

This section provides some numerical examples to illustrate and specify the general theoretical results from the previous sections.

The main theoretical results were obtained for general filter transforms and involve some complex functionals of the filters. The following two examples demonstrate that these results can be easily specialized for specific filters/wavelets and are feasibly computable.

Example 5.1. *Let us consider the Shannon father wavelet*

$$\psi_f(t) = \operatorname{sinc}(\pi t) := \begin{cases} \frac{\sin(\pi t)}{\pi t}, & t \neq 0, \\ 1, & t = 0. \end{cases}$$

Its Fourier transform is

$$\widehat{\psi}_f(\eta) = \mathbb{1}_{[-\pi,\pi]}(\eta) := \begin{cases} 1, & \eta \in [-\pi,\pi], \\ 0, & \eta \notin [-\pi,\pi]. \end{cases}$$

It is clear that Assumption 5.2 is satisfied. The corresponding integrals are

$$\int_{\mathbb{R}} \left|\widehat{\psi}_f(\eta)\right|^2 d\eta = 2\pi \qquad \text{and} \qquad \int_{\mathbb{R}} \eta^2 \left|\widehat{\psi}_f(\eta)\right|^2 d\eta = \frac{2}{3}\pi^2.$$

Let $I(c)$ denote the integral

$$I(c) := \int_{-c\pi}^{c\pi} \left| \sum_{n\in Z} \left|\widehat{\psi}_f(\eta + 2nc\pi)\right|^2 \right|^2 d\eta = \int_{-c\pi}^{c\pi} \left| \sum_{n\in Z} \mathbb{1}_{[-\pi,\pi]}(\eta + 2nc\pi) \right|^2 d\eta.$$

Then, for $c \geq 1$ one gets $I(c) = 2\pi$.

If $c < 1$, by solving the inequality $c\pi + 2n^* c\pi \leq \pi$ we obtain $n^* = \left[\frac{1-c}{2c}\right]$. Then, the solution of $\eta^* + 2(n^*+1)c\pi = \pi$ is $\eta^* = \pi\left(1 - 2c\left(1 + \left[\frac{1-c}{2c}\right]\right)\right)$. Therefore, for $\eta^* < 0$ it holds

$$\begin{aligned} I(c) &= \int_{\eta^*}^{-\eta^*} (2n^*+1)^2 d\eta + 2\int_{-c\pi}^{\eta^*} (2n^*+2)^2 d\eta \\ &= -2\eta^*(2n^*+1)^2 + 2(c\pi + \eta^*)(2n^*+2)^2 \end{aligned}$$

and for $\eta^* \geq 0$

$$\begin{aligned} I(c) &= \int_{-\eta^*}^{\eta^*} (2n^*+3)^2 d\eta + 2\int_{-c\pi}^{-\eta^*} (2n^*+2)^2 d\eta \\ &= 2\eta^*(2n^*+3)^2 + 2(c\pi - \eta^*)(2n^*+2)^2. \end{aligned}$$

Thus,

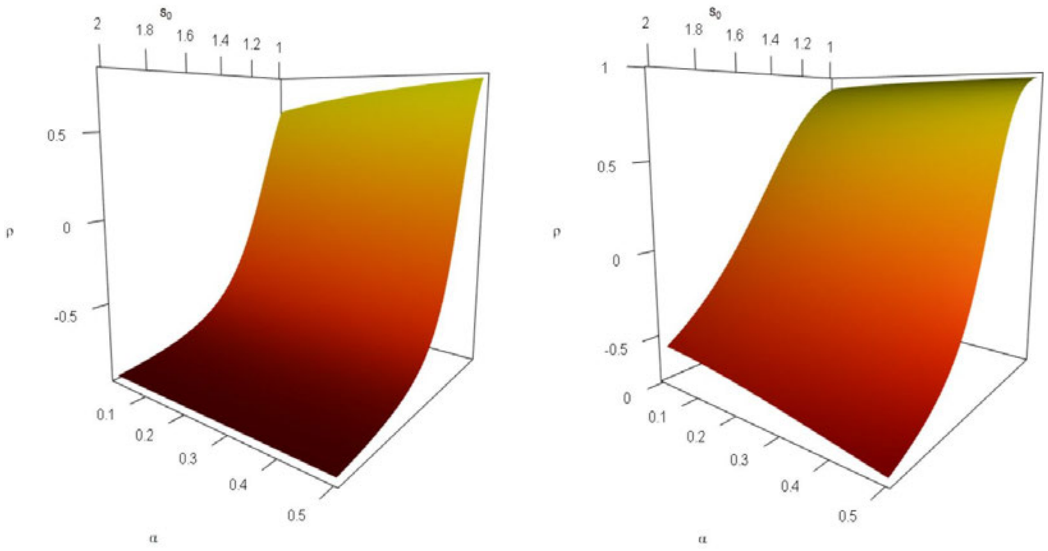
$$I(c) = \begin{cases} 2\pi, & c \geq 1, \\ 2\left|\eta^*\right|(2n^* + 2 + \operatorname{sign}(\eta^*))^2 + 2\left(c\pi - |\eta^*|\right)(2n^*+2)^2, & c < 1. \end{cases}$$

Hence, one can explicitly compute the covariance matrix $V_{s_0,\alpha}$ in Theorem 5.4. For exam-

ple, the correlation of the components of the asymptotic vector equals

$$\rho = \frac{\frac{1}{4\pi^2 s_0}(1-4\alpha\ln s_0)\alpha(4\alpha+2) - \frac{18}{\pi^4}\alpha {s_0}^3 \ln s_0}{\sqrt{\left(\frac{1}{4\pi^2}(1-4\alpha\ln s_0)^2 + \frac{18}{\pi^4}{s_0}^4(\ln s_0)^2\right)\left(\frac{1}{4\pi^2 {s_0}^2}\alpha^2(4\alpha+2)^2 + \frac{18}{\pi^4}\alpha^2 {s_0}^2\right)}}$$

and is plotted in Figure 5.3a as a function of s_0 and α. The plot shows that the components are highly correlated if s_0 is close to 1 and their correlation decreases as s_0 increases.



(a) Shannon father wavelet case

(b) Meyer father wavelet case

Figure 5.3: Asymptotic correlation of $\widehat{s_0}$ and $\widehat{\alpha}$.

Example 5.2. *Let us consider the Meyer father wavelet (Meyer (1992)). It satisfies Assumption 5.2 as its Fourier transform equals*

$$\widehat{\psi}_f(\eta) = \begin{cases} 1, & |\eta| \le \frac{2\pi}{3}, \\ \cos\left(\frac{\pi}{2}\,\nu\left(\frac{3|\eta|}{4\pi} - 1\right)\right), & \frac{2\pi}{3} \le |\eta| \le \frac{4\pi}{3}, \\ 0, & \textit{otherwise}, \end{cases}$$

where the function $\nu(\cdot)$ can be selected as

$$\nu(x) = \begin{cases} 0, & x < 0, \\ x, & x \in [0,1], \\ 1, & x > 1. \end{cases}$$

Its integrals are

$$\int_{\mathbb{R}} \left|\widehat{\psi}_f(\eta)\right|^2 d\eta = 2\pi \qquad \textit{and} \qquad \int_{\mathbb{R}} \eta^2 \left|\widehat{\psi}_f(\eta)\right|^2 d\eta = \frac{8}{9}\pi(\pi^2 - 2). \tag{5.54}$$

For example, for $c > \frac{4}{3}$ *one can easily compute that*

$$I(c) = \int_{-4\pi/3}^{4\pi/3} \left|\widehat{\psi}_f(\eta)\right|^4 d\eta = \frac{11}{6}\pi,$$

which with (5.54) *completely specifies the covariance matrix* $V_{s_0,\alpha}$. *The corresponding correlation is shown in Figure* 5.3b *as a function of* s_0 *and* α.

Comparing it with Figure 5.3a, *one can conclude that filters from Examples* 5.1 *and* 5.2 *produce similar correlation structures of the components of the asymptotic bivariate vector in Theorem* 5.4. *However, for the case of the Meyer father wavelet, the components exhibit higher correlations than for the Shannon one.*

The following example continues simulation studies from Alomari et al. (2020). Simulations in Alomari et al. (2020) demonstrated consistency of the filter-based estimators of the cyclic and long-memory parameters. In Example 5.3, we examine their asymptotic normality.

Note that the results in this chapter were derived for functional time series with continuous time. For computer simulations, one has to use discretized processes on finite grids. In the available literature, it is usually assumed that the corresponding discretization error is negligible with respect to the estimation error. In many cases, it can be rigorously proven, see for example, Alodat and Olenko (2020) and Ayache and Bertrand (2011).

Example 5.3. *In this example the Mexican hat wavelet was used as a filter. This wavelet and its Fourier transform are defined by, see* Liu (2010),

$$\psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{\frac{1}{4}}}\left(1 - \left(\frac{t}{\sigma}\right)^2\right) \mathrm{e}^{-\frac{t^2}{2\sigma^2}} \quad \text{and} \quad \widehat{\psi}(\eta) = \frac{\sqrt{8}\pi^{\frac{1}{4}}\sigma^{\frac{5}{2}}}{\sqrt{3}}\eta^2 \mathrm{e}^{-\frac{\sigma^2\eta^2}{2}}.$$

The value $\sigma = 1$ *was used for computations. The corresponding integrals are*

$$\int_{\mathbb{R}} \left|\widehat{\psi}(\eta)\right|^2 d\eta = 2 \qquad \textit{and} \qquad \int_{\mathbb{R}} \eta^2 \left|\widehat{\psi}(\eta)\right|^2 d\eta = 10.$$

The Fourier transform $\widehat{\psi}(\eta)$ *does not have a finite support, but has light tails that rapidly approaches zero when* $\eta \to +\infty$.

As $X(t), t \in \mathbb{Z}$, we selected the Gegenbauer random process, see Espejo et al. (2015)*. This stochastic process is defined by the following difference equation*

$$\Delta_u^d X(t) = \varepsilon(t), \quad |u| \leq 1, \ 0 < d < 1/2,$$

where $\varepsilon(t)$ is a zero-mean white noise with the common variance $E(\varepsilon^2(t)) = \sigma_\varepsilon^2$.

The fractional difference operator Δ_u^d is given by

$$\Delta_u^d = (1 - 2uB + B^2)^d,$$

where B denotes the time backward-shift operator, i.e. $BX(t) = X(t-1)$.

To simulate realizations of $X(t)$ we used truncated sums of the following infinite moving average representation of the Gegenbauer random process

$$X(t) = \sum_{n=0}^{\infty} C_n^{(d)}(u)\varepsilon(t-n), \quad t \in \mathbb{Z}, \tag{5.55}$$

with the coefficients given by the Gegenbauer polynomial

$$C_n^{(d)}(u) = \sum_{k=0}^{[n/2]} (-1)^k \frac{(2u)^{n-2k}\Gamma(d-k+n)}{k!(n-2k)!\Gamma(d)},$$

where $[n/2]$ is the integer part of $n/2$, and $\Gamma(\cdot)$ is the gamma function.

The chosen for simulations parameters values $d = 0.1$ and $u = 0.3$ correspond to s_0 and α inside of the admissible region $\mathcal{D}$. The realizations of $X(t)$ were approximated by truncated sums with 100 terms in (5.55)*. To compute the statistics $\overline{\delta}_j^{(2,m_j)}$ and $\Delta\overline{\delta}_{j+1}^{(2,M_j)}$ the values $a_j = j$, $b_{jk} = k$, $\gamma_j = 1$, and $m_j = a_j^9$, $j = 1, ..., 7$, were used. In* Alomari et al. (2020) *these values were used to illustrate convergence of the estimates to the true values of parameters.*

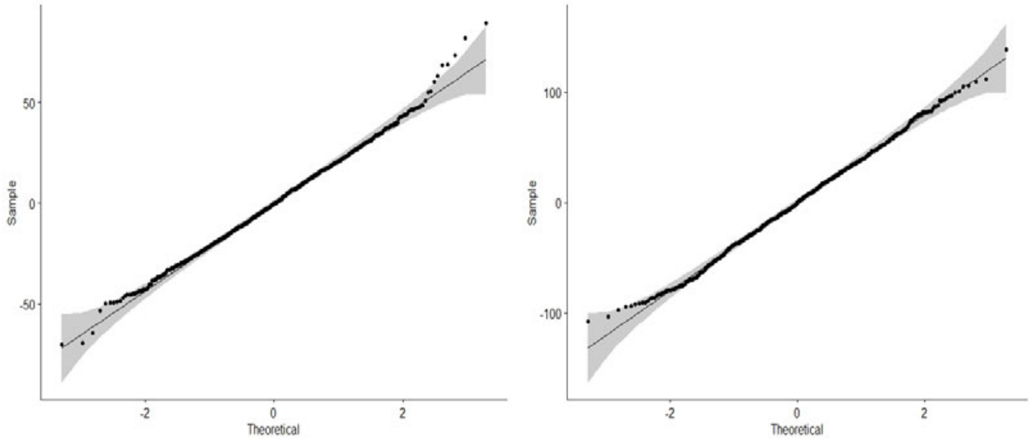
For $j = 7$, the subplots in Figures 5.4a *and* 5.4b *show Q-Q plots of the first two normalised statistics*

$$S_1 := \sqrt{m_j}\left(\frac{\overline{\delta}_j^{(2,m_j)}}{\int_{\mathbb{R}} |\widehat{\psi}(\eta)|^2 \, d\eta} - s_0^{-4\alpha}\right)$$

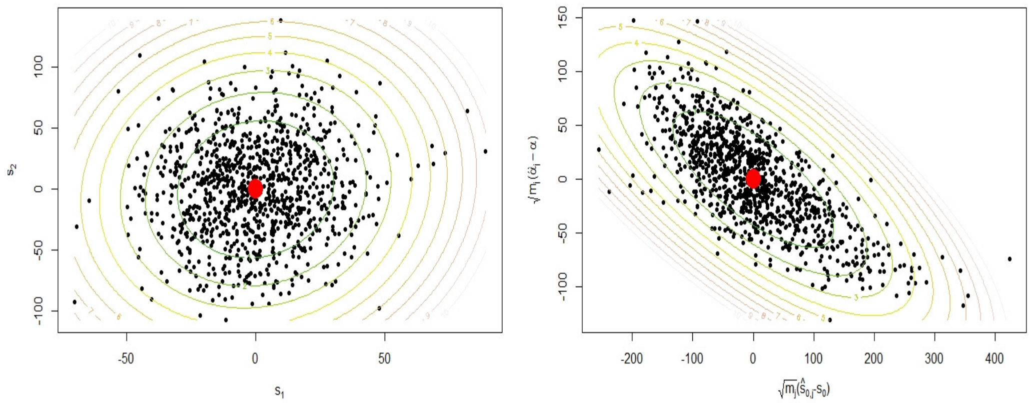
and

$$S_2 := \sqrt{m_j}\left(\frac{\Delta\overline{\delta}_{j+1}^{(2,M_j)}}{2\int_{\mathbb{R}} \eta^2 |\widehat{\psi}(\eta)|^2 \, d\eta} - \alpha s_0^{-4\alpha-2}\right).$$

These plots demonstrate that these statistics have distributions close to Gaussian ones,



(a) Q-Q plot of S_1 (b) Q-Q plot of S_2



(c) Density ellipsoid of (S_1, S_2) (d) Density ellipsoid of $\sqrt{m_j}\left(\widehat{(s_0, \alpha)}_j - (s_0, \alpha)\right)$

Figure 5.4: Realizations of normalised statistics

which is also confirmed by the Shapiro-Wilk test for normality with the corresponding p-values 0.613 and 0.262. Moreover, the estimated correlation matrix $\begin{pmatrix} 1 & 0.084 \\ 0.084 & 1 \end{pmatrix}$ *of these statistics and density ellipsoids in Figure 5.4c underpin the result in (5.47) about asymptotically bivariate normal distribution with uncorrelated components. Finally, Figure 5.4d gives density ellipsoids and realizations of the random vector* $\sqrt{m_j}\left(\widehat{(s_0, \alpha)}_j - (s_0, \alpha)\right)$ *which suggest an asymptotically bivariate normal distribution as in Theorem 5.4.*

The simulation studies suggest that the theoretical results are likely valid for wider classes of filters with light tails. They also demonstrate that the estimators exhibit approximately normal behaviour even for relatively small values of j. A separate publication will be devoted to comprehensive numerical studies.

5.7 Conclusion

The chapter developed statistical inference of semiparametric models of functional time series. It was proved that the generalized filtered method-of-moment estimators of cyclic long-memory models are consistent and asymptotically normal. New adjusted simultaneous statistics were suggested and investigated. A rather general semiparametric class of models satisfies the assumptions of the theorems. In particular, Gegenbauer-type processes belong to this class.

Some interesting areas for future investigations are:

– Applying the approach to the case of multiple singularities, see Arteche (2020) and Klykavka et al. (2012);

– Adapting the methodology to models with other types of spectral singularities;

– Investigating discretization errors for the case when $X(t)$ is observed on a finite grid, see Ayache and Bertrand (2011) and Bardet and Bertrand (2010);

– Investigating the case of random fields, i.e. when the index set of $X(t)$ is multidimensional, see Ayache (2018), Espejo et al. (2015) and Klykavka et al. (2012);

– Continuing simulation studies to empirically compare the proposed approach with least squares and likelihood-type methods, see Beaumont and Smallwood (2019), Ferrara and Guégan (2001) and Whitcher (2004).

Chapter 6

Conclusion

This chapter outlines the main contributions made by this thesis to the theory and statistics of stochastic processes and spherical random fields. It also proposes some future research problems that need consideration and are yet to be solved.

6.1 Main contribution

This thesis studied the stochastic modelling and statistical analysis of spatial and long-range dependent data, modelled by spherical random fields and functional time series.

The first and the second directions of research were carried out with the main motivation of checking the non-Gaussianity and other anomalies in the spatial data. The implemented methodologies were applied to the cosmic microwave background radiation (CMB) data from the Planck mission. First, the multifractal behaviour of spherical random fields was studied. In this study, the existing multifractal models in the literature were investigated and new theoretical models and formulas of multifractal spectra were developed. Next, CMB data from different sky windows located at various places and of different sizes were used to fit the multifractal models. Then, the thesis studied the multifractionality of spherical random fields with cosmological applications and examined probable CMB anomalies with the proposed multifractional approaches.

The third direction of research studied the asymptotic behaviour of simultaneous estimators of cyclic long-memory processes considering a wide semiparametric class of models. Numerical examinations for Meyer, Shannon and Mexican hat wavelets and substantial simulation studies were carried out to exemplify the theoretical findings.

The obtained results and the developed computing techniques in this thesis are novel for spatial and long-range dependent data. They may be applied to other spherical, geoscience, directional, environmental and medical imaging data as well.

The synopsis of the main contributions and results of this thesis are as follows:

- Developing multifractal models for spherical random fields such that the Rényi functions and the multifractal spectra can be computed explicitly;
- Deriving proofs to verify the convergence of the stochastic measure μ_k in the Rényi function computations;
- Investigating the behaviour of the Rényi functions and multifractal spectra of the considered models and the dependence of the Rényi functions on the scaling parameter of these models;
- Conducting extensive numerical studies of the Rényi function for CMB sky windows of different sizes which are located at different places of the CMB sky sphere;

- Simulating realizations of multifractal random fields and computing the empirical Rényi functions to verify the consistency of the actual CMB data with the considered theoretical models;
- Fitting the existing and the newly developed multifractal models to actual CMB data to check the non-Gaussianity in the CMB data from the Planck mission;
- Confirming that the CMB data from the Planck mission has no significant or very minor multifractality;
- Implementing the developed methodology as R programs to detect multifractality in spherical random fields;
- Developing two approaches to investigate the multifractionality of spherical random fields by using the HEALPix ring and nested ordering geometrical structure;
- Computing the pointwise Hölder exponent values for one- and two-dimensional regions of the CMB data considering the HEALPix ring and nested ordering schemes respectively;
- Employing the developed multifractional approaches for statistical analysis, detecting potential CMB anomalies and comparing the obtained results with the other methods available in the literature;
- Revealing that the CMB data from the Planck mission has some multifractionality;
- Implementing computing techniques via R codes to distinguish the multifractional presence in spherical random fields;
- Advancing inferential statistics for semiparametric models of functional time series;
- Proving the asymptotic normality and consistency of the generalized filtered method-of-moment estimators for cyclic long-memory models;
- Proposing novel adjusted simultaneous estimators for cyclic long-memory processes and investigating their properties;
- Conducting extensive numerical and simulation studies to verify the theoretical findings of the developed statistical inference procedure for semiparametric models.

6.2 Directions for future research

In this thesis, multifractal and multifractional behaviours of spherical random fields were investigated with specific applications to CMB data from the Planck mission. In future, it would be interesting to apply and extend the developed computational methodologies to new high resolution CMB data which will result from next generation CMB missions such as CMB-S4 (Abazajian et al. (2019)), Euclid (European Space Agency (2021a) and Racca et al. (2016)) and LiteBIRD (Matsumura et al. (2014)) to further investigate non-Gaussianity and anomalies in the CMB data. Also, it would be important to discover other approaches to detect the multifractality and multifractionality of spherical random fields and to compare the efficiency of various techniques. For example, one can consider the partition-function-based approach to develop computing techniques to examine multifractal presence.

Chapters 3 and 4 show that most of the new models and results were derived considering scalar random fields. In this context, only CMB temperature measurements were considered, but CMB also has polarization data. Therefore, real CMB data are vector random fields on the sphere. Thus, in future, it would be interesting to investigate other models based on vector random fields.

Further, in Chapter 3, the formulas of Rényi functions and theory behind them were developed considering $q \in [1, 2]$. In future, it would be interesting to prove that the results hold true for $q \notin [1, 2]$, see Denisov and Leonenko (2016).

It is known that in many real applications, spherical random fields are evolving over time. Stochastic partial differential equations (SPDEs) can be used to model such temporal changes, see Anh et al. (2018), Broadbridge et al. (2019), Broadbridge et al. (2020) and Restrepo et al. (2021) for more details. Therefore, in future, it would be interesting to study the variations of Rényi functions and Hölder exponents for random fields driven by the SPDEs on the sphere.

In Chapter 3, we examined the Rényi function to detect multifractality of spherical random fields. In future, it would be important to implement statistical tests for various types of Rényi functions. Another noteworthy query is to develop hypothesis tests to check the equality of local Hölder exponent values considering the random field's dependence structure. Deriving the distribution theory for the Hölder exponent estimators is another interesting problem.

It is also important to develop methods for investigating convergence rates of the asymptotics in Chapters 3 and 4 and to compare them with the classical convergence rates in central and non-central limit theorems.

In Chapter 4, detailed discussions were carried out regarding various inconsistent estimators of Hölder exponent values. In future, it would be interesting to examine and compare the reliability of different Hölder exponent estimators with applications in particular for CMB data.

Chapter 4 identified potential CMB anomalies which were consistent with the ones encountered in the TMASK region using the developed multifractional approaches. It would be interesting to further explore the corresponding cosmic objects and properties of such anomalous CMB sky regions.

Chapter 5 of this thesis considered cyclic long-memory processes with spectral singularities at non-zero frequencies. It would be interesting to extend the developed methodology for the case of multiple singularities.

In Chapter 5, novel simultaneous estimators were proposed for cyclic long-memory processes with Gegenbauer-type spectral densities. In future, it would be important to modify the existing methodology to suit models with different types of spectral densities.

Also, it would be interesting to extend the developed methodology for the multidimensional case considering random fields, see Ayache (2018), Espejo et al. (2015) and Klykavka et al. (2012) for more details.

Appendix A

Codes used to produce figures and simulations in Chapter 3

A.1 Maple code used to produce Figure 3.1

The Maple code in this section was used to produce Figure 3.1 in Chapter 3. The code in this section visualizes example Rényi functions and multifractal spectra for Models 1, 2 and 3 as shown in Chapter 3.

```
plot(q*(1 + 1/(4*ln(2))) - q^2*1/(4*ln(2)) - 1, q = 0.1 .. 3,
    axesfont = ["Times New Roman", "ARIAL", 12], labels = ["q", "
    T(q)"], titlefont = ["Times New Roman", 13])

alpha := 1 + 1/(4*ln(2)) - q/(2*ln(2))

falpha := 1 - q^2*1/(4*ln(2))

plot([alpha, falpha, q = 0.1 .. 3], axesfont = ["Times New Roman
    ", "ARIAL", 12], labels = ['alpha', f('alpha')], title = "
    Model 1", titlefont = ["Times New Roman", 13])

plot(q*(1 - log[2](1 - 1/3)) + log[2](1 - q/3) - 1, q = 0.1 ..
    3, axesfont = ["Times New Roman", "ARIAL", 12], labels = ["q"
    , "T(q)"], titlefont = ["Times New Roman", 13])

alpha := 1 - log[2](1 - 1/3) + 1/((q - 3)*ln(2))

falpha := 1 + q/((q - 3)*ln(2)) - log[2](1 - q/3)

plot([alpha, falpha, q = 0.1 .. 2.6], axesfont = ["Times New
    Roman", "ARIAL", 12], labels = ['alpha', f('alpha')], title =
    "Model 2", titlefont = ["Times New Roman", 13])

plot((q*(1 - ln(GAMMA(2)/(2*3*BesselK(2, 2*sqrt(3))))/(2*ln(2)))
    -1/2*log[2](q*BesselK(2, 2*sqrt(3*q)))) + (-1 - log[2](2*3/
   GAMMA(2))/2), q = 0.1..3, axesfont=["Times New Roman","ARIAL"
    ,12],labels=["q","T(q)"], titlefont=["Times New Roman",13])
```

```
alpha := 1 - ln(GAMMA(2)/(2*3*BesselK(2, 2*sqrt(3))))/(2*ln(2))
  - 1/(2*ln(2)*q) + sqrt(3)/2*(BesselK(3, 2*sqrt(3*q)) +
   BesselK(1, 2*sqrt(3*q)))/(2*ln(2)*BesselK(2, 2*sqrt(3*q))*
   sqrt(q))

falpha := 1 + log[2](2*3/GAMMA(2))/2 - 1/(2*ln(2)) + log[2](q*
   BesselK(2, 2*sqrt(3*q)))/2 + sqrt(3*q)/2*(BesselK(3, 2*sqrt(3
   *q)) + BesselK(1, 2*sqrt(3*q)))/(2*ln(2)*BesselK(2, 2*sqrt(3*
   q)))

plot([alpha, falpha, q = 0.1 .. 3], axesfont = ["Times New Roman
   ", "ARIAL", 12], labels = ['alpha', f('alpha')], title = "
   Model 3", titlefont = ["Times New Roman", 13])
```

A.2 Maple code used to produce Figure 3.2

The Maple code in this section was used to produce Figure 3.2 in Chapter 3. The code in this section visualizes example Rényi functions and multifractal spectra for Models 4, 5 and 6 as shown in Chapter 3.

```
plot(q - 1 - 1/2*log[2](2^q*GAMMA(q + 1/2)/sqrt(Pi)), q = 0.1 ..
    3, axesfont = ["Times New Roman", "ARIAL", 12], labels = ["q
   ", "T(q)"], titlefont = ["Times New Roman", 13])

alpha := 1 - 1/2 - Psi(q + 0.5)/(2*ln(2))

falpha := 1 + 1/2*log[2](GAMMA(q + 0.5)/sqrt(Pi)) - q*Psi(q +
   0.5)/(2*ln(2))

plot([alpha, falpha, q = 0.1 .. 3], axesfont = ["Times New Roman
   ", "ARIAL", 12], labels = ['alpha', f('alpha')], title = "
   Model 4", titlefont = ["Times New Roman", 13])

plot(q - 1 - 1/2*log[2](2^(2*q)*GAMMA(2*q + 1/2)/sqrt(Pi)), q =
   0.1 .. 3, axesfont = ["Times New Roman", "ARIAL", 12], labels
    = ["q", "T(q)"], titlefont = ["Times New Roman", 13])

alpha := -Psi(2*q + 0.5)/ln(2)

falpha := 1 + 1/2*log[2](GAMMA(2*q + 0.5)/sqrt(Pi)) - q*Psi(2*q
  + 0.5)/ln(2)

plot([alpha, falpha, q = 0.1 .. 3], axesfont = ["Times New Roman
   ", "ARIAL", 12], labels = ['alpha', f('alpha')], title = "
   Model 5", titlefont = ["Times New Roman", 13])

plot(q*(1 - 1/2*log[2](2/2)) - 1 - 1/2*log[2](2^q*GAMMA(q + 2/2)
  /GAMMA(2/2)), q = 0.1 .. 3, axesfont = ["Times New Roman", "
  ARIAL", 12], labels = ["q", "T(q)"], titlefont = ["Times New
  Roman", 13])

alpha := 1 - 1/2 - Psi(q + 1)/(2*ln(2))
```

```
falpha := 1 + 1/2*log[2](GAMMA(q + 1)) - q*Psi(q + 1)/(2*ln(2))

plot([alpha, falpha, q = 0.1 .. 3], axesfont = ["Times New Roman
    ", "ARIAL", 12], labels = ['alpha', f('alpha')], title = "
    Model 6", titlefont = ["Times New Roman", 13])
```

A.3 Maple code used to produce Figure 3.3

The Maple code in this section was used to produce Figure 3.3 in Chapter 3. The code in this section visualizes dependence of the Rényi functions for Models 1, 2, 3, 4, 5 and 6 on the parameter b as shown in Chapter 3.

```
plot3d(q*(1 + 1/(4*ln(b))) - q^2*1/(4*ln(b)) - 1, q = 0.1 .. 3,
    b = 1.1 .. 10, axesfont = ["Times New Roman", "ARIAL", 12],
    labels = ["q", "b", "T(q)"])

plot3d(q*(1 - log[b](1 - 1/3)) + log[b](1 - q/3) - 1, q = 0.1 ..
     3, b = 1.1 .. 10, axesfont = ["Times New Roman", "ARIAL",
    12], labels = ["q", "b", "T(q)"])

plot3d((q*(1 - ln(GAMMA(2)/(2*3*BesselK(2, 2*sqrt(3)))))/(2*ln(b)
    )) - 1/2*log[b](q*BesselK(2, 2*sqrt(3*q)))) + (-1 - log[b](2*
    3/GAMMA(2))/2), q = 0.1 .. 3, b = 1.1 .. 10, labels = ["q", "
    b", "T(q)"])

plot3d(q - 1 - 1/2*log[b](2^q*GAMMA(q + 1/2)/sqrt(Pi)), q = 0.1
    .. 3, b = 1.1 .. 10, axesfont = ["Times New Roman", "ARIAL",
    12], labels = ["q", "b", "T(q)"])

plot3d(q - 1 - 1/2*log[b](2^(2*q)*GAMMA(2*q + 1/2)/sqrt(Pi)), q
    = 0.1 .. 3, b = 1.1 .. 10, axesfont = ["Times New Roman", "
    ARIAL", 12], labels = ["q", "b", "T(q)"])

plot3d(q(1 - 1/2*log[b](1)) - 1 - 1/2*log[b](2^q*GAMMA(q + 1)/
   GAMMA(1)), q = 0.1 .. 3, b = 1.1 .. 10, axesfont = ["Times
   New Roman", "ARIAL", 12], labels = ["q", "b", "T(q)"])
```

A.4 R code used to produce Figure 3.4

The R code in this section was used to produce Figure 3.4 in Chapter 3. The code in this section visualizes a realization of a multifractal random field (Figure 3.4a), sample Rényi function with the fitted log-normal model (Figure 3.4b) and the plot of $f(\alpha)$ versus α (Figure 3.4c) as shown in Chapter 3.

```
library(sp)
library(RandomFieldsUtils)
library(RandomFields)
library(stats)
```

```
library(rcosmo)
library(colf)

#The function fRen computes the Renyi function for a chosen
   window of sky data(e.g.:- whole sky, large, medium,etc.)
fRen1 <- function(cmbdf, q.min = 1.01, q.max = 10, N = 20, k.box
     = log2(nside(cmbdf)) - 3, intensities = "I")
     {
     if (!is.CMBDataFrame(cmbdf))
     {
         stop("Argument must be a CMBDataFrame")
     }
     ns1 <- nside(cmbdf)
     pixind <- pix(cmbdf)
     nagrpix <- setdiff(1:(12 * ns1^2), pixind)
     field.comp <- rep(0, 12 * ns1^2)
     field.in <- cmbdf[, "I", drop = T]
     field.in <- field.in - minint
     field.final <- replace(field.comp, pixind, field.in)
     res.max <- log2(ns1)
     k.box <- log2(nside(cmbdf)) - 3
     npix <- 12 * 4^k.box
     lev.diff <- 4^(res.max - k.box)
     if (res.max - k.box > 0)
     {
         nagrpix <- unique(ancestor(nagrpix, res.max - k.box))
     }
     agrpix <- setdiff((1:npix), nagrpix)
     mu <- vector(mode = "numeric", length = length(agrpix))
     field.total <- 0
     i <- 1
     for (j in agrpix)
     {
         pixd <- (lev.diff * (j - 1) + 1):(lev.diff * j)
         field.total <- field.total + sum(field.final[pixd])
         mu[i] <- sum(field.final[pixd])
         i <- i + 1
     }
     mu <- mu/field.total
     Q <- seq(q.min, q.max, length.out = N)
     Tq <- vector(mode = "numeric", length = N)
     delta <- (1/length(agrpix))
     ri <- 1
     for (q in Q)
     {
         Tq[ri] <- log2(sum(mu^q))/log2(delta)
         ri <- ri + 1
     }
     Tqf <- data.frame(q = Q, tq = Tq)
     return(Tqf)
}

alp1 <- function(cmbdf, q.min = 1.01, q.max = 10, N = 20, k.box
    = log2(nside(cmbdf)) - 3, intensities = "I")
     {
     if (!is.CMBDataFrame(cmbdf))
     {
         stop("Argument must be a CMBDataFrame")
     }
```

```
    ns1 <- nside(cmbdf)
    pixind <- pix(cmbdf)
    nagrpix <- setdiff(1:(12 * ns1^2), pixind)
    field.comp <- rep(0, 12 * ns1^2)
    field.in <- cmbdf[, "I", drop = T]
    field.in <- field.in - minint
    field.final <- replace(field.comp, pixind, field.in)
    res.max <- log2(ns1)
    k.box <- log2(nside(cmbdf)) - 3
    npix <- 12 * 4^k.box
    lev.diff <- 4^(res.max - k.box)
    if (res.max - k.box > 0)
    {
        nagrpix <- unique(ancestor(nagrpix, res.max - k.box))
    }
    agrpix <- setdiff((1:npix), nagrpix)
    mu <- vector(mode = "numeric", length = length(agrpix))
    field.total <- 0
    i <- 1
    for (j in agrpix)
    {
        pixd <- (lev.diff * (j - 1) + 1):(lev.diff * j)
        field.total <- field.total + sum(field.final[pixd])
        mu[i] <- sum(field.final[pixd])
        i <- i + 1    }
    mu <- mu/field.total
    Q <- seq(q.min, q.max, length.out = N)
    alp <- vector(mode = "numeric", length = N)
    delta <- (1/length(agrpix))
    ri <- 1
    for (q in Q)
    {
        alp[ri] <- sum((mu^q/sum(mu^q)) * log2(mu))/log2(delta)
        ri <- ri + 1
    }
    alp0 <- data.frame(q = Q, tq = alp)
    return(alp0)
}

falp1 <- function(cmbdf, q.min = 1.01, q.max = 10, N = 20, k.box
    = log2(nside(cmbdf)) - 3, intensities = "I")
    {
    if (!is.CMBDataFrame(cmbdf))
    {
        stop("Argument must be a CMBDataFrame")
    }
    ns1 <- nside(cmbdf)
    pixind <- pix(cmbdf)
    nagrpix <- setdiff(1:(12 * ns1^2), pixind)
    field.comp <- rep(0, 12 * ns1^2)
    field.in <- cmbdf[, "I", drop = T]
    field.in <- field.in - minint
    field.final <- replace(field.comp, pixind, field.in)
    res.max <- log2(ns1)
    k.box <- log2(nside(cmbdf)) - 3
    npix <- 12 * 4^k.box
    lev.diff <- 4^(res.max - k.box)

    if (res.max - k.box > 0)
```

```
    {
        nagrpix <- unique(ancestor(nagrpix, res.max - k.box))
    }
    agrpix <- setdiff((1:npix), nagrpix)
    mu <- vector(mode = "numeric", length = length(agrpix))
    field.total <- 0
    i <- 1
    for (j in agrpix)
    {
        pixd <- (lev.diff * (j - 1) + 1):(lev.diff * j)
        field.total <- field.total + sum(field.final[pixd])
        mu[i] <- sum(field.final[pixd])
        i <- i + 1
    }
    mu <- mu/field.total
    Q <- seq(q.min, q.max, length.out = N)
    falp <- vector(mode = "numeric", length = N)
    delta <- (1/length(agrpix))
    ri <- 1
    for (q in Q)
    {
        falp[ri] <- sum((mu^q/sum(mu^q)) * log2(mu^q/sum(mu^q)))
    /log2(delta)
        ri <- ri + 1
    }
    falp0 <- data.frame(q = Q, tq = falp)
    return(falp0)
}

#Here, a realization of a multifractal random field is simulated
    in a large spherical window using a Gaussian mother random
    field $Y(x)$ with the exponential covariance model

cmbdf <- CMBDataFrame(nside = 1024, I = rep(0, 12 * 1024^2))

q.min <- 0.5
q.max <- 3
N <- 40
win <- CMBWindow(theta = c(3 * pi/6, 3 * pi/6, pi/4, pi/4), phi
   = c(0, pi/2, pi/2, 0))
cmbdf2 <- window(cmbdf, new.window = win)
df2 <- coords(cmbdf2, new.coords = "cartesian")

K <- 40
for (i in 0:K)
{
    model1 <- RMexp(var = 2, scale = (3^i))
    f1 <- RFsimulate(x = df2$x, y = df2$y, z = df2$z, model =
   model1, spConform = FALSE)
    cmbdf2$I <- cmbdf2$I + f1
}
save.image(file = "SimMultField1.RData")
load("SimMultField1.RData")
cmbdf2$I <- exp(cmbdf2$I - K - 1)

#Figure 3.4(a)-This figure gives the plot of the simulated
   multifractal random field. The field is rescaled to match
   rcosmo colour pallet
cmbdf1 <- cmbdf2
```

```
cmbdf1$I <- (cmbdf2$I - mean(cmbdf2$I)) * 10^{15.6}
plot(cmbdf1,back.col = "white", ylab = "", xlab = "", zlab = "")

minint <- min(cmbdf2$I)
Tq <- fRen1(cmbdf2, q.min, q.max, N)

#This figure plots the sample Renyi function with the linear
   function for the simulated large window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", main = "Sample
    Renyi function and linear function", pch = 20, col = "blue",
    cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[N, 1], Tq[N, 2], lwd = (2), col
   = "red")

Tq[, 3] <- (Tq[, 1] * (Tq[1, 2] - Tq[N, 2]) + Tq[N, 2] * Tq[1,
   1] - Tq[1, 2] * Tq[N, 1])/(Tq[1, 1] - Tq[N, 1])

#This figure plots the difference between the sample Renyi
   function and the linear function for the simulated large
   window of sky data
plot(Tq[, 1], Tq[, 2] - Tq[, 3], ylab = "difference", xlab = "q"
   , pch = 20, col = "blue")

x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)

QM1 <- colf_nls(y~0 + I(-(x^2) + x), data = Ren1, lower = c(0))
coef(QM1)
Q11 <- (fitted(QM1) + x - 1)

#Figure 3.4(b)-This figure gives the plot of the sample Renyi
   function with the fitted log-normal model for the simulated
   large window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
   = "blue")
lines(Tq[, 1], Q11, col = "red", lwd = 2)

#This figure gives the plot of the difference between the sample
    Renyi function and the fitted log-normal model for the
   simulated large window of sky data
plot(Tq[, 1], Q11 - Tq[, 2], ylab = "difference", xlab = "q",
   pch = 20, col = "blue")
residuals <- Q11 - Tq[, 2]
sqrt(mean(residuals^2))

q.min <- -10
q.max <- 10
#Computes the alpha function for the simulated large window of
   sky data
Alp <- alp1(cmbdf2, q.min, q.max, N)
#This figure gives the plot of the function alpha versus q for
   the simulated large window of sky data
plot(Alp[, 1], Alp[, 2], ylab = expression(paste(alpha(q))),
   xlab = "q", main = expression(paste("Sample ", alpha, "
   versus ", q)), pch = 20, col = "blue", cex.main = 1.25, cex.
   lab = 1.25, cex.axis = 1)
min(Alp[, 2])
```

```
max(Alp[, 2])
max(Alp[, 2]) - min(Alp[, 2])

#Computes the falpha function for the simulated large window of
    sky data
Fq <- falp1(cmbdf2, q.min, q.max, N)
plot(Fq[, 1], Fq[, 2], ylab = expression(paste(f[alpha](q))),
    xlab = "q", main = expression(paste("Sample ", f[alpha], "
    function")), pch = 20, col = "blue", cex.main = 1.25, cex.lab
     = 1.25, cex.axis = 1)

#Figure 3.4(c)-This figure gives the plot of falpha versus alpha
     for the simulated large window of sky data
plot(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))), xlab
     = expression(paste(alpha)), pch = 20, col = "red", cex.main
    = 1.25, cex.lab = 1.25, cex.axis = 1, type = "l")
points(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))),
    xlab = expression(paste(alpha)), pch = 19, col = "blue", cex.
    main = 1.25, cex.lab = 1.25, cex.axis = 1)
```

A.5 R code used to produce Figure 3.5

The R code in this section was used to produce Figure 3.5 in Chapter 3. The code in this section visualizes a large window of surface area 1.231, a medium window of surface area 0.4056, a small window of surface area 0.0596 and a very small window of surface area 0.0017 on the CMB sky sphere as shown in Chapter 3.

```
library(stats)
library(rcosmo)

cmbdf <- CMBDataFrame("CMB_map_smica1024.fits")
#Figure 3.5(a)-This figure gives the plot of a large window on
    the CMB sky sphere
plot(cmbdf,back.col="white",ylab="",xlab="",zlab="")
win <- CMBWindow(theta = c(3 * pi / 6, 3 * pi / 6, pi / 4, pi /
    4), phi = c(0, pi / 2, pi / 2, 0))
plot(win, col = "red", lwd = 3)

#Figure 3.5(b)-This figure gives the plot of a medium window on
    the CMB sky sphere
plot(cmbdf,back.col="white",ylab="",xlab="",zlab="")
win <- CMBWindow(theta = c(pi / 3.5, pi / 3.5, pi / 10, pi / 10)
    , phi = c(0, pi / 2, pi / 2, 0))
plot(win, col = "red", lwd = 3)

#Figure 3.5(c)-This figure gives the plot of a small and a very
    small window on the CMB sky sphere
plot(cmbdf,back.col="white",ylab="",xlab="",zlab="")
win <- CMBWindow(theta = c(pi / 6, pi / 6, pi / 12, pi / 12),
    phi = c(0, pi / 5, pi / 5, 0))
plot(win, col = "red", lwd = 3)
win <- CMBWindow(theta = c(pi / 15, pi / 15, pi / 20, pi / 20),
    phi = c(0, pi / 18, pi / 18, 0))
plot(win, col = "red", lwd = 3)
```

A.6 R code used to produce Figure 3.6

The R code in this section was used to produce Figure 3.6 and to obtain some values of Table 3.1 in Chapter 3.

The following code was used to obtain Figure 3.6 and visualizes the plots obtained with respect to the whole sky data analysis. It gives the plots of sample Rényi function versus linear function (Figure 3.6a), difference of sample Rényi function and linear function (Figure 3.6b), $\alpha(q)$ versus q (Figure 3.6c), $f(\alpha)$ versus α (Figure 3.6d), sample Rényi function with the fitted log-normal model (Figure 3.6e) and difference between sample Rényi function and the fitted log-normal model (Figure 3.6f) as shown in Chapter 3.

```
library(stats)
library(rcosmo)
library(sp)
library(colf)

#For the whole sky
cmbdf <- CMBDataFrame("CMB_map_smica1024.fits")
minint <- min(cmbdf[, "I", drop = T])
q.min <- 1.01
q.max <- 2
N <- 20

# Figure 3.6(a)-This figure gives the plot of the sample Renyi
    function with the linear function
Tq <- fRen1(cmbdf, q.min, q.max, N)
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[20, 1], Tq[20, 2], lwd = (2),
    col = "red")
Tq[, 3] <- (Tq[, 1] * (Tq[1, 2] - Tq[N, 2]) + Tq[N, 2] * Tq[1,
    1] - Tq[1, 2] * Tq[N, 1])/(Tq[1, 1] - Tq[N, 1])

#Figure 3.6(b)-This figure gives the plot of the difference
    between the sample Renyi function and the linear function
plot(Tq[, 1], Tq[, 2] - Tq[, 3], ylab = "Difference", xlab = "q"
    , pch = 20, col = "blue")

x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)

QM1 <- colf_nls(y ~ 0 + I(-(x^2) + x), data = Ren1, lower =c(0))
coef(QM1)
Q11 <- (fitted(QM1) + x - 1)

#Figure 3.6(e)-This figure gives the plot of the sample Renyi
    function with the fitted log-normal model
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue")
lines(Tq[, 1], Q11, col = "red", lwd = 2)
```

```
#Figure 3.6(f)-This figure gives the plot of the difference
    between the sample Renyi function and the fitted log-normal
    model
plot(Tq[, 1], Q11 - Tq[, 2], ylab = "Difference", xlab = "q",
    pch = 20, col = "blue")

#Root Mean Square Error (RMSE)
residuals <- Q11 - Tq[, 2]
sqrt(mean(residuals^2))

q.min <- -10
q.max <- 10

#Computes the alpha function for whole sky data
Alp <- alp1(cmbdf, q.min, q.max, N)

#Figure 3.6(c)-This figure gives the plot of the function alpha
    versus q for whole sky data
plot(Alp[, 1], Alp[, 2], ylab = expression(paste(alpha(q))),
    xlab = "q", pch = 20, col = "red", cex.main = 1.25, cex.lab =
     1.25, cex.axis = 1, type = "l")
points(Alp[, 1], Alp[, 2], ylab = expression(paste(alpha(q))),
    xlab = "q", pch = 19, col = "blue", cex.main = 1.25, cex.lab
    = 1.25, cex.axis = 1)

#alpha interval
min(Alp[, 2])
max(Alp[, 2])
max(Alp[, 2]) - min(Alp[, 2])

# Computes the falpha function for whole sky data
Fq <- falp1(cmbdf, q.min, q.max, N)
plot(Fq[, 1], Fq[, 2], ylab = expression(paste(f[alpha](q))),
    xlab = "q", main = expression(paste("Sample ", f[alpha], "
    function")), pch = 20, col = "blue", cex.main = 1.25, cex.lab
     = 1.25, cex.axis = 1)

#Figure 3.6(d)-This figure gives the plot of the function falpha
     versus alpha for whole sky data
plot(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))), xlab
     = expression(paste(alpha)), pch = 20, col = "red", cex.main
    = 1.25, cex.lab = 1.25, cex.axis = 1, type = "l")
points(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))),
    xlab = expression(paste(alpha)), pch = 19, col = "blue", cex.
    main = 1.25, cex.lab = 1.25, cex.axis = 1)
```

A.7 R code used to produce Figure 3.7

The R code in this section was used to produce Figure 3.7 and to obtain some values of Table 3.1 in Chapter 3.

The following code was used to obtain Figure 3.7 and visualizes the plots obtained with respect to the large and small sky windows data analysis. It gives the plots of $f(\alpha)$

versus α for large window (Figure 3.7a), difference with linear function for large window (Figure 3.7b), difference with Model 1 for large window (Figure 3.7c), $f(\alpha)$ versus α for small window (Figure 3.7d), difference with linear function for small window (Figure 3.7e) and difference with Model 1 for small window (Figure 3.7f) as shown in Chapter 3.

```
#For a large window near the pole of the sphere
library(stats)
library(rcosmo)
library(sp)
library(colf)

cmbdf <- CMBDataFrame("CMB_map_smica1024.fits")
minint <- min(cmbdf[, "I", drop = T])
q.min <- 1.01
q.max <- 2
N <- 20
#Choosing a large window of CMB sky data
win <- CMBWindow(theta = c(3 * pi/6, 3 * pi/6, pi/4, pi/4), phi
   = c(0, pi/2, pi/2, 0))
cmbdf1 <- window(cmbdf, new.window = win)
Tq <- fRen1(cmbdf1, q.min, q.max, N)

#This figure gives the plot of the sample Renyi function with
   the linear function for a large window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
   = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[20, 1], Tq[20, 2], lwd = (2),
   col = "red")
Tq[, 3] <- (Tq[, 1] * (Tq[1, 2] - Tq[N, 2]) + Tq[N, 2] * Tq[1,
   1] - Tq[1, 2] * Tq[N, 1])/(Tq[1, 1] - Tq[N, 1])

#Figure 3.7(b)-This figure gives the plot of the difference
   between the sample Renyi function and the linear function for
    a large window of sky data
plot(Tq[, 1], Tq[, 2] - Tq[, 3], ylab = "Difference", xlab = "q"
   , pch = 20, col = "blue")

x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)
QM1 <- colf_nls(y~0 + I(-(x^2) + x), data = Ren1, lower = c(0))
coef(QM1)
Q11 <- (fitted(QM1) + x - 1)

#This figure gives the plot of the sample Renyi function with
   the fitted log-normal model for a large window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
   = "blue")
lines(Tq[, 1], Q11, col = "red", lwd = 2)

#Figure 3.7(c)-This figure gives the plot of the difference
   between the sample Renyi function and the fitted log-normal
   model for a large window of sky data
plot(Tq[, 1], Q11 - Tq[, 2], ylab = "Difference", xlab = "q",
   pch = 20, col = "blue")
```

```
residuals <- Q11 - Tq[, 2]
sqrt(mean(residuals^2))

q.min <- -10
q.max <- 10
#Computes the alpha function for a large window of sky data
Alp <- alp1(cmbdf1, q.min, q.max, N)
#This figure gives the plot of the function alpha versus q for a
    large window of sky data
plot(Alp[, 1], Alp[, 2], ylab = expression(paste(alpha(q))),
   xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab
   = 1.25, cex.axis = 1)
min(Alp[, 2])
max(Alp[, 2])
max(Alp[, 2]) - min(Alp[, 2])

#Computes the falpha function for a large window of sky data
Fq <- falp1(cmbdf1, q.min, q.max, N)
plot(Fq[, 1], Fq[, 2], ylab = expression(paste(f[alpha](q))),
   xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab
   = 1.25, cex.axis = 1)

#Figure 3.7(a)-This figure gives the plot of the function falpha
    versus alpha for a large window of sky data
plot(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))), xlab
    = expression(paste(alpha)), pch = 20, col = "red", cex.main
   = 1.25, cex.lab = 1.25, cex.axis = 1, type = "l")
points(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))),
   xlab = expression(paste(alpha)), pch = 19, col = "blue", cex.
   main = 1.25, cex.lab = 1.25, cex.axis = 1)

#For a small window near the pole of the sphere
q.min <- 1.01
q.max <- 2
N <- 20

#Choosing a small window of CMB sky data
win <- CMBWindow(theta = c(pi/6, pi/6, pi/12, pi/12), phi = c(0,
    pi/5, pi/5, 0))
cmbdf4 <- window(cmbdf, new.window = win)
Tq <- fRen1(cmbdf4, q.min, q.max, N)

#This figure gives the plot of the sample Renyi function with
   the linear function for a small window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
   = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[20, 1], Tq[20, 2], lwd = (2),
   col = "red")
Tq[, 3] <- (Tq[, 1] * (Tq[1, 2] - Tq[N, 2]) + Tq[N, 2] * Tq[1,
   1] - Tq[1, 2] * Tq[N, 1])/(Tq[1, 1] - Tq[N, 1])

#Figure 3.7(e)-This figure gives the plot of the difference
   between the sample Renyi function and the linear function for
    a small window of sky data
plot(Tq[, 1], Tq[, 2] - Tq[, 3], ylab = "Difference", xlab = "q"
   , pch = 20, col = "blue")
```

```r
x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)

QM1 <- colf_nls(y~0 + I(-(x^2) + x), data = Ren1, lower = c(0))
#Coef(QM1) gives the estimated parameter a
coef(QM1)
Q11 <- (fitted(QM1) + x - 1)

#This figure gives the plot of the sample Renyi function with
   the fitted log-normal model for a small window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
   = "blue")
lines(Tq[, 1], Q11, col = "red", lwd = 2)

#Figure 3.7(f)-This figure gives the plot of the difference
   between the sample Renyi function and the fitted log-normal
   model for a small window of sky data
plot(Tq[, 1], Q11 - Tq[, 2], ylab = "Difference", xlab = "q",
   pch = 20, col = "blue")

residuals <- Q11 - Tq[, 2]
sqrt(mean(residuals^2))

q.min <- -10
q.max <- 10

#Computes the alpha function for a small window of sky data
Alp <- alp1(cmbdf4, q.min, q.max, N)
#This figure gives the plot of the function alpha versus q for a
    small window of sky data
plot(Alp[, 1], Alp[, 2], ylab = expression(paste(alpha(q))),
   xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab
   = 1.25, cex.axis = 1)

min(Alp[, 2])
max(Alp[, 2])
max(Alp[, 2]) - min(Alp[, 2])

#Computes the falpha function for a small window of sky data
Fq <- falp1(cmbdf4, q.min, q.max, N)
plot(Fq[, 1], Fq[, 2], ylab = expression(paste(f[alpha](q))),
   xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab
   = 1.25, cex.axis = 1)

#Figure 3.7(d)-This figure gives the plot of the function falpha
    versus alpha for a small window of sky data
plot(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))), xlab
    = expression(paste(alpha)), pch = 20, col = "red", cex.main
   = 1.25, cex.lab = 1.25, cex.axis = 1, type = "l")
points(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))),
   xlab = expression(paste(alpha)), pch = 19, col = "blue", cex.
   main = 1.25, cex.lab = 1.25, cex.axis = 1)

#For a medium window near the pole of the sphere
q.min <- 1.01
q.max <- 2
N <- 20
```

```r
#Choosing a medium window of CMB sky data
win <- CMBWindow(theta = c(pi/3.5, pi/3.5, pi/10, pi/10), phi = c(0, pi/2, pi/2, 0))
cmbdf2 <- window(cmbdf, new.window = win)
Tq <- fRen1(cmbdf2, q.min, q.max, N)

#This figure gives the plot of the sample Renyi function with the linear function for a medium window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[20, 1], Tq[20, 2], lwd = (2), col = "red")
Tq[, 3] <- (Tq[, 1] * (Tq[1, 2] - Tq[N, 2]) + Tq[N, 2] * Tq[1, 1] - Tq[1, 2] * Tq[N, 1])/(Tq[1, 1] - Tq[N, 1])

#This figure gives the plot of the difference between the sample Renyi function and the linear function for a medium window of sky data
plot(Tq[, 1], Tq[, 2] - Tq[, 3], ylab = "difference", xlab = "q", pch = 20, col = "blue")

x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)

QM1 <- colf_nls(y~0 + I(-(x^2) + x), data = Ren1, lower = c(0))
#Coef(QM1) gives the estimated parameter a
coef(QM1)
Q11 <- (fitted(QM1) + x - 1)

#This figure gives the plot of the sample Renyi function with the fitted log-normal model for a medium window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col = "blue")
lines(Tq[, 1], Q11, col = "red", lwd = 2)

residuals <- Q11 - Tq[, 2]
sqrt(mean(residuals^2))

q.min <- -10
q.max <- 10

#Computes the alpha function for a medium window of sky data
Alp <- alp1(cmbdf2, q.min, q.max, N)
#This figure gives the plot of the function alpha versus q for a medium window of sky data
plot(Alp[, 1], Alp[, 2], ylab = expression(paste(alpha(q))), xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)

min(Alp[, 2])
max(Alp[, 2])
max(Alp[, 2]) - min(Alp[, 2])

#Computes the falpha function for a medium window of sky data
Fq <- falp1(cmbdf2, q.min, q.max, N)
plot(Fq[, 1], Fq[, 2], ylab = expression(paste(f[alpha](q))),
```

```
    xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab
    = 1.25, cex.axis = 1)

#This figure gives the plot of the function falpha versus alpha
    for a medium window of sky data
plot(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))), xlab
     = expression(paste(alpha)), pch = 20, col = "red", cex.main
    = 1.25, cex.lab = 1.25, cex.axis = 1, type = "l")
points(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))),
    xlab = expression(paste(alpha)), pch = 19, col = "blue", cex.
    main = 1.25, cex.lab = 1.25, cex.axis = 1)

#For a very small window near the pole of the sphere
q.min <- 1.01
q.max <- 2
N <- 20

#Choosing a very small window of CMB sky data
win <- CMBWindow(theta = c(pi/15, pi/15, pi/20, pi/20), phi = c
    (0, pi/18, pi/18, 0))
cmbdf9 <- window(cmbdf, new.window = win)
Tq <- fRen1(cmbdf9, q.min, q.max, N)

#This figure gives the plot of the sample Renyi function with
    the linear function for a very small window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[20, 1], Tq[20, 2], lwd = (2),
    col = "red")
Tq[, 3] <- (Tq[, 1] * (Tq[1, 2] - Tq[N, 2]) + Tq[N, 2] * Tq[1,
    1] - Tq[1, 2] * Tq[N, 1])/(Tq[1, 1] - Tq[N, 1])

#This figure gives the plot of the difference between the sample
     Renyi function and the linear function for a very small
    window of sky data
plot(Tq[, 1], Tq[, 2] - Tq[, 3], ylab = "difference", xlab = "q"
    , pch = 20, col = "blue")

x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)

QM1 <- colf_nls(y~0 + I(-(x^2) + x), data = Ren1, lower = c(0))
#Coef(QM1) gives the estimated parameter a
coef(QM1)
Q11 <- (fitted(QM1) + x - 1)

#This figure gives the plot of the sample Renyi function with
    the fitted log-normal model for a very small window of sky
    data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue")
lines(Tq[, 1], Q11, col = "red", lwd = 2)

residuals <- Q11 - Tq[, 2]
sqrt(mean(residuals^2))

q.min <- -10
```

```
q.max <- 10

#Computes the alpha function for a very small window of sky data
Alp <- alp1(cmbdf9, q.min, q.max, N)

#This figure gives the plot of the function alpha versus q for a
     very small window of sky data
plot(Alp[, 1], Alp[, 2], ylab = expression(paste(alpha(q))),
    xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab
    = 1.25, cex.axis = 1)

min(Alp[, 2])
max(Alp[, 2])
max(Alp[, 2]) - min(Alp[, 2])

#Computes the falpha function for a very small window of sky
    data
Fq <- falp1(cmbdf9, q.min, q.max, N)
plot(Fq[, 1], Fq[, 2], ylab = expression(paste(f[alpha](q))),
    xlab = "q", pch = 20, col = "blue", cex.main = 1.25, cex.lab
    = 1.25, cex.axis = 1)

#This figure gives the plot of the function falpha versus alpha
    for a very small window of sky data
plot(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))), xlab
     = expression(paste(alpha)), pch = 20, col = "red", cex.main
    = 1.25, cex.lab = 1.25, cex.axis = 1, type = "l")
points(Alp[, 2], Fq[, 2], ylab = expression(paste(f(alpha))),
    xlab = expression(paste(alpha)), pch = 19, col = "blue", cex.
    main = 1.25, cex.lab = 1.25, cex.axis = 1)
```

A.8 R code used to produce Figure 3.8

The R code in this section was used to produce Figure 3.8 in Chapter 3. The following code visualizes the plots of differences between the sample Rényi functions and the fitted models 2, 3, 4, 5, and 6 correspondingly by Figures 3.8a, 3.8b, 3.8c, 3.8d and 3.8e as shown in Chapter 3.

```
library(stats)
library(rcosmo)
library(sp)
library(minpack.lm)

#Model 2: large window near the pole of the sphere
cmbdf <- CMBDataFrame("CMB_map_smica1024.fits")
minint <- min(cmbdf[, "I", drop = T])
q.min <- 1.01
q.max <- 2
N <- 20
win <- CMBWindow(theta = c(3 * pi/6, 3 * pi/6, pi/4, pi/4), phi
    = c(0, pi/2, pi/2, 0))
cmbdf1 <- window(cmbdf, new.window = win)
```

```
Tq <- fRen1(cmbdf, q.min, q.max, N)
#This figure gives the plot of the sample Renyi function with
    the linear function for a large window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[20, 1], Tq[20, 2], lwd = (2),
    col = "red")

x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)

QM2 <- nlsLM(y ~ (log(1 - (x * B)) - x * log(1 - B))/A, start =
    list(A = 2, B = 0.1), data = Ren1, trace = TRUE, control =
    nls.control(maxiter = 1000, tol = 0.01, printEval = FALSE,
    warnOnly = TRUE))
#Coef(QM2) gives the estimated parameters A and B
coef(QM2)

#Verifying that the estimated parameters satisfy the assumptions
     of Theorem 3.5
#Evaluating b(L.H.S)
exp(coef(QM2)[1])
#Evaluating R.H.S
1 + coef(QM2)[2]^2/(1 - 2 * coef(QM2)[2])

Q21 <- (fitted(QM2) + x - 1)

#This figure gives the plot of the sample Renyi function with
    the fitted Model 2
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue")
lines(Tq[, 1], Q21, col = "red", lwd = 2)

#Figure 3.8(a)-This figure gives the plot of the difference
    between the sample Renyi function and the fitted Model 2
plot(Tq[, 1], Q21 - Tq[, 2], ylab = "Difference", xlab = "q",
    pch = 20, col = "blue")
residuals <- Q21 - Tq[, 2]
sqrt(mean(residuals^2))

#Model 3: large window near the pole of the sphere
QM3 <- nlsLM(y ~ 1/A * (x * log(2 * (C^B) * besselK(2 * C, B,
    expon.scaled = FALSE)/gamma(B)) - B * log(x)/2 - log(besselK
    (2 * C * sqrt(x), B, expon.scaled = FALSE)) - log(2 * C^B/
    gamma(B))), start = list(A = 0.1, B = 2, C = 1.4), data =
    Ren1, trace = TRUE, lower = c(0.01, 0.1, 0.1), upper = c(10,
    10, 10), control = nls.control(maxiter = 1000, tol = 0.01,
    printEval = FALSE))
#Coef(QM3) gives the estimated parameters A, B and C
coef(QM3)

#Verifying that the estimated parameters satisfy the assumptions
     of Theorem 3.6
#Evaluating b(L.H.S)
exp((coef(QM3)[1])/2)
#Evaluating R.H.S
```

```
sqrt((gamma(coef(QM3)[2])*2^((coef(QM3)[2]/2)-1)*besselK(2*sqrt
    (2)*(coef(QM3)[3]), coef(QM3)[2], expon.scaled = FALSE))/(((
    coef(QM3)[3])^coef(QM3)[2])*((besselK((2*(coef(QM3)[3])),
    coef(QM3)[2], expon.scaled = FALSE))^2)))

Q31 <- (fitted(QM3) + x - 1)

#This figure gives the plot of the sample Renyi function with
    the fitted Model 3
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue")
lines(Tq[, 1], Q31, col = "red", lwd = 2)

#Figure 3.8(b)-This figure gives the plot of the difference
    between the sample Renyi function and the fitted Model 3
plot(Tq[, 1], Q31 - Tq[, 2], ylab = "Difference", xlab = "q",
    pch = 20, col = "blue")
residuals <- Q31 - Tq[, 2]
sqrt(mean(residuals^2))

#Model 4: large window near the pole of the sphere
win <- CMBWindow(theta = c(3 * pi/6, 3 * pi/6, pi/4, pi/4), phi
    = c(0, pi/2, pi/2, 0))
cmbdf1 <- window(cmbdf, new.window = win)
Tq <- fRen1(cmbdf1, q.min, q.max, N)

#This figure gives the plot of the sample Renyi function with
    the linear function for a large window of sky data
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue", cex.main = 1.25, cex.lab = 1.25, cex.axis = 1)
segments(Tq[1, 1], Tq[1, 2], Tq[20, 1], Tq[20, 2], lwd = (2),
    col = "red")

x <- Tq[, 1]
b <- rep(1, 20)
y <- (Tq[, 2] - x + b)
Ren1 <- data.frame(x, y)

QM4 <- nls(y ~ A * (x + log2(gamma(x + (1/2))) - 0.5 * log2(pi))
    , start = list(A = 0.2), data = Ren1, control = nls.control(
    maxiter = 1000, tol = 0.01, minFactor = 1/1024, printEval =
    FALSE, warnOnly = TRUE), trace = TRUE)
#Coef(QM4) gives the estimated parameter A
coef(QM4)

Q14 <- (fitted(QM4) + x - 1)

#This figure gives the plot of the sample Renyi function with
    the fitted Model 4
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue")
lines(Tq[, 1], Q14, col = "red", lwd = 2)

#Figure 3.8(c)-This figure gives the plot of the difference
    between the sample Renyi function and the fitted Model 4
plot(Tq[, 1], Q14 - Tq[, 2], ylab = "Difference", xlab = "q",
    pch = 20, col = "blue")
residuals <- Q14 - Tq[, 2]
sqrt(mean(residuals^2))
```

```
#Model 5: large window near the pole of the sphere
QM5 <- nls(y ~ A * (B * x + log2(gamma(B * x + (1/2))) - 0.5 *
    log2(pi)), start = list(A = 0.2, B = 2), data = Ren1, control
     = nls.control(maxiter = 1e+07, tol = 0.01, minFactor = 1/
    1024, printEval = FALSE, warnOnly = TRUE), trace = TRUE)
#Coef(QM5) gives the estimated parameters A and B
coef(QM5)

Q15 <- (fitted(QM5) + x - 1)

#This figure gives the plot of the sample Renyi function with
    the fitted Model 5
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue")
lines(Tq[, 1], Q15, col = "red", lwd = 2)

#Figure 3.8(d)-This figure gives the plot of the difference
    between the sample Renyi function and the fitted Model 5
plot(Tq[, 1], Q15 - Tq[, 2], ylab = "Difference", xlab = "q",
    pch = 20, col = "blue")
residuals <- Q15 - Tq[, 2]
sqrt(mean(residuals^2))

#Model 6: large window near the pole of the sphere
QM6 <- nls(y ~ A * (x * log2(B) - x - log2(gamma(x + B)) + log2(
    gamma(B))), start = list(A = 0.2, B = 3), data = Ren1,
    control = nls.control(maxiter = 1e+07, tol = 0.01, minFactor
    = 1/1024, printEval = FALSE, warnOnly = TRUE), trace = TRUE,
    algorithm = "port", lower = c(0, 1), upper = c(100, 300))
#Coef(QM6) gives the estimated parameters A and B
coef(QM6)

Q16 <- (fitted(QM6) + x - 1)

#This figure gives the plot of the sample Renyi function with
    the fitted Model 6
plot(Tq[, 1], Tq[, 2], ylab = "T(q)", xlab = "q", pch = 20, col
    = "blue")
lines(Tq[, 1], Q16, col = "red", lwd = 2)

#Figure 3.8(e)-This figure gives the plot of the difference
    between the sample Renyi function and the fitted Model 6
plot(Tq[, 1], Q16 - Tq[, 2], ylab = "Difference", xlab = "q",
    pch = 20, col = "blue")
residuals <- Q16 - Tq[, 2]
sqrt(mean(residuals^2))
```

Appendix B

Codes used to produce figures and simulations in Chapter 4

B.1 R code used to produce Figure 4.1

The R code in this section was used to produce Figure 4.1 in Chapter 4. The code in this section visualizes a HEALPix ring ordering scheme (Figure 4.1a) and a HEALPix nested ordering scheme (Figure 4.1b) as shown in Chapter 4.

```
library(rcosmo)
library(rgl)

#Figure 4.1(a)-This figure gives the plot of the HEALPix ring
   ordering visualization
#Generating a CMB data frame with nside=8 and "ring" ordering
cmbdf <- CMBDataFrame(nside = 8, ordering = "ring")
plot(cmbdf, type = "l", col = "black", back.col = "white", xlab
   = "", ylab = "", zlab = "")

#Labeling the HEALPix values 1, 100:107 and 768
tolabel <- c(1, 100:107, 768)
plot(cmbdf[tolabel,], labels = tolabel, col = "red", add = TRUE)

um <- matrix(c(0.8948848, -0.4459228, -0.01829224, 0, 0.1114479,
    0.1835916, 0.97666484, 0, -0.4321588, -0.8760409,
   0.21399030, 0, 0, 0, 0, 1), byrow = TRUE, nrow = 4, ncol = 4)
view3d(userMatrix = um)
rgl.snapshot("Figure41a.png")

#Figure 4.1(b)-This figure gives the plot of the HEALPix nested
   ordering visualization with nside=2
ns <- 256
rand <- rnorm(12 * ns^2)
cmbdf <- CMBDataFrame(nside = 64, I = rnorm(12 * 64 ^ 2),
   ordering = "nested")
w21 <- window(CMBDataFrame(nside = ns, I = rand, ordering = "
   nested"), in.pixels = 1)
```

```
displayPixelBoundaries(nside = 2, ordering = "nested", incl.
    labels = c(1, 2, 3, 4), nums.col = "red", col = "black")
plot(w21, col = "light blue", back.col = "white", add = TRUE,
    size = 1.2)
plot(window(cmbdf, in.pixels = 2), col = "green", add = TRUE)
plot(window(cmbdf, in.pixels = 4), col = "purple", add = TRUE)
plot(window(cmbdf, in.pixels = 5), col = "orange", add = TRUE)
plot(window(cmbdf, in.pixels = 6), col = "red", add = TRUE)
um <- matrix(c(-0.6893843, 0.7242531, -0.01436349, 0,
    -0.5146448, -0.4757210, 0.71332407, 0, 0.5097938, 0.4991464,
    0.70068759, 0, 0, 0, 0, 1), byrow = TRUE, nrow = 4, ncol = 4)
view3d(userMatrix = um)
rgl.snapshot("Figure41b.png")
```

B.2 R code used to produce Figure 4.2 and results in Tables 4.1 and 4.2

The R code in this section was used to produce Figure 4.2 and to obtain results in Tables 4.1 and 4.2 in Chapter 4. The code in this section visualizes examples of scaled intensities and $\hat{H}(t)$ values for one-dimensional CMB regions. It was also used to obtain $\hat{H}(t)$ values for pixels in different rings of the CMB sky sphere using the Hölder exponent approach (Table 4.1) and R/S method (Table 4.2) as shown in Chapter 4.

```
library(rcosmo)
library(fractal)

#Link to download the CMB data set with resolution 2048
#URL = 'http://irsa.ipac.caltech.edu/data/Planck/release_2/all-
    sky-maps/maps/component-maps/cmb/COM_CMB_IQU-smica-field-Int_
    2048_R2.01_full.fits'
#downloadCMBMap(foreground = "smica", nside = 2048)

#Generating CMB data frame with nside=2048
cmbdf <- CMBDataFrame("CMB_map_smica2048.fits", ordering="ring")
df1 <- coords(cmbdf, new.coords = "spherical")
#CMB_coord-coordinates of the CMB data
CMB_coord <- data.frame(cbind(theta1 = df1$theta, phi1=df1$phi))

#Storing the intensities of pixels into a vector
Int1 <- cmbdf$I
#Computing the scaled CMB intensities
Scale <- max(abs(Int1 - mean(Int1)))
Int1 <- Int1 / max(abs(Int1))

#d-dimension
d <- 1
N_side <- 2048

#Tot_Ring is the total number of rings having the N_side value
Tot_Ring <- ((4 * (N_side)) - 1)
```

```r
#Tot_RingP is the total number of pixels in the CMB sky sphere
Tot_RingP <- (12 * (N_side)^2)

#Tot_UPR is the total number of rings in the upper part
Tot_UPR <- (N_side - 1)

#Tot_ULPRP is the total number of pixels in the upper part =
   last pixel number of the upper part = Total number of pixels
   in the lower part
Tot_ULPRP <- (2 * (N_side) * (N_side - 1))

#Tot_MPR is the total number of rings in the middle part
Tot_MPR <- ((2 * N_side) + 1)

#Tot_UMPRP is the total number of pixels in the upper and middle
    parts together
Tot_UMPRP <- ((Tot_RingP) - (Tot_ULPRP))

#Tot_LPR is the total number of rings in the lower part
Tot_LPR <- (N_side - 1)

#These functions give the ring number of the specific pixel
#Upperpolar_Ringno function gives the ring number of the pixel
   if the pixel belongs to the upper part
Upperpolar_Ringno <- function(CMB_row) {
  U_ring <- 0
  ppixel_count <- 0
  pixelring_count <- 0
  for (i in 1:(N_side - 1)) {
    pixelring_count <- pixelring_count + 4
    ppixel_count <- ppixel_count + pixelring_count
    if (CMB_row <= ppixel_count) {
      U_ring <- i
      break
    }
  }
  return(U_ring)
}

#Middlepolar_Ringno function gives the ring number of the pixel
   if the pixel belongs to the middle part
Middlepolar_Ringno <- function(CMB_row) {
  M_ring <- as.integer((CMB_row - Tot_ULPRP - 1) / (4 * N_side))
    + N_side
  return(M_ring)
}

#Lowerpolar_Ringno function gives the ring number of the pixel
   if the pixel belongs to the lower part
#upCMB_row is the updated pixel number of the lower part but in
   reverse order
Lowerpolar_Ringno <- function(CMB_row) {
  upCMB_row <- (Tot_RingP - CMB_row) + 1
  U_ring <- Upperpolar_Ringno(upCMB_row)
  L_ring <- ((Tot_Ring - (U_ring)) + 1)
  return(L_ring)
}
```

```
#These functions find the first and last members of the ring and
    get all their member pixels into a vector
Upperpolar_firstlast <- function(Ring_no) {
  first_member <- (2 * (Ring_no - 1) * Ring_no) + 1
  last_member <- (2 * Ring_no * (Ring_no + 1))
  FL <- c(first_member, last_member)
  return(FL)
}

Middlepolar_firstlast <- function(Ring_no) {
  first_member <- ((Tot_ULPRP) + (((Ring_no - N_side) * (4 * N_
   side)) + 1))
  last_member <- ((Tot_ULPRP) + (((Ring_no - N_side) + 1) * (4 *
    N_side)))
  FL <- c(first_member, last_member)
  return(FL)
}

Lowerpolar_firstlast <- function(Ring_no) {
  first_member <- Tot_RingP - (2 * (Tot_Ring - Ring_no + 1) * (
   Tot_Ring - Ring_no + 2)) + 1
  last_member <- Tot_RingP - ((2 * (Tot_Ring - Ring_no) * (Tot_
   Ring - Ring_no + 1)) + 1) + 1
  FL <- c(first_member, last_member)
  return(FL)
}

#Checking upper part rings
#[I]-Checking the random pixel "552300", where ring_no=525
#The function Ring_pix gives the ring number, total number of
   pixels and the member pixels into which the given pixel
   belongs
Ring_pix <- function(CMB_row) {
  if (CMB_row <= Tot_ULPRP) {
    Ring_no <- Upperpolar_Ringno(CMB_row)
    Tot_pixelR <- (4 * Ring_no)
    FLm <- c(Upperpolar_firstlast(Ring_no))
    MemberP <- c(seq(FLm[1], FLm[2], 1))
    print("Upper Polar")
  } else if (CMB_row > Tot_UMPRP) {
    Ring_no <- Lowerpolar_Ringno(CMB_row)
    Tot_pixelR <- (4 * (Tot_Ring - Ring_no + 1))
    FLm <- c(Lowerpolar_firstlast(Ring_no))
    MemberP <- c(seq(FLm[1], FLm[2], 1))
    print("Lower polar")
  } else {
    Ring_no <- Middlepolar_Ringno(CMB_row)
    Tot_pixelR <- (4 * N_side)
    FLm <- c(Middlepolar_firstlast(Ring_no))
    MemberP <- c(seq(FLm[1], FLm[2], 1))
    print("Middle part")
  }
  return(cbind(Ring_no, Tot_pixelR, MemberP))
}
Ring_info <- data.frame(Ring_pix(552300))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP
```

```
#Once we know the ring number of the given pixel, finding the
   radius of the ring into which the given pixel belongs
#Let Rd be the ring deviation from center
Rd <- abs(Ring_no - (2 * N_side))

#Let R be the unit radius of the sphere
R <- 1

#Let Rp be the radius of the ring into which the given pixel
   belongs
Rp <- (R * cos((pi / (4 * N_side)) * Rd))

#Finding the number of pixels in the half-circle
N1 <- round((Tot_pixelR / (2)), 0)

#Let lp be the distance between two pixels in this interval
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D1 <- rep(0, N2)
HExp1D1c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
   integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp1D1[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  HExp1D1c[l] <- RoverS(Int1[A])
  l <- l + 1
}
HExp1D1
min(HExp1D1)
max(HExp1D1)
max(HExp1D1) - min(HExp1D1)
mean(HExp1D1)

HExp1D1c
min(HExp1D1c)
max(HExp1D1c)
max(HExp1D1c) - min(HExp1D1c)
mean(HExp1D1c)
```

```
#This figure gives the plot of scaled CMB intensities of ring
   525
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
   = "I", type = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 525
plot(x = df1, y = HExp1D1, xlab = "CMB pixel", ylab = "H", type
   = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 525 by R/S
   method
plot(x = df1, y = HExp1D1c, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")

#[II]-Checking the random pixel "1533000", where ring_no=875
Ring_info <- data.frame(Ring_pix(1533000))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)
r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D2 <- rep(0, N2)
HExp1D2c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix]-Pix_int,df1[CMB_pix]+Pix_int,1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
   integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1 }
  VNt <- tot
  HExp1D2[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  HExp1D2c[l] <- RoverS(Int1[A])
  l <- l + 1 }
HExp1D2
min(HExp1D2)
max(HExp1D2)
max(HExp1D2) - min(HExp1D2)
mean(HExp1D2)

HExp1D2c
min(HExp1D2c)
max(HExp1D2c)
max(HExp1D2c) - min(HExp1D2c)
```

```
mean(HExp1D2c)

#This figure gives the plot of scaled CMB intensities of ring
    875
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
    = "I", type = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 875
plot(x = df1, y = HExp1D2, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 875 by R/S
    method
plot(x = df1, y = HExp1D2c, xlab = "CMB pixel", ylab = "H", type
     = "l", lwd = 1, col = "blue")

#[III]-Checking the random pixel "3253800", where ring_no=1275
Ring_info <- data.frame(Ring_pix(3253800))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D3 <- rep(0, N2)
HExp1D3c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
   integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp1D3[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  HExp1D3c[l] <- RoverS(Int1[A])
  l <- l + 1
}
HExp1D3
min(HExp1D3)
max(HExp1D3)
max(HExp1D3) - min(HExp1D3)
```

```
mean(HExp1D3)

HExp1D3c
min(HExp1D3c)
max(HExp1D3c)
max(HExp1D3c) - min(HExp1D3c)
mean(HExp1D3c)

#Figure 4.2(a)-This figure gives the plot of scaled CMB
    intensities of ring 1275
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
    = "I", type = "l", lwd = 1, col = "blue")

#Figure 4.2(c)-This figure gives the plot of H values of ring
    1275
plot(x = df1, y = HExp1D3, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 1275 by R/S
    method
plot(x = df1, y = HExp1D3c, xlab = "CMB pixel", ylab = "H", type
     = "l", lwd = 1, col = "blue")

#Checking middle part rings
#[IV]-Checking the random pixel "10047488", where ring_no=2250
Ring_info <- data.frame(Ring_pix(10047488))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D4 <- rep(0, N2)
HExp1D4c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
   integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
```

```
  HExp1D4[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  HExp1D4c[l] <- RoverS(Int1[A])
  l <- l + 1
}
HExp1D4
min(HExp1D4)
max(HExp1D4)
max(HExp1D4) - min(HExp1D4)
mean(HExp1D4)

HExp1D4c
min(HExp1D4c)
max(HExp1D4c)
max(HExp1D4c) - min(HExp1D4c)
mean(HExp1D4c)

#This figure gives the plot of scaled CMB intensities of ring
   2250
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
   = "I", type = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 2250
plot(x = df1, y = HExp1D4, xlab = "CMB pixel", ylab = "H", type
   = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 2250 by R/S
   method
plot(x = df1, y = HExp1D4c, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")

#[V]-Choosing a pixel near the equator/great circle, where ring_
   no=5000
Ring_info <- data.frame(Ring_pix(32575488))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D5 <- rep(0, N2)
HExp1D5c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
   , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
```

```
    tot <- 0
    for (j in 1:(N - 2)) {
      Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
      integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
      tot <- tot + Increment1
    }
    VNt <- tot
    HExp1D5[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
      N1)))
    HExp1D5c[l] <- RoverS(Int1[A])
    l <- l + 1
}
HExp1D5
min(HExp1D5)
max(HExp1D5)
max(HExp1D5) - min(HExp1D5)
mean(HExp1D5)

HExp1D5c
min(HExp1D5c)
max(HExp1D5c)
max(HExp1D5c) - min(HExp1D5c)
mean(HExp1D5c)

#This figure gives the plot of scaled CMB intensities of ring
    5000
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
    = "I", type = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 5000
plot(x = df1, y = HExp1D5, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 5000 by R/S
    method
plot(x = df1, y = HExp1D5c, xlab = "CMB pixel", ylab = "H", type
     = "l", lwd = 1, col = "blue")

#[VI]-Checking the random pixel "39948288", where ring_no=5900
Ring_info <- data.frame(Ring_pix(39948288))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D6 <- rep(0, N2)
```

```
HExp1D6c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
   integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp1D6[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  HExp1D6c[l] <- RoverS(Int1[A])
  l <- l + 1
}
HExp1D6
min(HExp1D6)
max(HExp1D6)
max(HExp1D6) - min(HExp1D6)
mean(HExp1D6)

HExp1D6c
min(HExp1D6c)
max(HExp1D6c)
max(HExp1D6c) - min(HExp1D6c)
mean(HExp1D6c)

#Figure 4.2(b)-This figure gives the plot of scaled CMB
    intensities of ring 5900
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
    = "I", type = "l", lwd = 1, col = "blue")

#Figure 4.2(d)-This figure gives the plot of H values of ring
    5900
plot(x = df1, y = HExp1D6, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 5900 by R/S
    method
plot(x = df1, y = HExp1D6c, xlab = "CMB pixel", ylab = "H", type
     = "l", lwd = 1, col = "blue")

#Checking lower part rings
#[VII]-Checking the random pixel "47656664", where ring_no=7035
Ring_info <- data.frame(Ring_pix(47656664))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)

r <- 0.08
```

```
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D7 <- rep(0, N2)
HExp1D7c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
  integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp1D7[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  HExp1D7c[l] <- RoverS(Int1[A])
  l <- l + 1
}
HExp1D7
min(HExp1D7)
max(HExp1D7)
max(HExp1D7) - min(HExp1D7)
mean(HExp1D7)

HExp1D7c
min(HExp1D7c)
max(HExp1D7c)
max(HExp1D7c) - min(HExp1D7c)
mean(HExp1D7c)

#This figure gives the plot of scaled CMB intensities of ring
   7035
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
   = "I", type = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 7035
plot(x = df1, y = HExp1D7, xlab = "CMB pixel", ylab = "H", type
   = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 7035 by R/S
   method
plot(x = df1, y = HExp1D7c, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")

#[VIII]-Checking the random pixel "48651704, where ring_no=7275"
Ring_info <- data.frame(Ring_pix(48651704))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
```

```
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D8 <- rep(0, N2)
HExp1D8c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
    integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp1D8[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  HExp1D8c[l] <- RoverS(Int1[A])
  l <- l + 1
}
HExp1D8
min(HExp1D8)
max(HExp1D8)
max(HExp1D8) - min(HExp1D8)
mean(HExp1D8)

HExp1D8c
min(HExp1D8c)
max(HExp1D8c)
max(HExp1D8c) - min(HExp1D8c)
mean(HExp1D8c)

#This figure gives the plot of scaled CMB intensities of ring
   7275
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
   = "I", type = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 7275
plot(x = df1, y = HExp1D8, xlab = "CMB pixel", ylab = "H", type
   = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 7275 by R/S
   method
plot(x = df1, y = HExp1D8c, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")
```

```
#[XI]-Checking the random pixel "49375304", where ring_no=7500
Ring_info <- data.frame(Ring_pix(49375304))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP

Rd <- abs(Ring_no - (2 * N_side))
R <- 1
Rp <- (R * cos((pi / (4 * N_side)) * Rd))
N1 <- round((Tot_pixelR / (2)), 0)
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1D9 <- rep(0, N2)
HExp1D9c <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
   integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp1D9[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
    N1)))
  HExp1D9c[l] <- RoverS(Int1[A])
  l <- l + 1
}
HExp1D9
min(HExp1D9)
max(HExp1D9)
max(HExp1D9) - min(HExp1D9)
mean(HExp1D9)

HExp1D9c
min(HExp1D9c)
max(HExp1D9c)
max(HExp1D9c) - min(HExp1D9c)
mean(HExp1D9c)

#This figure gives the plot of scaled CMB intensities of ring
    7500
plot(x = MemberP1, y = Int1[MemberP1], xlab = "CMB pixel", ylab
    = "I", type = "l", lwd = 1, col = "blue")

#This figure gives the plot of H values of ring 7500
plot(x = df1, y = HExp1D9, xlab = "CMB pixel", ylab = "H", type
    = "l", lwd = 1, col = "blue")
```

```
#This figure gives the plot of H values of ring 7500 by R/S
   method
plot(x = df1, y = HExp1D9c, xlab = "CMB pixel", ylab = "H", type
   = "l", lwd = 1, col = "blue")
save.image(file = "HolderExponent1D_2048.RData")
```

B.3 R code used to produce Figure 4.3 and results in Table 4.3

The R code in this section was used to produce Figure 4.3 and to obtain results in Table 4.3 in Chapter 4. The code in this section visualizes the distribution of $\hat{H}(t)$ values of four rim segments. It was also used to obtain p-values for Wilcoxon tests between different rings as shown in Table 4.3 of Chapter 4.

```
library(ggplot2)
load("HolderExponent1D_2048.RData")

Test1 <- cbind(R = rep(1, length(HExp1D1)), H = HExp1D1)
Test3 <- cbind(R = rep(3, length(HExp1D3)), H = HExp1D3)
Test4 <- cbind(R = rep(4, length(HExp1D4)), H = HExp1D4)
Test9 <- cbind(R = rep(9, length(HExp1D9)), H = HExp1D9)

Hexp1 <- data.frame(rbind(Test1, Test3, Test4, Test9))
Hexp1$R <- factor(Hexp1$R, levels = c("1", "3", "4", "9"),
   labels = c(525, 1275, 2250, 7500))

#Figure 4.3-This figure gives the plot of the distribution of $\
   hat{H}(t)$ values of four rim segments
ggplot(Hexp1, aes(x = R, y = H, fill = R)) +
  geom_boxplot(width = 0.45, position = position_dodge(width =
   0.9), fatten = NULL) +
  labs(x = "Ring no.", y = "H(t)") +
  scale_fill_manual(name = "Ring no.", labels = c("525", "1275",
    "2250", "7500"), values = c("#CC0033", "#FFCC33", "#00AFBB",
    "#9933CC")) +
  stat_summary(fun = mean, geom = "errorbar", aes(ymax = ..y..,
   ymin = ..y..), width = 0.45, linetype = "solid", size = 1) +
  theme_bw()+theme(legend.position = "none")

#Checking for statistical significance in difference of means
#[1]-Difference between rings 1 & 3
#Using the proposed H method and all H(t) values
t.test(HExp1D1, HExp1D3)
wilcox.test(HExp1D1, HExp1D3, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D1[seq(1, length(HExp1D1), by = 50)])
shapiro.test(HExp1D3[seq(1, length(HExp1D3), by = 50)])
```

```
#Carrying out the T-test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D1[seq(1, length(HExp1D1), by = 50)], HExp1D3[seq(1,
    length(HExp1D3), by = 50)])
wilcox.test(HExp1D1[seq(1, length(HExp1D1), by = 50)], HExp1D3[
   seq(1, length(HExp1D3), by = 50)], alternative = "two.sided")

#Using the R/S method and all H(t) values
t.test(HExp1D1c, HExp1D3c)
wilcox.test(HExp1D1c, HExp1D3c, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)])
shapiro.test(HExp1D3c[seq(1, length(HExp1D3c), by = 50)])

#Carrying out the T-test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)], HExp1D3c[seq
   (1, length(HExp1D3c), by = 50)])
wilcox.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)],
   HExp1D3c[seq(1, length(HExp1D3c), by = 50)], alternative = "
   two.sided")

#[2]-Difference between rings 1 & 4
#Using the proposed H method and all H(t) values
t.test(HExp1D1, HExp1D4)
wilcox.test(HExp1D1, HExp1D4, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D1[seq(1, length(HExp1D1), by = 50)])
shapiro.test(HExp1D4[seq(1, length(HExp1D4), by = 50)])

#Carrying out the T-test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D1[seq(1, length(HExp1D1), by = 50)], HExp1D4[seq(1,
    length(HExp1D4), by = 50)])
wilcox.test(HExp1D1[seq(1, length(HExp1D1), by = 50)], HExp1D4[
   seq(1, length(HExp1D4), by = 50)], alternative = "two.sided")

#Using the R/S method and all H(t) values
t.test(HExp1D1c, HExp1D4c)
wilcox.test(HExp1D1c, HExp1D4c, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)])
shapiro.test(HExp1D4c[seq(1, length(HExp1D4c), by = 50)])

#Carrying out the T-test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)], HExp1D4c[seq
   (1, length(HExp1D4c), by = 50)])
wilcox.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)],
   HExp1D4c[seq(1, length(HExp1D4c), by = 50)], alternative = "
   two.sided")
```

```
#[3]−Difference between rings 1 & 9
#Using the proposed H method and all H(t) values
t.test(HExp1D1, HExp1D9)
wilcox.test(HExp1D1, HExp1D9, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D1[seq(1, length(HExp1D1), by = 50)])
shapiro.test(HExp1D9[seq(1, length(HExp1D9), by = 50)])

#Carrying out the T−test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D1[seq(1, length(HExp1D1), by = 50)], HExp1D9[seq(1,
    length(HExp1D9), by = 50)])
wilcox.test(HExp1D1[seq(1, length(HExp1D1), by = 50)], HExp1D9[
   seq(1, length(HExp1D9), by = 50)], alternative = "two.sided")

#Using the R/S method and all H(t) values
t.test(HExp1D1c, HExp1D9c)
wilcox.test(HExp1D1c, HExp1D9c, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)])
shapiro.test(HExp1D9c[seq(1, length(HExp1D6c), by = 50)])

#Carrying out the T−test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)], HExp1D9c[seq
   (1, length(HExp1D9c), by = 50)])
wilcox.test(HExp1D1c[seq(1, length(HExp1D1c), by = 50)],
   HExp1D9c[seq(1, length(HExp1D9c), by = 50)], alternative = "
   two.sided")

#[4]−Difference between rings 3 & 4
# Using the proposed H method and all H(t) values
t.test(HExp1D3, HExp1D4)
wilcox.test(HExp1D3, HExp1D4, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D3[seq(1, length(HExp1D3), by = 50)])
shapiro.test(HExp1D4[seq(1, length(HExp1D4), by = 50)])

#Carrying out the T−test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D3[seq(1, length(HExp1D3), by = 50)], HExp1D4[seq(1,
    length(HExp1D4), by = 50)])
wilcox.test(HExp1D3[seq(1, length(HExp1D3), by = 50)], HExp1D4[
   seq(1, length(HExp1D4), by = 50)], alternative = "two.sided")

#Using the R/S method and all H(t) values
t.test(HExp1D3c, HExp1D4c)
wilcox.test(HExp1D3c, HExp1D4c, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D3c[seq(1, length(HExp1D3c), by = 50)])
shapiro.test(HExp1D4c[seq(1, length(HExp1D4c), by = 50)])
```

```
#Carrying out the T-test and Wilcoxon's test to check for
    difference in means
t.test(HExp1D3c[seq(1, length(HExp1D3c), by = 50)], HExp1D4c[seq
    (1, length(HExp1D4c), by = 50)])
wilcox.test(HExp1D3c[seq(1, length(HExp1D3c), by = 50)],
    HExp1D4c[seq(1, length(HExp1D4c), by = 50)], alternative = "
    two.sided")

#[5]-Difference between rings 3 & 9
# Using the proposed H method and all H(t) values
t.test(HExp1D3, HExp1D9)
wilcox.test(HExp1D3, HExp1D9, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
    observations using H(t) values with step 50
shapiro.test(HExp1D3[seq(1, length(HExp1D3), by = 50)])
shapiro.test(HExp1D9[seq(1, length(HExp1D9), by = 50)])

#Carrying out the T-test and Wilcoxon's test to check for
    difference in means
t.test(HExp1D3[seq(1, length(HExp1D3), by = 50)], HExp1D9[seq(1,
     length(HExp1D9), by = 50)])
wilcox.test(HExp1D3[seq(1, length(HExp1D3), by = 50)], HExp1D9[
    seq(1, length(HExp1D9), by = 50)], alternative = "two.sided")

#Using the R/S method and all H(t) values
t.test(HExp1D3c, HExp1D9c)
wilcox.test(HExp1D3c, HExp1D9c, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
    observations using H(t) values with step 50
shapiro.test(HExp1D3c[seq(1, length(HExp1D3c), by = 50)])
shapiro.test(HExp1D9c[seq(1, length(HExp1D9c), by = 50)])

#Carrying out the T-test and Wilcoxon's test to check for
    difference in means
t.test(HExp1D3c[seq(1, length(HExp1D3c), by = 50)], HExp1D9c[seq
    (1, length(HExp1D9c), by = 50)])
wilcox.test(HExp1D3c[seq(1, length(HExp1D3c), by = 50)],
    HExp1D9c[seq(1, length(HExp1D9c), by = 50)], alternative = "
    two.sided")

#[6]-Difference between rings 4 & 9
#Using the proposed H method and all H(t) values
t.test(HExp1D4, HExp1D9)
wilcox.test(HExp1D4, HExp1D9, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
    observations using H(t) values with step 50
shapiro.test(HExp1D4[seq(1, length(HExp1D4), by = 50)])
shapiro.test(HExp1D9[seq(1, length(HExp1D9), by = 50)])

#Carrying out the T-test and Wilcoxon's test to check for
    difference in means
t.test(HExp1D4[seq(1, length(HExp1D4), by = 50)], HExp1D9[seq(1,
     length(HExp1D9), by = 50)])
wilcox.test(HExp1D4[seq(1, length(HExp1D4), by = 50)], HExp1D9[
    seq(1, length(HExp1D9), by = 50)], alternative = "two.sided")
```

```
#Using the R/S method and all H(t) values
t.test(HExp1D4c, HExp1D9c)
wilcox.test(HExp1D4c, HExp1D9c, alternative = "two.sided")

#Checking the normality assumption ensuring the independence of
   observations using H(t) values with step 50
shapiro.test(HExp1D4c[seq(1, length(HExp1D4c), by = 50)])
shapiro.test(HExp1D9c[seq(1, length(HExp1D9c), by = 50)])

#Carrying out the T-test and Wilcoxon's test to check for
   difference in means
t.test(HExp1D4c[seq(1, length(HExp1D4c), by = 50)], HExp1D9c[seq
   (1, length(HExp1D9c), by = 50)])
wilcox.test(HExp1D4c[seq(1, length(HExp1D4c), by = 50)],
   HExp1D9c[seq(1, length(HExp1D9c), by = 50)], alternative = "
   two.sided")
```

B.4 R code used to produce Figure 4.4

The R code in this section was used to produce Figure 4.4 in Chapter 4. The code in this section visualizes examples of pixels with 7 and 8 neighbours for $N_{side} = 4$ as shown in Chapter 4.

```
library(rcosmo)
library(rgl)

#This function plots the neighbouring pixels for a given base
   pixel
demoNeighbours <- function(p, j, fcol) {
  neighbours(p, j)
  displayPixels(
    boundary.j = j, j = j, plot.j = j + 3,
    spix = neighbours(p, j),
    boundary.col = "grey48",
    boundary.lwd = 1,
    incl.labels = neighbours(p, j),
    col = toString(fcol),
    size = 3
  )
}

#Figure 4.4-This figure gives the plot of examples of pixels
   with 7 and 8 neighbours for N_side=4
#Plotting the neighbouring pixels for pixel index 6 at N_side=4
demoNeighbours(6, 2, 3)
#Plotting the neighbouring pixels for pixel index 72 at N_side=4
demoNeighbours(72, 2, 4)
um <- matrix(c(0.7560294, -0.6384937, -0.1440324, 0, -0.1466970,
    -0.3797444, 0.9133860, 0, -0.6378868, -0.6694176,
   -0.3807631, 0, 0, 0, 0, 1), byrow = TRUE, nrow = 4, ncol = 4)
view3d(userMatrix = um)
clipplanes3d(1, 1, 0.65, d = 0.6)
rgl.snapshot("Figure44.png")
```

```
#Generating CMB data frame with nside=2048
Nside <- 2048
df <- CMBDataFrame(nside = Nside, I = rep(0, 12 * Nside^2),
    ordering = "nested")
N <- 12 * Nside^2
j1 <- as.integer(log(Nside) / log(2))

#This function computes the number of neighbouring pixels of a
    given pixel
fincr2d <- function(x) {
  length(neighbours(x, j1))
}
#Vectorizing the "fincr2d" function
fincr2dv <- Vectorize(fincr2d)

#Computing the number of neighbouring pixels of all pixel
    indices
df$I <- fincr2dv(1:N)

#Finding the number of pixels having 8 and 9 neighbours
    respectively
sum(df$I == 8)
sum(df$I == 9)
```

B.5 R code used to produce Figures 4.5, 4.6 and results in Table 4.4

The R code in this section was used to produce Figures 4.5, 4.6 and to obtain results in Table 4.4 in Chapter 4. The code in this section visualizes the sky windows used for computations (Figure 4.5) and local estimates $\hat{H}(t)$ for two-dimensional regions (Figure 4.6). It was also used to obtain $\hat{H}(t)$ values for pixels in different sky windows of the CMB sky sphere as shown in Table 4.4 of Chapter 4.

```
library(rcosmo)
library(rgl)
library(akima)

#Generating CMB data frame with nside=2048
cmbdf <- CMBDataFrame("CMB_map_smica2048.fits")
df1 <- coords(cmbdf, new.coords = "cartesian")

#CMB_coord-coordinates of the CMB data
CMB_coord <- data.frame(cbind(x = df1$x, y = df1$y, z = df1$z))

#Storing the intensities of pixels into a vector
Int1 <- cmbdf$I
# Computing the scaled CMB intensities
Int1 <- Int1 / max(abs(Int1))

#k is changing with nside. If nside=2048=2^(11)=2^(k)
k <- 11
```

```
#d-dimension
d <- 2

#N_pix is the no. of pixels in which we compute the Holder
   exponent values
N_pix <- 1000

#For a warm region
#Figure 4.5(a)-This figure gives the plot of a sky window from
   the warm region
plot(cmbdf, back.col = "white", ylab = "", xlab = "", zlab = "")
CMB_row <- (29990264)
win <- CMBWindow(x = df1[CMB_row, ]$x, y = df1[CMB_row, ]$y, z =
    df1[CMB_row, ]$z, r = 0.23)
plot(win, col = "red", lwd = 3)

cmbdf11 <- window(cmbdf, new.window = win)
length(pix(cmbdf11))
min(cmbdf11$I)
max(cmbdf11$I)
avg1 <- mean(cmbdf11$I)
df11 <- coords(cmbdf11, new.coords = "cartesian")
df_sample1 <- df11[seq(1, length(pix(cmbdf11)), by = (length(pix
   (cmbdf11)) / N_pix)), ]
df12 <- coords(df_sample1, new.coords = "cartesian")
df13 <- data.frame(cbind(x = df12$x, y = df12$y, z = df12$z))

r <- 0.01
N11 <- length(pix(cmbdf11))
N12 <- sqrt(N11)
gamma <- (-(log((sqrt(pi) * r) / 2) / log((N12))))
HExp2D1 <- rep(0, N_pix)
l <- 1
for (CMB_row in 1:N_pix) {
  win1 <- CMBWindow(x = df13[CMB_row, ]$x, y = df13[CMB_row, ]$y
    , z = df13[CMB_row, ]$z, r = 0.01)
  cmbdf1 <- window(cmbdf, new.window = win1)
  N <- length(pix(cmbdf1))
  Pix_number <- as.integer(pix(cmbdf1))
  tot <- 0
  for (i in 1:N) {
    A <- neighbours(Pix_number[i], k)
    B <- neighbours(as.integer(A[1]), k)
    Increment1 <- (((Int1[as.integer(B[1])]) - (2 * Int1[as.
   integer(B[2])]) + (Int1[as.integer(B[3])]) -
      (2 * Int1[as.integer(B[4])]) + (Int1[as.integer(B[5])]) -
   (2 * Int1[as.integer(B[6])]) +
      (Int1[as.integer(B[7])]) - (2 * Int1[as.integer(B[8])]) +
   (4 * Int1[as.integer(B[9])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp2D1[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N12)))
  l <- l + 1
}
HExp2D1
min(HExp2D1)
max(HExp2D1)
```

```
max(HExp2D1) - min(HExp2D1)
mean(HExp2D1)

df14 <- coords(df_sample1, new.coords = "spherical")
n_interpolation <- 500

x <- df14$theta
y <- df14$phi
z <- HExp2D1

spline_interpolated <- interp(x, y, z, xo = seq(min(x), max(x),
   length = n_interpolation), yo = seq(min(y), max(y), length =
   n_interpolation))

x.si <- spline_interpolated$x
y.si <- spline_interpolated$y
z.si <- spline_interpolated$z

nbcol <- 50
color <- rev(rainbow(nbcol, start = 0, end = 1))
zcol <- cut(z.si, nbcol)

#Figure 4.6(a)-This figure gives the plot of $\hat{H}(t)$ values
    from the warm region
persp3d(x.si, y.si, z.si, xlab = expression(theta), ylab =
   expression(varphi), zlab = "H", col = color[zcol])

# For a cold region
# Figure 4.5(b)-This figure gives the plot of a sky window from
   the cold region
plot(cmbdf, back.col = "white", ylab = "", xlab = "", zlab = "")
CMB_row <- (45045200)
win <- CMBWindow(x = df1[CMB_row, ]$x, y = df1[CMB_row, ]$y, z =
    df1[CMB_row, ]$z, r = 0.23)
plot(win, col = "red", lwd = 3)

cmbdf12 <- window(cmbdf, new.window = win)
length(pix(cmbdf12))
min(cmbdf12$I)
max(cmbdf12$I)
avg2 <- mean(cmbdf12$I)
df21 <- coords(cmbdf12, new.coords = "cartesian")
df_sample2 <- df21[seq(1, length(pix(cmbdf12)), by = (length(pix
   (cmbdf12)) / N_pix)), ]
df22 <- coords(df_sample2, new.coords = "cartesian")
df23 <- data.frame(cbind(x = df22$x, y = df22$y, z = df22$z))

r <- 0.01
N21 <- length(pix(cmbdf12))
N22 <- sqrt(N21)
gamma <- (-(log((sqrt(pi) * r) / 2) / log((N22))))
HExp2D2 <- rep(0, N_pix)
l <- 1
for (CMB_row in 1:N_pix) {
  win1 <- CMBWindow(x = df23[CMB_row, ]$x, y = df23[CMB_row, ]$y
    , z = df23[CMB_row, ]$z, r = 0.01)
  cmbdf1 <- window(cmbdf, new.window = win1)
  N <- length(pix(cmbdf1))
  Pix_number <- as.integer(pix(cmbdf1))
```

```
  tot <- 0
  for (i in 1:N) {
    A <- neighbours(Pix_number[i], k)
    B <- neighbours(as.integer(A[1]), k)
    Increment1 <- (((Int1[as.integer(B[1])]) - (2 * Int1[as.
   integer(B[2])]) + (Int1[as.integer(B[3])]) -
      (2 * Int1[as.integer(B[4])]) + (Int1[as.integer(B[5])]) -
   (2 * Int1[as.integer(B[6])]) +
      (Int1[as.integer(B[7])]) - (2 * Int1[as.integer(B[8])]) +
   (4 * Int1[as.integer(B[9])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp2D2[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N22)))
  l <- l + 1
}
HExp2D2
min(HExp2D2)
max(HExp2D2)
max(HExp2D2) - min(HExp2D2)
mean(HExp2D2)

df24 <- coords(df_sample2, new.coords = "spherical")
n_interpolation <- 500

x <- df24$theta
y <- df24$phi
z <- HExp2D2

spline_interpolated <- interp(x, y, z, xo = seq(min(x), max(x),
   length = n_interpolation), yo = seq(min(y), max(y), length =
   n_interpolation))

x.si <- spline_interpolated$x
y.si <- spline_interpolated$y
z.si <- spline_interpolated$z

nbcol <- 50
color <- rev(rainbow(nbcol, start = 0, end = 1))
zcol <- cut(z.si, nbcol)

#Figure 4.6(b)-This figure gives the plot of $\hat{H}(t)$ values
    from the cold region
persp3d(x.si, y.si, z.si, xlab = expression(theta), ylab =
   expression(varphi), zlab = "H", col = color[zcol])

#For a mixture of warm and cold regions
#Figure 4.5(c)-This figure gives the plot of a sky window from
   the mixture region
plot(cmbdf, back.col = "white", ylab = "", xlab = "", zlab = "")
CMB_row <- (25163208)
win <- CMBWindow(x = df1[CMB_row, ]$x, y = df1[CMB_row, ]$y, z =
    df1[CMB_row, ]$z, r = 0.23)
plot(win, col = "red", lwd = 3)

cmbdf13 <- window(cmbdf, new.window = win)
length(pix(cmbdf13))
min(cmbdf13$I)
```

```
max(cmbdf13$I)
avg3 <- mean(cmbdf13$I)
df31 <- coords(cmbdf13, new.coords = "cartesian")
df_sample3 <- df31[seq(1, length(pix(cmbdf13)), by = (length(pix
    (cmbdf13)) / N_pix)), ]
df32 <- coords(df_sample3, new.coords = "cartesian")
df33 <- data.frame(cbind(x = df32$x, y = df32$y, z = df32$z))

r <- 0.01
N31 <- length(pix(cmbdf13))
N32 <- sqrt(N31)
gamma <- (-(log((sqrt(pi) * r) / 2) / log((N32))))
HExp2D3 <- rep(0, N_pix)
l <- 1
for (CMB_row in 1:N_pix) {
  win1 <- CMBWindow(x = df33[CMB_row, ]$x, y = df33[CMB_row, ]$y
    , z = df33[CMB_row, ]$z, r = 0.01)
  cmbdf1 <- window(cmbdf, new.window = win1)
  N <- length(pix(cmbdf1))
  Pix_number <- as.integer(pix(cmbdf1))
  tot <- 0
  for (i in 1:N) {
    A <- neighbours(Pix_number[i], k)
    B <- neighbours(as.integer(A[1]), k)
    Increment1 <- (((Int1[as.integer(B[1])]) - (2 * Int1[as.
    integer(B[2])]) + (Int1[as.integer(B[3])]) -
       (2 * Int1[as.integer(B[4])]) + (Int1[as.integer(B[5])]) -
    (2 * Int1[as.integer(B[6])]) +
       (Int1[as.integer(B[7])]) - (2 * Int1[as.integer(B[8])]) +
    (4 * Int1[as.integer(B[9])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp2D3[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
    N32)))
  l <- l + 1
}
HExp2D3
min(HExp2D3)
max(HExp2D3)
max(HExp2D3) - min(HExp2D3)
mean(HExp2D3)

df34 <- coords(df_sample3, new.coords = "spherical")
n_interpolation <- 500

x <- df34$theta
y <- df34$phi
z <- HExp2D3

spline_interpolated <- interp(x, y, z, xo = seq(min(x), max(x),
    length = n_interpolation), yo = seq(min(y), max(y), length =
    n_interpolation))

x.si <- spline_interpolated$x
y.si <- spline_interpolated$y
z.si <- spline_interpolated$z

nbcol <- 50
```

```
color <- rev(rainbow(nbcol, start = 0, end = 1))
zcol <- cut(z.si, nbcol)

#Figure 4.6(c)-This figure gives the plot of $\hat{H}(t)$ values
    from the region with mixture of temperatures
persp3d(x.si, y.si, z.si, xlab = expression(theta), ylab =
   expression(varphi), zlab = "H", col = color[zcol])

#For a borderline region
#Figure 4.5(d)-This figure gives the plot of a sky window from
   the borderline region
plot(cmbdf, back.col = "white", ylab = "", xlab = "", zlab = "")
CMB_row <- (42662192)
win <- CMBWindow(x = df1[CMB_row, ]$x, y = df1[CMB_row, ]$y, z =
    df1[CMB_row, ]$z, r = 0.23)
plot(win, col = "red", lwd = 3)

cmbdf14 <- window(cmbdf, new.window = win)
length(pix(cmbdf14))
min(cmbdf14$I)
max(cmbdf14$I)
avg4 <- mean(cmbdf14$I)
df41 <- coords(cmbdf14, new.coords = "cartesian")
df_sample4 <- df41[seq(1, length(pix(cmbdf14)), by = (length(pix
   (cmbdf14)) / N_pix)), ]
df42 <- coords(df_sample4, new.coords = "cartesian")
df43 <- data.frame(cbind(x = df42$x, y = df42$y, z = df42$z))

r <- 0.01
N41 <- length(pix(cmbdf14))
N42 <- sqrt(N41)
gamma <- (-(log((sqrt(pi) * r) / 2) / log((N42))))
HExp2D4 <- rep(0, N_pix)
l <- 1
for (CMB_row in 1:N_pix) {
  win1 <- CMBWindow(x = df43[CMB_row, ]$x, y = df43[CMB_row, ]$y
    , z = df43[CMB_row, ]$z, r = 0.01)
  cmbdf1 <- window(cmbdf, new.window = win1)
  N <- length(pix(cmbdf1))
  Pix_number <- as.integer(pix(cmbdf1))
  tot <- 0
  for (i in 1:N) {
    A <- neighbours(Pix_number[i], k)
    B <- neighbours(as.integer(A[1]), k)
    Increment1 <- (((Int1[as.integer(B[1])]) - (2 * Int1[as.
   integer(B[2])]) + (Int1[as.integer(B[3])]) -
      (2 * Int1[as.integer(B[4])]) + (Int1[as.integer(B[5])]) -
   (2 * Int1[as.integer(B[6])]) +
      (Int1[as.integer(B[7])]) - (2 * Int1[as.integer(B[8])]) +
   (4 * Int1[as.integer(B[9])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp2D4[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N42)))
  l <- l + 1
}
HExp2D4
min(HExp2D4)
```

```
max(HExp2D4)
max(HExp2D4) - min(HExp2D4)
mean(HExp2D4)

df44 <- coords(df_sample4, new.coords = "spherical")
n_interpolation <- 500

x <- df44$theta
y <- df44$phi
z <- HExp2D4

spline_interpolated <- interp(x, y, z, xo = seq(min(x), max(x),
   length = n_interpolation), yo = seq(min(y), max(y), length =
   n_interpolation))

x.si <- spline_interpolated$x
y.si <- spline_interpolated$y
z.si <- spline_interpolated$z

nbcol <- 50
color <- rev(rainbow(nbcol, start = 0, end = 1))
zcol <- cut(z.si, nbcol)

#Figure 4.6(d)-This figure gives the plot of $\hat{H}(t)$ values
    from the borderline region
persp3d(x.si, y.si, z.si, xlab = expression(theta), ylab =
   expression(varphi), zlab = "H", col = color[zcol])
save.image(file = "HolderExponent2Dall_2048.RData")
```

B.6 R code used to produce Figure 4.7 and results in Table 4.5

The R code in this section was used to produce Figure 4.7 and to obtain results in Table 4.5 in Chapter 4. The code in this section visualizes the distribution of $\hat{H}(t)$ values for chosen sky windows. It was also used to obtain p-values for Wilcoxon tests between chosen sky windows as shown in Table 4.5 of Chapter 4.

```
library(ggplot2)
load("HolderExponent2Dall_2048.RData")

Test1 <- cbind(R = rep(1, length(HExp2D1)), H = HExp2D1)
Test2 <- cbind(R = rep(2, length(HExp2D2)), H = HExp2D2)
Test3 <- cbind(R = rep(3, length(HExp2D3)), H = HExp2D3)
Test4 <- cbind(R = rep(4, length(HExp2D4)), H = HExp2D4)
Hexp1 <- data.frame(rbind(Test1, Test2, Test3, Test4))
Hexp1$R <- factor(Hexp1$R, levels = c("1", "2", "3", "4"),
   labels = c("warm", "cold", "mixture", "borderline"))

#Figure 4.7-This figure gives the plot of the distribution of $\
   hat{H}(t)$ values for chosen sky windows
ggplot(Hexp1, aes(x = R, y = H, fill = R)) + geom_boxplot(width
   = 0.45, position = position_dodge(width = 0.9), fatten=NULL)
```

```
    + labs(x = "Type of sky window", y = "H") + scale_fill_manual
    (name = "Type of sky window", labels = c("warm", "cold", "
    mixture", "borderline"), values = c("#CC0033", "#FFCC33", "
    #00AFBB", "#9933CC")) + stat_summary(fun = mean, geom = "
    errorbar", aes(ymax = ..y.., ymin = ..y..), width = 0.45,
    linetype = "solid", size = 1) + theme_bw() + theme(legend.
    position = "none")

#Checking for statistical significance in difference of means
#[1]-Difference between warm and cold regions
#Checking the normality assumption of H(t) values
shapiro.test(HExp2D1)
shapiro.test(HExp2D2)

t.test(HExp2D1, HExp2D2)
wilcox.test(HExp2D1, HExp2D2, alternative = "two.sided")

#[2]-Difference between warm and mixture regions
#Checking the normality assumption of H(t) values
shapiro.test(HExp2D1)
shapiro.test(HExp2D3)

t.test(HExp2D1, HExp2D3)
wilcox.test(HExp2D1, HExp2D3, alternative = "two.sided")

#[3]-Difference between warm and borderline regions
#Checking the normality assumption of H(t) values
shapiro.test(HExp2D1)
shapiro.test(HExp2D4)

t.test(HExp2D1, HExp2D4)
wilcox.test(HExp2D1, HExp2D4, alternative = "two.sided")

#[4]-Difference between cold and mixture regions
#Checking the normality assumption of H(t) values
shapiro.test(HExp2D2)
shapiro.test(HExp2D3)

t.test(HExp2D2, HExp2D3)
wilcox.test(HExp2D2, HExp2D3, alternative = "two.sided")

#[5]-Difference between cold and borderline regions
#Checking the normality assumption of H(t) values
shapiro.test(HExp2D2)
shapiro.test(HExp2D4)

t.test(HExp2D2, HExp2D4)
wilcox.test(HExp2D2, HExp2D4, alternative = "two.sided")

#[6]-Difference between mixture and borderline regions
#Checking the normality assumption of H(t) values
shapiro.test(HExp2D3)
shapiro.test(HExp2D4)

t.test(HExp2D3, HExp2D4)
wilcox.test(HExp2D3, HExp2D4, alternative = "two.sided")
```

B.7 R code used to produce Figures 4.8a, 4.8b and results in Table 4.6

The R code in this section was used to produce Figures 4.8a, 4.8b and to obtain results in Table 4.6 in Chapter 4. The code in this section visualizes scaled intensities (Figure 4.8a) and $\hat{H}(t)$ values (Figure 4.8b) of great circle/ring 4096. It was also used to obtain some results in the analysis of CMB intensities near the equatorial region (Table 4.6) as shown in Chapter 4.

```
library(rcosmo)

#Generating CMB data frame with nside=2048
cmbdf <- CMBDataFrame("CMB_map_smica2048.fits", ordering="ring")
df1 <- coords(cmbdf, new.coords = "spherical")
#CMB_coord-coordinates of the real CMB data
CMB_coord <- data.frame(cbind(theta1 = df1$theta, phi1=df1$phi))

#Storing the intensities of pixels into a vector
Int1 <- cmbdf$I
Scale <- max(abs(Int1 - mean(Int1)))
Int1 <- Int1 / max(abs(Int1))

#d-dimension
d <- 1
N_side <- 2048

#Equator ring analysis
#Checking the random pixel "25161729", where Ring_no = 4096
#Equator ring
CMB_row <- (2 * (N_side) * (N_side - 1)) + (4 * ((N_side)^2))+1

Ring_info <- data.frame(Ring_pix(25161729))
Ring_no <- Ring_info$Ring_no[1]
Tot_pixelR <- Ring_info$Tot_pixelR[1]
MemberP <- Ring_info$MemberP
mean(cmbdf$I[MemberP])

#Once we know the ring number of the given pixel, finding the
   radius of the ring into which the given pixel belongs
#Let Rd be the ring deviation from center
Rd <- abs(Ring_no - (2 * N_side))

#Let R be the unit radius of the sphere
R <- 1

#Let Rp be the radius of the ring into which the given pixel
   belongs
Rp <- (R * cos((pi / (4 * N_side)) * Rd))

#Finding the number of pixels in the half-circle
N1 <- round((Tot_pixelR / (2)), 0)

#Let lp be the distance between two pixels in this interval
```

```
lp <- (1 / N1)

r <- 0.08
gamma <- (-(log(r) / log(N1)))
Pix_int <- as.integer(r / lp)
MemberP1 <- MemberP[(Pix_int + 1):((Pix_int) + N1)]
Int2 <- Int1[MemberP1]
df1 <- MemberP1[(Pix_int + 1):(N1 - (Pix_int))]
N2 <- length(df1)

HExp1De <- rep(0, N2)
l <- 1
for (CMB_pix in 1:N2) {
  N <- length(seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int
    , 1))
  A <- seq(df1[CMB_pix] - Pix_int, df1[CMB_pix] + Pix_int, 1)
  tot <- 0
  for (j in 1:(N - 2)) {
    Increment1 <- (((Int1[as.integer(A[j])]) - 2 * (Int1[as.
   integer(A[j + 1])]) + (Int1[as.integer(A[j + 2])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp1De[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N1)))
  l <- l + 1
}
HExp1De
min(HExp1De)
max(HExp1De)
max(HExp1De) - min(HExp1De)
mean(HExp1De)

#Figure 4.8(a)-This figure gives the plot of scaled CMB
    intensities of ring 4096
plot(x = MemberP1[721:3769], y = Int1[MemberP1[721:3769]], xlab
    = "CMB pixel", ylab = "I", type = "l", lwd = 1, col = "blue")

#Figure 4.8(b)-This figure gives the plot of H values of ring
    4096
plot(x = df1[394:3442], y = HExp1De[394:3442], xlab = "CMB pixel
    ", ylab = "H", type = "l", lwd = 1, col = "blue")

#The Holder exponent range which is constant = HExp1De[826:1470]
#The CMB pixel range in the original file = 25163208:25163852

#Table 4.6-This table gives the analysis of CMB intensities near
     the equator region
CMB_coord_abnormal <- CMB_coord[25163208:25163852, ]
min(cmbdf$I[25163208:25163852])
max(cmbdf$I[25163208:25163852])
max(cmbdf$I[25163208:25163852]) - min(Int1[25163208:25163852])
mean(cmbdf$I[25163208:25163852])
var(cmbdf$I[25163208:25163852])

min(HExp1De[826:1470])
max(HExp1De[826:1470])
max(HExp1De[826:1470]) - min(HExp1De[826:1470])
mean(HExp1De[826:1470])
```

```
#Finding the galactic coordinates of the unusual H range in one-
   dimensional space
#Finding the first pair of coordinates in the unusual H range
l1 <- 65.02
b1 <- (90 - 89.99)
#(l1,b1) = (65.02, 0.01)

#Finding the last pair of coordinates in the unusual H range
l2 <- 93.32
b2 <- (90 - 89.99)
#(l2,b2) = (93.32, 0.01)

#One-dimensional region excluding the region of unusual values
Remain_int <- c(df1[1:825], df1[1471:3442])
min(cmbdf$I[Remain_int])
max(cmbdf$I[Remain_int])
max(cmbdf$I[Remain_int]) - min(Int1[Remain_int])
mean(cmbdf$I[Remain_int])
var(cmbdf$I[Remain_int])

Remain_H <- c(HExp1De[1:825], HExp1De[1471:3442])
min(Remain_H)
max(Remain_H)
max(Remain_H) - min(Remain_H)
mean(Remain_H)
save.image(file = "HolderExponent1De1_2048.RData")
```

B.8 R code used to produce Figures 4.8c, 4.8d and results in Table 4.6

The R code in this section was used to produce Figures 4.8c, 4.8d and to obtain some results in Table 4.6 in Chapter 4. The code in this section visualizes scaled intensities (Figure 4.8c) and $\hat{H}(t)$ values (Figure 4.8d) of equator region. It was also used to obtain some results in the analysis of CMB intensities near the equatorial region (Table 4.6) as shown in Chapter 4.

```
library(rcosmo)
library(rgl)
library(akima)

#Generating CMB data frame with nside=2048
cmbdf <- CMBDataFrame("CMB_map_smica2048.fits")
df1 <- coords(cmbdf, new.coords = "cartesian")
#CMB_coord-coordinates of the real CMB data
CMB_coord <- data.frame(cbind(x = df1$x, y = df1$y, z = df1$z))

#Storing the intensities of pixels into a vector
Int1 <- cmbdf$I
Int1 <- Int1 / max(abs(Int1))
```

```
#k is changing with nside. If nside=2048=2^(11)=2^(k)
k <- 11

#d-dimension
d <- 2

#N_pix is the no. of pixels in which we compute the Holder
   exponent values
N_pix <- 1000

#For the equator region
CMB_row <- (23439718)
win <- CMBWindow(x = df1[CMB_row, ]$x, y = df1[CMB_row, ]$y, z =
     df1[CMB_row, ]$z, r = 0.23)
cmbdf11 <- window(cmbdf, new.window = win)
length(pix(cmbdf11))
min(cmbdf11$I)
max(cmbdf11$I)
avg1 <- mean(cmbdf11$I)
df11e <- coords(cmbdf11, new.coords = "cartesian")
df_sample1e <- df11e[seq(1, length(pix(cmbdf11)), by = (length(
    pix(cmbdf11)) / N_pix)), ]
df12e <- coords(df_sample1e, new.coords = "cartesian")
df13e <- data.frame(cbind(x = df12e$x, y = df12e$y, z=df12e$z))

r <- 0.01
N11e <- length(pix(cmbdf11))
N12e <- sqrt(N11e)
gamma <- (-(log((sqrt(pi) * r) / 2) / log((N12e))))
HExp2De <- rep(0, N_pix)
l <- 1
for (CMB_row in 1:N_pix) {
  win1 <- CMBWindow(x = df13e[CMB_row, ]$x, y = df13e[CMB_row, ]
   $y, z = df13e[CMB_row, ]$z, r = 0.01)
  cmbdf1 <- window(cmbdf, new.window = win1)
  N <- length(pix(cmbdf1))
  Pix_number <- as.integer(pix(cmbdf1))
  tot <- 0
  for (i in 1:N) {
    A <- neighbours(Pix_number[i], k)
    B <- neighbours(as.integer(A[1]), k)
    Increment1 <- (((Int1[as.integer(B[1])]) - (2 * Int1[as.
   integer(B[2])]) + (Int1[as.integer(B[3])]) -
      (2 * Int1[as.integer(B[4])]) + (Int1[as.integer(B[5])]) -
   (2 * Int1[as.integer(B[6])]) +
      (Int1[as.integer(B[7])]) - (2 * Int1[as.integer(B[8])]) +
   (4 * Int1[as.integer(B[9])]))^2)
    tot <- tot + Increment1
  }
  VNt <- tot
  HExp2De[l] <- (1 / 2) * ((d * (1 - gamma)) - (log(VNt) / log(
   N12e)))
  l <- l + 1
}
HExp2De
min(HExp2De)
max(HExp2De)
max(HExp2De) - min(HExp2De)
mean(HExp2De)
```

```
df14e <- data.frame(coords(df_sample1e, new.coords="spherical"))
n_interpolation <- 500

x <- df14e$theta
y <- df14e$phi
z <- HExp2De

spline_interpolated <- interp(x, y, z, xo = seq(min(x), max(x),
   length = n_interpolation), yo = seq(min(y), max(y), length =
   n_interpolation))

x.si <- spline_interpolated$x
y.si <- spline_interpolated$y
z.si <- spline_interpolated$z

nbcol <- 50
color <- rev(rainbow(nbcol, start = 0, end = 1))
zcol <- cut(z.si, nbcol)

#Figure 4.8(d)-This figure gives the plot of $\hat{H}(t)$ values
    from the equator region
persp3d(x.si, y.si, z.si, xlab = expression(theta), ylab =
   expression(varphi), zlab = "H", col = color[zcol])

df_sample2 <- df11e[seq(1, length(pix(cmbdf11)), by = (length(
   pix(cmbdf11)) / 100000)), ]
df15 <- data.frame(coords(df_sample2, new.coords = "spherical"))
n_interpolation <- 500

x <- df15$theta
y <- df15$phi
z <- df15$I / max(abs(df15$I))

spline_interpolated <- interp(x, y, z, xo = seq(min(x), max(x),
   length = n_interpolation), yo = seq(min(y), max(y), length =
   n_interpolation))

x.si <- spline_interpolated$x
y.si <- spline_interpolated$y
z.si <- spline_interpolated$z

nbcol <- 50
color <- rev(rainbow(nbcol, start = 0, end = 1))
zcol <- cut(z.si, nbcol)

#Figure 4.8(c)-This figure gives the plot of scaled CMB
   intensities of the equator region
persp3d(x.si, y.si, z.si, xlab = expression(theta), ylab =
   expression(varphi), zlab = "I", col = color[zcol])

df16 <- coords(cmbdf, new.coords = "spherical")
df17 <- data.frame(cbind(theta1 = df16$theta, phi1 = df16$phi))
equi_region <- data.frame(subset(df17, (theta1 >1.50 & theta1
   <1.60) & (phi1 >1.34 & phi1 <1.50), select = c(theta1, phi1)))

min(equi_region$phi1)
max(equi_region$phi1)
```

```r
min(equi_region$theta1)
max(equi_region$theta1)

#Finding the galactic coordinates of the unusual H range in two-
   dimensional space
#Finding the first pair of coordinates in the unusual H range of
    spherical surface,theta1=1.599772, phi1=1.499466
#Pix_value = 23404309
l1 <- 85.91
b1 <- (90 - 91.66)
#(l1,b1) = (85.91, -1.66)

#Finding the second pair of coordinates in the unusual H range
   of spherical surface, theta1=1.599772, phi1=1.340699
#Pix_value = 23391936
l2 <- 76.82
b2 <- (90 - 91.66)
#(l2,b2) = (76.82, -1.66)

#Finding the third pair of coordinates in the unusual H range of
    spherical surface, theta1=1.500099, phi1=1.340699
#Pix_value = 23564929
l3 <- 76.82
b3 <- (90 - 85.95)
#(l3,b3) = (76.82, 4.05)

#Finding the final pair of coordinates in the unusual H range of
    spherical surface, theta1=1.500099, phi1=1.499466
#Pix_value = 24158424
l4 <- 85.91
b4 <- (90 - 85.95)
#(l4,b4) = (85.91, 4.05)

#Table 4.6-This table gives the analysis of CMB intensities near
    the equator region
equidata_total <- cbind(coords(df_sample1e, new.coords = "
   spherical"), H = HExp2De)

#Analysis of two-dimensional region around unusual values
equidata_unusual <- data.frame(subset(equidata_total, (theta >
   1.50 & theta < 1.60) & (phi > 1.34 & phi < 1.50), select = c(
   theta, phi, I, H)))

min(equidata_unusual$I)
max(equidata_unusual$I)
max(equidata_unusual$I) - min(equidata_unusual$I)
avg_eq <- mean(equidata_unusual$I)
var_eq <- var(equidata_unusual$I)

min(equidata_unusual$H)
max(equidata_unusual$H)
max(equidata_unusual$H) - min(equidata_unusual$H)
mean(equidata_unusual$H)

#Analysis of two-dimensional region excluding the region of
   unusual values
win_eq_exterior <- CMBWindow(theta = c(1.599772, 1.599772,
   1.500099, 1.500099), phi = c(1.340699, 1.499466, 1.499466,
   1.340699), set.minus = TRUE)
```

```
CMB_row <- (23439718)
win_remain <- list(win_eq_exterior, CMBWindow(x = df1[CMB_row, ]
    $x, y = df1[CMB_row, ]$y, z = df1[CMB_row, ]$z, r = 0.23))
cmbdf_eqrem <- window(cmbdf, new.window = win_remain)
plot(cmbdf_eqrem)

equidata_remain <- data.frame(subset(equidata_total, !((theta >
    1.50 & theta < 1.60) & (phi > 1.34 & phi < 1.50)), select = c
    (theta, phi, I, H)))

min(equidata_remain$I)
max(equidata_remain$I)
max(equidata_remain$I) - min(equidata_remain$I)
avg_eqr <- mean(equidata_remain$I)
var_eqr <- var(equidata_remain$I)

min(equidata_remain$H)
max(equidata_remain$H)
max(equidata_remain$H) - min(equidata_remain$H)
mean(equidata_remain$H)
save.image(file = "H-2Deqregion2-smica2048.RData")
```

B.9 R code used to produce Figure 4.9

The R code in this section was used to produce Figure 4.9 in Chapter 4. The code in this section visualizes SMICA 2015 map with TMASK and the region of anomalies as shown in Chapter 4.

```
library(rcosmo)
library(rgl)

cmbdf <- CMBDataFrame("CMB_map_smica2048.fits", include.masks =
    TRUE)
cmbdf$TMASK1 <- (1 - cmbdf$TMASK) + cmbdf$I

#Figure 4.9(a)-This figure gives the plot of non-inpainted
    Planck 2015 CMB map with the anomalous sky window
plot(cmbdf, intensities = "TMASK1", back.col = "white", ylab = "
    ", xlab = "", zlab = "")
win_eq <- CMBWindow(theta = c(1.599772, 1.599772, 1.500099,
    1.500099), phi = c(1.340699, 1.499466, 1.499466, 1.340699))
plot(win_eq, col = "white", lwd = 3)

cmbdf_eq <- window(cmbdf, new.window = win_eq)
#Figure 4.9(b)-This figure gives the plot of the enlarged
    anomalous sky window
plot(cmbdf_eq, back.col = "white", ylab="", xlab="", zlab="")
```

B.10 R code used to produce Figure 4.10

The R code in this section was used to produce Figure 4.10 in Chapter 4. The code in this section visualizes discrepancy maps using the Hölder exponent approach (Figure 4.10a) and AC discrepancy approach (Figure 4.10b) for CMB intensities from SMICA 2015 as shown in Chapter 4.

```
#Figure 4.10(b)
library(rcosmo)
library(R.matlab)
library(rgl)

CMB_probe_smica<-readMat("ACDSMICA2015_L1500_HL1024_lag10.mat")
ACD_smica <- CMB_probe_smica[[1]]
cmbdf <- CMBDataFrame(nside = 1024, I = ACD_smica[1, ], ordering
    = "ring")
cmbdf$ACD1 <- ifelse(cmbdf$I > quantile(cmbdf$I, 0.95), 1, -1)

#Figure 4.10(b)-This figure gives the plot of the AC discrepancy
    map from SMICA 2015
plot(cmbdf, intensities = "ACD1", back.col = "white", ylab = "",
    xlab = "", zlab = "")
win_eq <- CMBWindow(theta = c(1.599772, 1.599772, 1.500099,
    1.500099), phi = c(1.340699, 1.499466, 1.499466, 1.340699))
plot(win_eq, col = "white", lwd = 3)

um <- matrix(c(-0.98451775, 0.17489178, -0.01173002, 0,
    -0.02096882, -0.05107139, 0.99847484, 0, 0.17402618,
    0.98326194, 0.05394801, 0, 0, 0, 0, 1), byrow = TRUE, nrow =
    4, ncol = 4)
view3d(userMatrix = um)
rgl.snapshot("Figure410b.png")

#Figure 10(a)
library(rcosmo)
library(rgl)
library(RcppRoll)

#Generating CMB data frame with nside=2048
cmbdf <- CMBDataFrame("CMB_map_smica2048.fits", ordering="ring")

L <- length(cmbdf$I)
Int1 <- cmbdf$I
Scale <- max(abs(Int1 - mean(Int1)))
Int1 <- Int1 / max(abs(Int1))

N_side <- 2048
d <- 1
N1 <- 8192 / 2
#Let lp be the distance between two pixels in this interval
lp <- (1 / N1)

Pix_int <- 30
r <- Pix_int * lp
gamma <- (-(log(r) / log(N1)))
```

```r
Increment1 <- rep(0, L - 2)
Increment1 <- (Int1[1:(L-2)] - 2 * Int1[2:(L-1)] + Int1[3:L])^2

anomshift <- roll_sum(Increment1, 2 * Pix_int + 1)
anom <- rep(mean(anomshift), L)
anom <- rep(1, L)
anom[(Pix_int + 1):(L - 2 - (Pix_int))] <- anomshift

HExp1De <- rep(0, L)
HExp1De <- (1 / 2) * ((d * (1 - gamma)) - (log(anom) / log(N1)))
cmbdf$H <- HExp1De - mean(HExp1De)
cmbdf$H1 <- ifelse(cmbdf$H < quantile(cmbdf$H, 0.05), 1, -1)

#Figure 4.10(a)-This figure gives the plot of the Holder
    exponent map from SMICA 2015
plot(cmbdf, intensities = "H1", back.col = "white", ylab = "",
    xlab = "", zlab = "")
win_eq <- CMBWindow(theta = c(1.599772, 1.599772, 1.500099,
    1.500099), phi = c(1.340699, 1.499466, 1.499466, 1.340699))
plot(win_eq, col = "white", lwd = 3)

um <- matrix(c(-0.98451775, 0.17489178, -0.01173002, 0,
    -0.02096882, -0.05107139, 0.99847484, 0, 0.17402618,
    0.98326194, 0.05394801, 0, 0, 0, 0, 1), byrow = TRUE, nrow =
    4, ncol = 4)
view3d(userMatrix = um)
rgl.snapshot("Figure410a.png")
```

B.11 R code used to produce Figure 4.11

The R code in this section was used to produce Figure 4.11 in Chapter 4. The code in this section visualizes $\hat{H}_{\Delta}$ discrepancy maps for CMB intensities from SMICA 2015 as shown in Chapter 4.

```r
library(rcosmo)
library(rgl)
library(RcppRoll)

#Figure 4.11(a)
cmbdf$Had <- rep(0, L)

cmbdf$Had[1:(L - 20)] <- pmin(
  abs(cmbdf$H[1:(L - 20)] - cmbdf$H[11:(L - 10)]), abs(cmbdf$H
   [1:(L - 20)] - cmbdf$H[12:(L - 9)]), abs(cmbdf$H[1:(L - 20)]
  - cmbdf$H[13:(L - 8)]),
  abs(cmbdf$H[1:(L - 20)] - cmbdf$H[14:(L - 7)]),
  abs(cmbdf$H[1:(L - 20)] - cmbdf$H[15:(L - 6)]), abs(cmbdf$H
   [1:(L - 20)] - cmbdf$H[16:(L - 5)]),
  abs(cmbdf$H[1:(L - 20)] - cmbdf$H[17:(L - 4)]),
  abs(cmbdf$H[1:(L - 20)] - cmbdf$H[18:(L - 3)]), abs(cmbdf$H
   [1:(L - 20)] - cmbdf$H[19:(L - 2)]),
  abs(cmbdf$H[1:(L - 20)] - cmbdf$H[20:(L - 1)]),
  abs(cmbdf$H[1:(L - 20)] - cmbdf$H[21:(L)]))
```

```
cmbdf$H2 <- ifelse(cmbdf$Had > quantile(cmbdf$Had, 0.95), 1, -1)

#Figure 4.11(a)-This figure gives the plot of the ${\hat{H}}_{\
    Delta}$ discrepancy map from SMICA 2015
plot(cmbdf, intensities = "H2", back.col = "white", ylab = "",
    xlab = "", zlab = "")
win_eq <- CMBWindow(theta = c(1.599772, 1.599772, 1.500099,
    1.500099), phi = c(1.340699, 1.499466, 1.499466, 1.340699))
plot(win_eq, col = "white", lwd = 3)
view3d(userMatrix = um)
rgl.snapshot("Figure411a.png")

#Figure 4.11(b)
cmbdf1 <- CMBDataFrame("CMB_map_smica2048.fits", include.masks =
     TRUE, ordering = "ring")
cmbdf$H3 <- cmbdf$H2 * pmax(cmbdf1$TMASK, cmbdf$H2)

#Figure 4.11(b)-This figure gives the plot of the ${\hat{H}}_{\
    Delta}$ discrepancies over TMASK from SMICA 2015
plot(cmbdf, intensities = "H3", back.col = "white", ylab = "",
    xlab = "", zlab = "")
win_eq <- CMBWindow(theta = c(1.599772, 1.599772, 1.500099,
    1.500099), phi = c(1.340699, 1.499466, 1.499466, 1.340699))
plot(win_eq, col = "black", lwd = 3)
view3d(userMatrix = um)
rgl.snapshot("Figure411b.png")
```

Appendix C

Codes used to produce figures and simulations in Chapter 5

C.1 R code used to produce Figure 5.1

The R code in this section was used to produce Figure 5.1 in Chapter 5. The code in this section visualizes a realization of a cyclic long-memory time series (Figure 5.1a), plot of its periodogram (Figure 5.1b), sample covariance function (Figure 5.1c) and the wavelet coefficients (Figure 5.1d) as shown in Chapter 5.

```
pck <- c("waveslim", "wmtsa", "latex2exp")
lapply(pck, library, character.only = TRUE)

set.seed(654321)

Ts <- 1000
n0 <- min(floor(Ts/10), 100)
n <- Ts + 2 * n0

#The function dwpt.sim simulates a seasonal persistent process (
    cyclic long-memory time series) using the discrete wavelet
    packet transform (DWPT)
ts1 <- dwpt.sim(n, "mb16", 0.4, 0.1, epsilon = 0.001)

#Figure 5.1(a)-This figure gives the plot of a realization of a
    cyclic long-memory time series
plot(ts1, type = "l", xlab = "Time", ylab = "Value")

#Figure 5.1(b)-This figure gives the plot of the periodogram
    corresponding to the simulated cyclic long-memory time series
plot(0:(n/2)/n, per(ts1), type = "l", xlab = "Frequency", ylab =
     "Value")

#Figure 5.1(c)-This figure gives the plot of the sample
    covariance function
acf(ts1, lag.max = 100, ylim = c(-0.6, 1), main = " ")
```

```
#The function wavCWT() computes the continuous wavelet transform
    of a time series
ts1.cwt <- wavCWT(ts1)

#Figure 5.1(d) - This figure gives the plot of the wavelet
    coefficients
plot(ts1.cwt)
```

C.2 Maple codes used in Section 5.6 of Chapter 5

The following code gives the Shannon father wavelet and the plot of the Shannon father wavelet respectively.

```
sfw := piecewise(t \neq 0, sin(Pi*t)/(Pi*t), t = 0, 1)

plots[display](plot(sfw, t = -2*Pi .. 2*Pi), plottools[line]([0,
    0], [0, 1]), color = red)
```

The following code gives the Meyer father wavelet and the plot of the Meyer father wavelet respectively.

```
Psi_fmf := piecewise(abs(lambda) <= (2*Pi)/3, 1, (2*Pi)/3 <= abs
    (lambda) and abs(lambda) <= (4*Pi)/3, cos(Pi/2*((3*abs(lambda
    ))/(2*Pi) - 1)), 0)
plot(Psi_fmf)
```

The following code was used to evaluate the integrals in Example 5.2 of Chapter 5.

```
int(1, lambda = (-2*Pi)/3 .. (2*Pi)/3) + 2*int(abs(cos(Pi/2*((3*
    abs(lambda))/(2*Pi) - 1)))^2, lambda = (2*Pi)/3 .. (4*Pi)/3)

int(2*lambda^2, lambda = (-2*Pi)/3 .. (2*Pi)/3) + 2*int(2*lambda
    ^2*abs(cos(Pi/2*((3*abs(lambda))/(2*Pi) - 1)))^2, lambda = (2
    *Pi)/3 .. (4*Pi)/3)

int(1, lambda = (-2*Pi)/3 .. (2*Pi)/3) + 2*int(abs(cos(Pi/2*((3*
    abs(lambda))/(2*Pi) - 1)))^4, lambda = (2*Pi)/3 .. (4*Pi)/3)
```

The following code gives the plot of the Mexican hat wavelet and the plot of the Fourier transform of the Mexican hat wavelet respectively.

```
plot(2*(-t^2 + 1)*exp(-t^2/2)/(Pi^0.25*sqrt(3)), t = -6 .. 6)
plot(sqrt(8)*Pi^0.25*lambda^2*exp(-lambda^2/2)/sqrt(3), lambda =
    -6 .. 6)
```

The following code evaluates the values of the constants c2 and c3.

```
c2 := int(abs(sqrt(8)*Pi^(-0.25)*lambda^2*exp(-lambda^2/2)/sqrt
    (3))^2, lambda = -10 .. 10)
c3 := int(2.*lambda^2*abs(sqrt(8)*Pi^(-0.25)*lambda^2*exp(-
    lambda^2/2)/sqrt(3))^2, lambda = -10 .. 10)
```

C.3 R code used to produce Figure 5.3

The R code in this section was used to produce Figure 5.3 in Chapter 5. The code in this section visualizes the plot of the asymptotic correlation of $\hat{s_0}$ and $\hat{\alpha}$ based on the Shannon father wavelet case (Figure 5.3a) and the plot of the asymptotic correlation of $\hat{s_0}$ and $\hat{\alpha}$ based on the Meyer father wavelet case (Figure 5.3b) respectively.

```
#Figure 5.3(a)
library(rgl)
library(latex2exp)

#Let s be the seasonality parameter
s <- seq(1, 2, length.out = 100)

#Let a be the long-memory parameter
a <- seq(0.01, 1/2, length.out = 100)

#This function computes the correlation of the components of the
     asymptotic vector
zcoor <- function(s, a) {
     ((((1 - (4 * a * log(s))) * a * (4 * a + 2) * s^(-1))/(4 * (
    pi^2))) - ((18 * a * s^3 * log(s))/(pi^4)))/(sqrt(((((1 - (4
    * a * log(s)))^2)/(4 * (pi^2))) + ((18 * (s^4) * (log(s))^2)/
    (pi^4))) * ((((a^2) * ((4 * a + 2)^2) * (s^(-2)))/(4 * (pi^2)
    )) + ((18 * (a^2) * (s^2))/(pi^4)))))
}

rho <- outer(s, a, zcoor)

#This jet.colors function interpolates 'red' and 'yellow' colors
     and creates a new color palette
jet.colors <- colorRampPalette(c("red", "yellow"))
pal <- jet.colors(100)
col.ind <- cut(rho, 100)

um <- matrix(c(-0.5988361, 0.8008499, 0.005933772, 0,
    -0.1430069, -0.1142176, 0.983109236, 0, 0.7880003, 0.5878722,
     0.182924747, 0, 0, 0, 0, 1), 4, 4, byrow = TRUE)
view3d(userMatrix = um)

#Figure 5.3(a) - This figure gives the plot of the asymptotic
    correlation of hat{s_{0}} and hat{\alpha} based on the
    Shannon father wavelet
persp3d(s, a, rho, col = pal[col.ind], box = FALSE, xlab = " ",
    ylab = " ", zlab = " ")
```

```
#These commands add axes labels of {s_{0}}, {\alpha} and {\rho}
   to their corresponding axes
mtext3d(TeX("s_0"), "x−+", line = 3)
mtext3d(TeX("$\\alpha$"), "y+−", line = 3)
mtext3d(TeX("$\\rho$"), "z+−", line = 3)

#This gives Figure 5.3(a) in the PNG format.
rgl.snapshot("Fig53a.png")

#Figure 5.3(b)
library(rgl)
library(latex2exp)

#Let s be the seasonality parameter
s <− seq(1, 2, length.out = 100)

#Let a be the long−memory parameter
a <− seq(0, 1/2, length.out = 100)

#This function computes the correlation of the components of the
    asymptotic vector
zcoor <− function(s, a) {
    (((((1 − (4 * a * log(s))) * a * (4 * a + 2) * (s^(−1)))/(4
   * (pi^2))) − ((8 * a * s^3 * log(s))/(((8 * pi * (pi^2 − 2))/
   9)^2)))/(sqrt(((((1 − (4 * a * log(s)))^2)/(4 * pi^2)) + ((8
   * s^4 * (log(s))^2)/(((8 * pi * (pi^2 − 2))/9)^2))) * (((a^2
   * ((4 * a + 2)^2) * (s^(−2)))/(4 * pi^2)) + ((8 * a^2 * s^2)/
   (((8 * pi * (pi^2 − 2))/9)^2))))))
}

rho <− outer(s, a, zcoor)

#This jet.colors function interpolates 'red' and 'yellow' colors
    and creates a new color palette
jet.colors <− colorRampPalette(c("red", "yellow"))
pal <− jet.colors(100)
col.ind <− cut(rho, 100)

um <− matrix(c(−0.5988361, 0.8008499, 0.005933772, 0,
   −0.1430069, −0.1142176, 0.983109236, 0, 0.7880003, 0.5878722,
    0.182924747, 0, 0, 0, 0, 1), 4, 4, byrow = TRUE)
view3d(userMatrix = um)

#Figure 5.3(b) − This figure gives the plot of the asymptotic
   correlation of hat{s_{0}} and hat{\alpha} based on the Meyer
   father wavelet
persp3d(s, a, rho, col = pal[col.ind], xlab = " ", box = FALSE,
   ylab = " ", zlab = " ")

#These commands add axes labels of {s_{0}}, {\alpha} and {\rho}
   to their corresponding axes
mtext3d(TeX("s_0"), "x−+", line = 3)
mtext3d(TeX("$\\alpha$"), "y+−", line = 3)
mtext3d(TeX("$\\rho$"), "z+−", line = 3)

#This gives Figure 5.3(b) in the PNG format
rgl.snapshot("Fig53b.png")
```

C.4 R code used to produce Figure 5.4

The R code in this section was used to obtain Figure 5.4 in Chapter 5. The code in this section visualizes the Q-Q plots of the first two normalised statistics S_1 (Figure 5.4a) and S_2 (Figure 5.4b) for $j = 7$, the plot of the density ellipsoids of (S_1, S_2) (Figure 5.4c) and the plot of the density ellipsoids and realizations of the random vector $\sqrt{m_j}\left(\widehat{(s_0,\alpha)}_j - (s_0,\alpha)\right)$ (Figure 5.4d).

```
if (!requireNamespace("BiocManager", quietly = TRUE)) {
  install.packages("BiocManager")
}
BiocManager::install("MassSpecWavelet")
pck <- c("polynom", "orthopolynom", "pracma", "MassSpecWavelet",
    "qboxplot", "latex2exp", "RcppRoll", "ggplot2", "dplyr", "
   ggpubr")
lapply(pck, library, character.only = TRUE)
set.seed(8654321)

M <- 10
k <- 100
fN <- 7

#Timing code
ptm <- proc.time()

delta1_mat1 <- matrix(ncol = fN, nrow = k * M)
delta2_mat1 <- matrix(ncol = fN - 1, nrow = k * M)

c2 <- 6.283185
c3 <- 31.41593

for (m in 1:M) {
    print(m)
    #Values used for simulations
    Ts <- 1e+07
    kt <- 100
    k <- 100
    n0 <- 100

    #Parameters of Gegenbauer time series
    u1 <- 0.3
    d1 <- 0.1
    alpha <- d1
    s0 <- acos(u1)

    #Multiplier to normalize h(0)=1
    A <- sqrt((2 * pi * 4^(2 * alpha) * (sin(s0/2))^(4 * alpha))
   /(s0^(4 * alpha)))

    #Gegenbauer polynomial
    gcoef <- function(i) {
        polynomial.values(gegenbauer.polynomials(i - 1, d1), u1)
   [[i]]
    }
```

```
gcoef <- Vectorize(gcoef)
coef <- seq(n0 + 1)
coef <- gcoef(coef)
coef1 <- rev(coef)
n <- Ts/kt + 200 + n0

#Creating epsilon:
epsil0 <- rnorm(n, mean = 0, sd = 1)

#Simulating Gengenbauer random process
ts1 <- seq(Ts/kt + 200)
ts1 <- A * roll_sum(epsil0, weights = coef1, normalize =
FALSE, align = "left")

#Range of scales and the number of scales
ns <- 11

#Calculating the continuous wavelet transform (CWT) of
Gengenbauer random process
ts1.cwt <- cwt(ts1, scales = seq(1, 11, 1), wavelet = "mexh"
)
a1 <- as.matrix(ts1.cwt[101:(Ts/kt + 100)])
aj <- seq(1, 11, 1)
squ_aj <- (aj)^(-2)

#Calculating aj^(-2)- aj+1^(-2)
denominator <- squ_aj[1:(fN - 1)] - squ_aj[2:(fN)]

delta1_mat <- matrix(ncol = fN, nrow = k)
delta2j <- rep(0, fN)
delta2_mat <- matrix(ncol = fN - 1, nrow = k)

ts1cwtmatr <- matrix(0, Ts - 100, 11)
#Creating matrix with wavelet coefficients
ts10 <- seq(Ts/kt + 100)
for (N in 1:k) {
    epsil <- rnorm(Ts/kt + 100 + n0, mean = 0, sd = 1)
    ts10 <- A * roll_sum(epsil, weights = coef1, normalize =
FALSE)
    ts10.cwt <- as.matrix(cwt(ts10, scales = seq(1, 11, 1),
wavelet = "mexh"))
    ts1cwtmatr[1:(Ts/kt - 100), ] <- ts10.cwt[101:(Ts/kt), ]
    ts1 <- ts10
    for (j in 1:(kt - 1)) {
        epsil <- c(epsil[(length(epsil) - n0 + 1):length(
epsil)], rnorm(Ts/kt, mean = 0, sd = 1))
        ts10 <- A * roll_sum(epsil, weights = coef1,
normalize = FALSE)
        ts1 <- c(ts1[(length(ts1) - 200 + 1):length(ts1)],
ts10[1:(Ts/kt)])
        ts10.cwt <- as.matrix(cwt(ts1, scales = seq(1, 11,
1), wavelet = "mexh"))
        ts1cwtmatr[(Ts/kt - 100 + 1 + (j - 1) * Ts/kt):(Ts/
kt - 100 + j * Ts/kt), ] <- ts10.cwt[101:(Ts/kt + 100), ]
    }
    ts1cwtmatr <- (ts1cwtmatr * ts1cwtmatr)

    #Calculating deltaj^(2)- deltaj+1^(2)
    for (j in 1:(fN)) {
```

```
            delta2j[j] <- round(mean(ts1cwtmatr[1:min(Ts - 100,
    200 * (j < 3) + (aj[j])^9), j]), 5)
          }
          delta1_mat[N, ] <- delta2j
          numerator <- delta2j[1:(fN - 1)] - delta2j[2:fN]

          #Calculating Delta(deltaj.^(2))
          delat_result <- numerator/denominator
          delta2_mat[N, ] <- delat_result

          delta1_mat1[(1 + k * (m - 1)):(m * k), ] <- delta1_mat
          delta2_mat1[(1 + k * (m - 1)):(m * k), ] <- delta2_mat
      }
      rm(list = setdiff(ls(), c("ptm", "m", "M", "k", "fN", "c2",
    "c3", "delta1_mat1", "delta2_mat1", "s_mat1", "alpha_mat1")))
}

timetaken1 <- (proc.time() - ptm)/60
save.image(file = "Longmemory_estimates1.RData")
load("Longmemory_estimates1.RData")

u1 <- 0.3
d1 <- 0.1
alpha <- d1
s0 <- acos(u1)
delta1_mat <- delta1_mat1
delta2_mat <- delta2_mat1

#Boxplots of bar(delta)j^(2), bar(Delta(deltaj.^(2))), s^hat_0j
    and alpha^hat_j
q1 <- 1

#Boxplot of bar(delta)j^(2) for j=1,2,3,4,5,6,7
qboxplot(data.frame(delta1_mat[, q1:(fN)]), probs = c(0.25, 0.5,
     0.75), range = 0, col = "bisque", xaxt = "n", medcol = "
    bisque", main = TeX("Boxplot of $\\bar{\\delta}_{j\\cdot
    }^{(2)}$"))
m1 <- apply(delta1_mat[, q1:(fN)], 2, mean)
points((q1:(fN)), m1, col = "red", pch = 18)
axis(1, at = (q1:(fN)), labels = (q1:(fN)))
abline(h = c2 * s0^(-4 * alpha), col = "red", lty = 3)

#Boxplot of bar(Delta(deltaj.^(2))) for j=1,2,3,4,5,6
qboxplot(data.frame(delta2_mat[, q1:(fN - 1)]), probs = c(0.25,
    0.5, 0.75), range = 0, col = "bisque", xaxt = "n", medcol = "
    bisque", main = TeX("Boxplot of $\\Delta \\bar{\\delta}_{j\\
    cdot}^{(2)}$"))
m1 <- apply(delta2_mat[, q1:(fN - 1)], 2, mean)
points((q1:(fN - 1)), m1, col = "red", pch = 18)
axis(1, at = (q1:(fN - 1)), labels = (q1:(fN - 1)))
abline(h = c3 * alpha * s0^(-4 * alpha - 2), col = "red", lty =
    3)

Ts <- 1e+07

shapiro.test(sqrt(Ts - 100) * ((delta1_mat1[, (fN)])/c2 - mean((
    delta1_mat1[, (fN)])/c2)))
shapiro.test(sqrt(Ts - 100) * ((delta2_mat1[, (fN - 1)])/c3 -
    mean((delta2_mat1[, (fN - 1)])/c3)))
```

```
#Figure 5.4(a) - This figure gives the Q-Q plot of S1
ggqqplot(sqrt(Ts - 100) * ((delta1_mat1[, (fN)])/c2 - mean((
   delta1_mat1[, (fN)])/c2)))

#Figure 5.4(b) - This figure gives the Q-Q plot of S2
ggqqplot(sqrt(Ts - 100) * ((delta2_mat1[, (fN - 1)])/c3 - mean((
   delta2_mat1[, (fN - 1)])/c3)))

p <- cbind((delta1_mat1[, (fN - 3)])/c2, (delta2_mat1[, (fN - 1)
   ])/c3)

sum(p[, 2] > 0)
sum(p[, 2] < 0.5 * p[, 1]^2)
sum(p[, 1] > 0)
sum(p[, 1] < 1)

p <- p[p[, 2] > 0 & p[, 2] < 0.5 * p[, 1]^2, ]

m10 <- mean(p[, 1])
m20 <- mean(p[, 2])

pnorm <- cbind(sqrt(Ts - 100) * (p[, 1] - m10), sqrt(Ts - 100) *
    (p[, 2] - m20))

sigma <- cov(pnorm)

#Computing the correlation matrix
cor(pnorm)

sigma.inv <- solve(sigma, matrix(c(1, 0, 0, 1), 2, 2))

ellipse <- function(s, t) {
    u <- c(s, t) - c(0, 0)
    u %*% sigma.inv %*% u/2
}

#Figure 5.4(c) - This figure gives the density ellipsoid of (S1,
   S2)
plot(pnorm, pch = 20, xlab = TeX("$S_1$"), ylab = TeX("$S_2$"))
points(0, 0, col = "red", lty = 2, pch = 18, cex = 3)

n <- 200
x <- seq(min(pnorm[, 1]), max(pnorm[, 1]), length.out = n)
y <- seq(min(pnorm[, 2]), max(pnorm[, 2]), length.out = n)
z <- mapply(ellipse, as.vector(rep(x, n)), as.vector(outer(rep
   (0, n), y, `+`)))
contour(x, y, matrix(z, n, n), levels = (0:10), col = terrain.
   colors(11), add = TRUE)

q <- p[, 2]/p[, 1]

hat_s0j <- exp(0.5 * lambertWp(log(1/p[, 1])/(2 * q)))
str(hat_s0j)

hat_alphaj <- q * (exp(lambertWp(log(1/p[, 1])/(2 * q))))
str(hat_alphaj)

m30 <- mean(hat_s0j)
```

```r
m40 <- mean(hat_alphaj)

pnorm1 <- cbind(sqrt(Ts - 100) * (hat_s0j - m30), sqrt(Ts - 100)
    * (hat_alphaj - m40))

sigma1 <- cov(pnorm1)
sigma.inv1 <- solve(sigma1, matrix(c(1, 0, 0, 1), 2, 2))

ellipse <- function(s, t) {
    u <- c(s, t) - c(0, 0)
    u %*% sigma.inv1 %*% u/2
}

#Figure 5.4(d) - This figure gives the density ellipsoid of sqrt
   {m_j}((hat{s_0,alpha}_j)_(s_0,alpha))
plot(pnorm1, pch = 20, xlab = TeX("$\\sqrt{m_j}(\\hat{s}_{0,j}-s
   _0)$"), ylab = TeX("$\\sqrt{m_j}(\\hat{\\alpha}_j-\\alpha)$")
   , mgp = c(2.6, 1, 0))
points(0, 0, col = "red", lty = 2, pch = 18, cex = 3)

n <- 200
x <- seq(min(pnorm1[, 1]), max(pnorm1[, 1]), length.out = n)
y <- seq(min(pnorm1[, 2]), max(pnorm1[, 2]), length.out = n)
z <- mapply(ellipse, as.vector(rep(x, n)), as.vector(outer(rep
   (0, n), y, `+`)))
contour(x, y, matrix(z, n, n), levels = (0:10), col = terrain.
   colors(11), add = TRUE)
```

Appendix D

Parallel computing using Gadi of NCI

This appendix consists of the parallelized version of the R code which was used to obtain Figure 5.4. The serial version of the code was parallelized using the R packages FOREACH (Microsoft and Weston (2020)), PARALLEL (R Core Team (2020)) and DOPARALLEL (Microsoft Corporation and Weston (2020)). Parallel computing was done using the super computer Gadi of National Computational Infrastructure (NCI). NCI is the major high performance computing facility of Australia that has the largest highly-integrated super computer and physically located in the Australian National University, Canberra. Gadi is the latest most powerful super computer of NCI launched in November 2019, which facilitates parallel computing. Numerous simulation studies were carried out using Gadi by setting the cluster size, ncpus (number of central processing units) $= 1, 2, ..., 20$ to observe the improvement in the execution time of the code. The speedup coefficients were calculated using the formula speedup = serial execution time/parallel execution time. The Table D.1 shows a summary of the carried out simulations by increasing the ncpus.

Figure D.1 depicts the variation in the speedup coefficients with the increasing number of cpus and it confirms that it follows Amdahl's law (Gustafson (1988)). That is, the speedup coefficient increases with the increasing number of cpus up to 10 and then it stays approximately constant. Therefore, for simulations related to results in Figure 5.4, it is enough to use ncpus=10.

No. of cpus	**Elapsed time (in hours)**	**Elapsed time (in seconds)**	**Speedup**
1	07:03:55	$2.5435 \cdot 10^4$	1.00
2	03:33:06	$1.2786 \cdot 10^4$	1.99
3	02:51:56	$1.0316 \cdot 10^4$	2.47
4	02:21:36	$8.4960 \cdot 10^3$	2.99
5	01:29:46	$5.3860 \cdot 10^3$	4.72
6	01:28:14	$5.2940 \cdot 10^3$	4.80
7	01:32:12	$5.5320 \cdot 10^3$	4.60
8	01:29:08	$5.3480 \cdot 10^3$	4.76
9	01:26:15	$5.1750 \cdot 10^3$	4.91
10	00:47:10	$2.8300 \cdot 10^3$	8.99
11	00:47:34	$2.8540 \cdot 10^3$	8.91
12	00:48:54	$2.9340 \cdot 10^3$	8.67
13	00:43:45	$2.6250 \cdot 10^3$	9.69
14	00:46:57	$2.8170 \cdot 10^3$	9.03
15	00:45:12	$2.7120 \cdot 10^3$	9.38
16	00:46:43	$2.8030 \cdot 10^3$	9.07
17	00:45:50	$2.7500 \cdot 10^3$	9.25
18	00:44:23	$2.6630 \cdot 10^3$	9.55
19	00:44:06	$2.6460 \cdot 10^3$	9.61
20	00:43:44	$2.6240 \cdot 10^3$	9.69

Table D.1: Analysis of execution time over the number of cpus

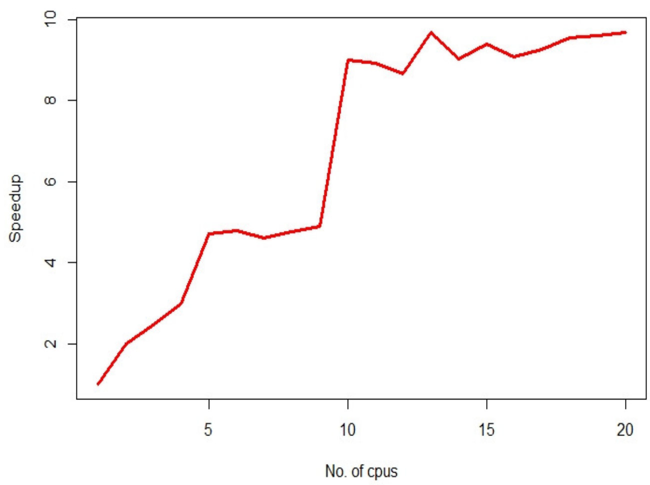


Figure D.1: Improvement in the speed of execution with the number of cpus

```
#Using parallel computation to run the loop where M=1,...,10.
pck <- c("polynom", "orthopolynom", "pracma", "MassSpecWavelet",
    "qboxplot", "latex2exp", "RcppRoll", "foreach", "parallel", "
    doParallel")
lapply(pck, library, character.only = TRUE)

M <- 10
k <- 100
```

```
fN <- 7

#Timing code
ptm <- proc.time()

delta1_mat1 <- matrix(ncol = fN, nrow = k * M)
delta2_mat1 <- matrix(ncol = fN - 1, nrow = k * M)

c2 <- 6.283185
c3 <- 31.41593

#Calculating the no. of cores
numCores <- detectCores()

#Creating a cluster of 10
cl <- makeCluster(10)
doParallel::registerDoParallel(cl)
set.seed(8654321, kind = "L'Ecuyer-CMRG")

#Loading the libraries to each cluster
clusterEvalQ(cl, pck <- c("polynom", "orthopolynom", "pracma", "
   MassSpecWavelet", "qboxplot", "latex2exp", "RcppRoll", "
   foreach", "parallel", "doParallel"))
clusterEvalQ(cl, lapply(pck, library, character.only = TRUE))

#Parameters of Gegenbauer time series
u1 <- 0.3
d1 <- 0.1
alpha <- d1
s0 <- acos(u1)

#Storing the parameter values in each cluster
clusterEvalQ(cl, {
    u1 <- 0.3
    d1 <- 0.1
    alpha <- d1
    s0 <- acos(u1)
})

A <- sqrt((2 * pi * 4^(2 * alpha) * (sin(s0/2))^(4 * alpha))/(s0
   ^(4 * alpha)))
clusterEvalQ(cl, A <- sqrt((2 * pi * 4^(2 * alpha) * (sin(s0/2))
   ^(4 *alpha))/(s0^(4* alpha))))

#Gegenbauer polynomial
clusterEvalQ(cl, gcoef <- function(i) {
    polynomial.values(gegenbauer.polynomials(i-1, d1), u1)[[i]]
})

#Evaluating the function on each cluster
clusterEvalQ(cl, gcoef <- Vectorize(gcoef))

out <- foreach(m = 1:M, .combine = rbind) %dopar% {
    print(m)
    #Values used for simulations
    Ts <- 1e+07
    kt <- 100
    k <- 100
    n0 <- 100
```

```
 A <- sqrt((2 * pi * 4^(2 * alpha) * (sin(s0/2))^(4 * alpha))
/(s0^(4 * alpha)))

 gcoef <- function(i) {
     polynomial.values(gegenbauer.polynomials(i - 1, d1), u1)
[[i]]
 }
 gcoef <- Vectorize(gcoef)
 coef <- seq(n0 + 1)
 coef <- gcoef(coef)
 coef1 <- rev(coef)
 n <- Ts/kt + 200 + n0

 #Creating epsilon
 epsil0 <- rnorm(n, mean = 0, sd = 1)

 #Simulating Gengenbauer random process
 ts1 <- seq(Ts/kt + 200)
 ts1 <- A * roll_sum(epsil0, weights = coef1, normalize =
FALSE, align = "left")
 ns <- 11

 #Calculating the continuous wavelet transform (CWT) of
Gengenbauer random process
 ts1.cwt <- cwt(ts1, scales = seq(1, 11, 1), wavelet = "mexh"
)
 a1 <- as.matrix(ts1.cwt[101:(Ts/kt + 100)])
 aj <- seq(1, 11, 1)
 squ_aj <- (aj)^(-2)

 denominator <- squ_aj[1:(fN - 1)] - squ_aj[2:(fN)]
 delta1_mat <- matrix(ncol = fN, nrow = k)
 delta2j <- rep(0, fN)
 delta2_mat <- matrix(ncol = fN - 1, nrow = k)

 ts1cwtmatr <- matrix(0, Ts - 100, 11)
 ts10 <- seq(Ts/kt + 100)
 for (N in 1:k) {
     epsil <- rnorm(Ts/kt + 100 + n0, mean = 0, sd = 1)
     ts10 <- A * roll_sum(epsil, weights = coef1, normalize =
 FALSE)
     ts10.cwt <- as.matrix(cwt(ts10, scales = seq(1, 11, 1),
wavelet = "mexh"))
     ts1cwtmatr[1:(Ts/kt - 100), ] <- ts10.cwt[101:(Ts/kt), ]
     ts1 <- ts10

     for (j in 1:(kt - 1)) {
          epsil <- c(epsil[(length(epsil) - n0 + 1):length(
epsil)], rnorm(Ts/kt, mean = 0, sd = 1))
          ts10 <- A * roll_sum(epsil, weights = coef1,
normalize = FALSE)
          ts1 <- c(ts1[(length(ts1) - 200 + 1):length(ts1)],
ts10[1:(Ts/kt)])
          ts10.cwt <- as.matrix(cwt(ts1, scales = seq(1, 11,
1), wavelet = "mexh"))
          ts1cwtmatr[(Ts/kt - 100 + 1 + (j - 1) * Ts/kt):(Ts/
kt - 100 + j * Ts/kt), ] <- ts10.cwt[101:(Ts/kt + 100), ]
     }
     ts1cwtmatr <- (ts1cwtmatr * ts1cwtmatr)
```

```
        #Calculating deltaj^(2)- deltaj+1^(2)
        for (j in 1:(fN)) {
            delta2j[j] <- round(mean(ts1cwtmatr[1:min(Ts - 100,
    200 * (j < 3) + (aj[j])^9), j]), 5)
        }
        delta1_mat[N, ] <- delta2j
        numerator <- delta2j[1:(fN - 1)] - delta2j[2:fN]

        #Calculating Delta(deltaj.^(2))
        delat_result <- numerator/denominator
        delta2_mat[N, ] <- delat_result
        delta1_mat1[(1 + k * (m - 1)):(m * k), ] <- delta1_mat
        delta2_mat1[(1 + k * (m - 1)):(m * k), ] <- delta2_mat
    }
    rm(list = setdiff(ls(), c("ptm", "m", "M", "k", "fN", "c2",
    "c3", "delta1_mat1", "delta2_mat1", "s_mat1", "alpha_mat1")))
    return(cbind(delta1_mat1[(1 + k * (m - 1)):(m * k), ],
    delta2_mat1[(1 + k * (m - 1)):(m * k), ]))
}

#Splitting the two matrices into delta1_mat1 and delta2_mat1
delta1_mat1 <- out[, 1:fN]
delta2_mat1 <- out[, (fN + 1):(2 * fN - 1)]

timetaken1 <- (proc.time() - ptm)/60
save.image(file = "Longmemory_estimates110.RData")

#Closing the cluster
parallel::stopCluster(cl)
```

Bibliography

1. Abazajian, K., Addison, G., Adshead, P., Ahmed, Z., Allen, S., Alonso, D., et al. (2019). CMB-S4 Science Case, Reference Design, and Project Plan. https://arxiv.org/abs/1907.04473.
2. Abramowitz, M. and Stegun, I. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York.
3. Abry, P. and Didier, G. (2018). Wavelet estimation for operator fractional Brownian motion. *Bernoulli*, 24(2):895–928.
4. Abry, P., Didier, G., and Li, H. (2019). Two-step wavelet-based estimation for Gaussian mixed fractional processes. *Statistical Inference for Stochastic Processes*, 22(2):157–185.
5. Abry, P. and Veitch, D. (1998). Wavelet analysis of long-range-dependent traffic. *IEEE Transactions on Information Theory*, 44(1):2–15.
6. Adachi, S., Aguilar Faúndez, M. A. O., Arnold, K., Baccigalupi, C., et al. (2020). A measurement of the CMB E-mode angular power spectrum at subdegree scales from 670 square degrees of POLARBEAR data. *The Astrophysical Journal*, 904(1):65.
7. Adam, R., Ade, P., Aghanim, N., Alves, M., Arnaud, M., Ashdown, M., et al. (2016a). Planck 2015 results - X. Diffuse component separation: Foreground maps. *Astronomy & Astrophysics*, 594:A10.
8. Adam, R., Ade, P. A. R., Aghanim, N., Akrami, Y., Alves, M. I. R., Argüeso, F., et al. (2016b). Planck 2015 results - I. Overview of products and scientific results. *Astronomy & Astrophysics*, 594(A1):38.
9. Ade, P. A. R., Aghanim, N., Akrami, Y., Aluri, P. K., Arnaud, M., Ashdown, M., et al. (2016). Planck 2015 results - XVI. Isotropy and statistics of the CMB. *Astronomy & Astrophysics*, 594(A16):62.
10. Ade, P. A. R., Aghanim, N., Alves, M. I. R., Armitage-Caplan, C., Arnaud, M., Ashdown, M., et al. (2014). Planck 2013 results - I. Overview of products and

scientific results. *Astronomy & Astrophysics*, 571(A1):48.

11. Adhikari, S., Shandera, S., and Erickcek, A. L. (2016). Large-scale anomalies in the cosmic microwave background as signatures of non-Gaussianity. *Physical Review D*, 93(2):023524.

12. Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., et al. (2020). Planck 2018 results-VI. Cosmological parameters. *Astronomy & Astrophysics*, 641(A6):67.

13. Ahmad, N., Kamal, A., Khan, M. M., Hushnud, and Tufail, A. (2014). A study of multifractal spectra and Rényi dimensions in 14.5A GeV/c ^{28}Si-Nucleus collisions. *Journal of Modern Physics*, 5(14):1288–1293.

14. Allen, L. J. S. (2010). *An Introduction to Stochastic Processes with Applications to Biology*. CRC Press, Boca Raton.

15. Alodat, T., Leonenko, N., and Olenko, A. (2020). Limit theorems for filtered long-range dependent random fields. *Stochastics*, 92(8):1175–1196.

16. Alodat, T. and Olenko, A. (2018). Weak convergence of weighted additive functionals of long-range dependent fields. *Theory of Probability and Mathematical Statistics*, 97:1–16.

17. Alodat, T. and Olenko, A. (2020). On asymptotics of discretized functionals of long-range dependent functional data. *To appear in Communications in Statistics - Theory and Methods*, pages 1–26.

18. Alomari, H. M., Ayache, A., Fradon, M., and Olenko, A. (2020). Estimation of cyclic long-memory parameters. *Scandinavian Journal of Statistics*, 47(1):104–133.

19. Alomari, H. M., Frías, M. P., Leonenko, N. N., Ruiz-Medina, M. D., Sakhno, L., and Torres, A. (2017). Asymptotic properties of parameter estimates for random fields with tapered data. *Electronic Journal of Statistics*, 11(2):3332–3367.

20. Alpher, R. A., Bethe, H., and Gamow, G. (1948a). The origin of chemical elements. *Physical Review*, 73(7):803.

21. Alpher, R. A., Herman, R., and Gamow, G. A. (1948b). Thermonuclear reactions in the expanding universe. *Physical Review*, 74(9):1198.

22. Anděl, J. (1986). Long-memory time series models. *Kybernetika*, 22(2):105–123.

23. Andersson, P. (1997). Characterization of pointwise Hölder regularity. *Applied and Computational Harmonic Analysis*, 4(4):429–443.

24. Angulo, J. and Esquivel, F. J. (2015). Multifractal dimensional dependence assess-

ment based on Tsallis mutual information. *Entropy*, 17(8):5382–5401.

25. Anh, V., Broadbridge, P., Olenko, A., and Wang, Y. G. (2018). On approximation for fractional stochastic partial differential equations on the sphere. *Stochastic Environmental Research and Risk Assessment*, 32(9):2585–2603.

26. Anh, V., Leonenko, N., and Olenko, A. (2015). On the rate of convergence to Rosenblatt-type distribution. *Journal of Mathematical Analysis and Applications*, 425(1):111–132.

27. Anh, V., Leonenko, N., Olenko, A., and Vaskovych, V. (2019a). On rate of convergence in non-central limit theorems. *Bernoulli*, 25(4A):2920–2948.

28. Anh, V., Leonenko, N., and Shieh, N.-R. (2008). Multifractality of products of geometric Ornstein-Uhlenbeck-type processes. *Advances in Applied Probability*, 40(4):1129–1156.

29. Anh, V., Olenko, A., and Vaskovych, V. (2019b). On LSE in regression model for long-range dependent random fields on spheres. *Statistics*, 53(5):1131–1151.

30. Anh, V. V., Angulo, J. M., and Ruiz-Medina, M. D. (1999). Possible long-range dependence in fractional random fields. *Journal of Statistical Planning and Inference*, 80(1-2):95–110.

31. Anh, V. V., Leonenko, N. N., and Sakhno, L. M. (2004). Quasi-likelihood-based higher-order spectral estimation of random fields with possible long-range dependence. *Journal of Applied Probability*, 41(A):35–53.

32. Anh, V. V., Leonenko, N. N., and Sakhno, L. M. (2007a). Minimum contrast estimation of random processes based on information of second and third orders. *Journal of Statistical Planning and Inference*, 137(4):1302–1331.

33. Anh, V. V., Leonenko, N. N., and Sakhno, L. M. (2007b). Statistical inference using higher-order information. *Journal of Multivariate Analysis*, 98(4):706–742.

34. Appel, M. and Pebesma, E. (2020). Spatiotemporal multi-resolution approximations for analyzing global environmental data. *Spatial Statistics*, 38:100465.

35. Arneodo, A., D'Aubenton-Carafa, Y., Audit, B., Bacry, E., Muzy, J.-F., and Thermes, C. (1998). What can we learn with wavelets about DNA sequences? *Physica A: Statistical Mechanics and its Applications*, 249(1-4):439–448.

36. Arteche, J. (2020). Exact local Whittle estimation in long memory time series with multiple poles. *Econometric Theory*, 36(6):1064–1098.

37. Arteche, J. and Robinson, P. M. (1999). Seasonal and cyclical long memory. In

Ghosh, S., editor, *Asymptotics, Nonparametrics, and Time Series*, pages 115–148. Marcel Dekker, Inc, New York.

38. Arteche, J. and Robinson, P. M. (2000). Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, 21(1):1–25.

39. Arteche, J. and Velasco, C. (2005). Trimming and tapering semi-parametric estimates in asymmetric long-memory time series. *Journal of Time Series Analysis*, 26(4):581–611.

40. Artiach, M. and Arteche, J. (2011). Estimation of the frequency in cyclical long-memory series. *Journal of Statistical Computation and Simulation*, 81(11):1627–1639.

41. Audit, B., Bacry, E., Muzy, J.-F., and Arneodo, A. (2002). Wavelet-based estimators of scaling behavior. *IEEE Transactions on Information Theory*, 48(11):2938–2954.

42. Ayache, A. (2002). The generalized multifractional field: a nice tool for the study of the generalized multifractional Brownian motion. *The Journal of Fourier Analysis and Applications*, 8(6):581–602.

43. Ayache, A. (2013). Continuous Gaussian multifractional processes with random pointwise Hölder regularity. *Journal of Theoretical Probability*, 26(1):72–93.

44. Ayache, A. (2018). *Multifractional Stochastic Fields: Wavelet Strategies in Multifractional Frameworks*. World Scientific, Singapore.

45. Ayache, A., Benassi, A., Cohen, S., and Véhel, J. L. (2005). Regularity and identification of generalized multifractional Gaussian processes. In Émery, M., Ledoux, M., and Yor, M., editors, *Séminaire de Probabilités XXXVIII*, pages 290–312. Springer, Berlin, Heidelberg.

46. Ayache, A. and Bertrand, P. R. (2011). Discretization error of wavelet coefficient for fractal like processes. *Advances in Pure and Applied Mathematics*, 2(2):297–321.

47. Ayache, A. and Esmili, Y. (2020). Wavelet-type expansion of the generalized rosenblatt process and its rate of convergence. *Journal of Fourier Analysis and Applications*, 26(3):1–35.

48. Ayache, A., Fradon, M., Nanayakkara, R., and Olenko, A. (2021). Asymptotic normality of simultaneous estimators of cyclic long-memory processes. *To appear in Electronic Journal of Statistics*, pages 1–32.

49. Ayache, A. and Jaffard, S. (2010). Hölder exponents of arbitrary functions. *Revista Matemática Iberoamericana*, 26(1):77–89.

50. Ayache, A., Jaffard, S., and Taqqu, M. S. (2007). Wavelet construction of generalized multifractional processes. *Revista Matematica Iberoamericana*, 23(1):327–370.

51. Ayache, A., Shieh, N.-R., and Xiao, Y. (2020). Wavelet series representation and geometric properties of harmonizable fractional stable sheets. *Stochastics*, 92(1):1–23.

52. Ayache, A. and Taqqu, M. S. (2005). Multifractional processes with random exponent. *Publicacions Matemátiques*, 49(2):459–486.

53. Ayache, A. and Véhel, J. (2004). On the identification of the pointwise Hölder exponent of the generalized multifractional Brownian motion. *Stochastic Processes and their Applications*, 111(1):119–156.

54. Ayache, A. and Véhel, J. L. (1999). Generalized multifractional Brownian motion: Definition and preliminary results. In Dekking, M., Véhel, J. L., Lutton, E., and Tricot, C., editors, *Fractals: Theory and Applications in Engineering*, pages 17–32. Springer, London.

55. Ayache, A. and Véhel, J. L. (2000). The generalized multifractional Brownian motion. *Statistical Inference for Stochastic Processes*, 3(1):7–18.

56. Bacry, E. and Muzy, J. F. (2003). Log-infinitely divisible multifractal processes. *Communications in Mathematical Physics*, 236(3):449–475.

57. Baiesi, M., Burigana, C., Conti, L., Falasco, G., Maes, C., Rondoni, L., et al. (2020). Possible nonequilibrium imprint in the cosmic background at low frequencies. *Physical Review Research*, 2(1):013210.

58. Baldi, P. and Marinucci, D. (2007). Some characterizations of the spherical harmonics coefficients for isotropic random fields. *Statistics & Probability Letters*, 77(5):490–496.

59. Barboza, L. A. and Viens, F. G. (2017). Parameter estimation of Gaussian stationary processes using the generalized method of moments. *Electronic Journal of Statistics*, 11(1):401–439.

60. Bardet, J. M. and Bertrand, P. R. (2010). A non-parametric estimator of the spectral density of a continuous-time Gaussian process observed at random times. *Scandinavian Journal of Statistics*, 37(3):458–476.

61. Bardet, J.-M. and Surgailis, D. (2013). Nonparametric estimation of the local Hurst function of multifractional Gaussian processes. *Stochastic Processes and their Applications*, 123(3):1004–1045.

62. Barral, J. and Mandelbrot, B. B. (2002). Multifractal products of cylindrical pulses.

Probability Theory and Related Fields, 124(3):409–430.

63. Barreiro, R. B., Hobson, M. P., Lasenby, A. N., Banday, A. J., Górski, K. M., and Hinshaw, G. (2000). Testing the Gaussianity of the COBE-DMR data with spherical wavelets. *Monthly Notices of the Royal Astronomical Society*, 318(2):475–481.

64. Bartlett, M. S. (1955). *An Introduction to Stochastic Processes: With Special Reference to Methods and Applications*. Cambridge University Press, Cambridge.

65. Bartolo, N., Matarrese, S., and Riotto, A. (2010). Non-Gaussianity and the cosmic microwave background anisotropies. *Advances in Astronomy*, 2010:68.

66. Beaumont, P. and Smallwood, A. (2019). Inference for likelihood-based estimators of generalized long-memory processes. MPRA Paper 96313, University Library of Munich, Germany. Retrieved from https://ideas.repec.org/p/pra/mprapa/96313.html.

67. Benassi, A., Bertrand, P., Cohen, S., and Istas, J. (2000). Identification of the Hurst index of a step fractional Brownian motion. *Statistical Inference for Stochastic Processes*, 3(1):101–111.

68. Benassi, A., Cohen, S., and Istas, J. (1998a). Identifying the multifractional function of a Gaussian process. *Statistics & Probability Letters*, 39(4):337–345.

69. Benassi, A., Cohen, S., Istas, J., and Jaffard, S. (1998b). Identification of filtered white noises. *Stochastic Processes and their Applications*, 75(1):31–49.

70. Benassi, A., Roux, D., and Jaffard, S. (1997). Elliptic Gaussian random processes. *Revista Matemática Iberoamericana*, 13(1):19–90.

71. Benzi, R., Paladin, G., Parisi, G., and Vulpiani, A. (1984). On the multifractal nature of fully developed turbulence and chaotic systems. *Journal of Physics A: Mathematical and General*, 17(18):3521.

72. Beran, J. (1992). Statistical methods for data with long-range dependence. *Statistical Science*, 7(4):404–416.

73. Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman & Hall/CRC, Boca Raton.

74. Beran, J., Feng, Y., Ghosh, S., and Kulik, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer, Berlin.

75. Beran, J., Ghosh, S., and Schell, D. (2009). On least squares estimation for long-memory lattice processes. *Journal of Multivariate Analysis*, 100(10):2178–2194.

76. Beran, J. and Shumeyko, Y. (2012). On asymptotically optimal wavelet estimation of trend functions under long-range dependence. *Bernoulli*, 18(1):137–176.

77. Biagini, F., Hu, Y., Øksendal, B., and Zhang, T. (2008). *Stochastic Calculus for Fractional Brownian Motion and Applications*. Springer-Verlag, London.

78. Bianchi, S. (2005). Pathwise identification of the memory function of multifractional Brownian motion with application to finance. *International Journal of Theoretical and Applied Finance*, 8(2):255–281.

79. Bianchi, S. and Pianese, A. (2007). Modelling stock price movements: multifractality or multifractionality? *Quantitative Finance*, 7(3):301–319.

80. Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular Variation*. Cambridge University Press, Cambridge.

81. Bjorken, J. D. (2003). Cosmology and the standard model. *Physical Review D*, 67(4):043508.

82. Boniece, B. C., Didier, G., and Sabzikar, F. (2021). Tempered fractional Brownian motion: Wavelet estimation, modeling and testing. *Applied and Computational Harmonic Analysis*, 51:461–509.

83. Boubaker, H. (2015). Wavelet estimation of Gegenbauer processes: Simulation and empirical application. *Computational Economics*, 46(4):551–574.

84. Boubaker, H. and Sghaier, N. (2015). Semiparametric generalized long-memory modeling of some mena stock market returns: A wavelet approach. *Economic Modelling*, 50:254–265.

85. Bradley, R. C. (2007). *Introduction to Strong Mixing Conditions Vol. 1,2 and 3*. Kendrick Press, Heber City.

86. Broadbridge, P. and Deutscher, K. (2020). Solution of non-autonomous Schrödinger equation for quantized de sitter Klein-Gordon oscillator modes undergoing attraction-repulsion transition. *Symmetry*, 12(6):943.

87. Broadbridge, P., Kolesnik, A. D., Leonenko, N., and Olenko, A. (2019). Random spherical hyperbolic diffusion. *Journal of Statistical Physics*, 177(5):889–916.

88. Broadbridge, P., Kolesnik, A. D., Leonenko, N., Olenko, A., and Omari, D. (2020). Spherically restricted random hyperbolic diffusion. *Entropy*, 22(2):217.

89. Broadbridge, P., Nanayakkara, R., and Olenko, A. (2021). On multifractionality of spherical random fields with cosmological applications. https://arxiv.org/abs/2104.13945.

90. Bromley, B. C. and Tegmark, M. (1999). Is the cosmic microwave background really non-Gaussian? *The Astrophysical Journal*, 524(2):L79–L82.

91. Brown, G., Michon, G., and Peyriére, J. (1992). On the multifractal analysis of measures. *Journal of Statistical Physics*, 66(3):775–790.

92. Buchert, T., France, M. J., and Steiner, F. (2017). Model-independent analyses of non-Gaussianity in Planck CMB maps using Minkowski functionals. *Classical and Quantum Gravity*, 34(9):094002.

93. Calcagni, G. (2019). Multifractional spacetimes from the standard model to cosmology. *International Journal of Geometric Methods in Modern Physics*, 16(supp01):1940004.

94. Calcagni, G. (2021). Multifractional theories: an updated review. *Modern Physics Letters A*, 36(14):2140006.

95. Calcagni, G., Kuroyanagi, S., and Tsujikawa, S. (2016). Cosmic microwave background and inflation in multifractional spacetimes. *Journal of Cosmology and Astroparticle Physics*, 2016(8):039.

96. Calvet, L. and Fisher, A. (2002). Multifractality in asset returns: Theory and evidence. *The Review of Economics and Statistics*, 84(3):381–406.

97. Calvet, L. E., Fisher, A. J., and Mandelbrot, B. B. (1997). Large deviations and the distribution of price changes. Cowles Foundation Discussion Paper No. 1165, Sauder School of Business Working Paper.

98. Cammarota, V. and Marinucci, D. (2015). The stochastic properties of l^1-regularized spherical gaussian fields. *Applied and Computational Harmonic Analysis*, 38(2):262–283.

99. Caniego, F. J., Espejo, R., Martín, M. A., and José, F. S. (2005). Multifractal scaling of soil spatial variability. *Ecological Modelling*, 182(3-4):291–303.

100. Carlin, J. B. and Dempster, A. P. (1989). Sensitivity analysis of seasonal adjustments: Empirical case studies. *Journal of the American Statistical Association*, 84(405):6–20.

101. Carmona, R. and Torresani, B. (2021). *Rwave: Time-Frequency Analysis of 1-D Signals*. R package version 2.5-0. Retrieved from https://CRAN.R-project.org/package=Rwave.

102. Castelvecchi, D. (2019). The quest to unlock the secrets of the baby universe. *Nature*, 572(7769):298–301.

103. Cayón, L., Sanz, J. L., Martínez-González, E., Banday, A. J., Argüeso, F., Gallegos, J. E., et al. (2001). Spherical Mexican hat wavelet: an application to detect non-Gaussianity in the COBE-DMR maps. *Monthly Notices of the Royal

Astronomical Society, 326(4):1243–1248.

104. Cazelles, B., Chavez, M., Magny, G. C. d., Guégan, J.-F., and Hales, S. (2007). Time-dependent spectral analysis of epidemiological time-series with wavelets. *Journal of the Royal Society Interface*, 4(15):625–636.

105. Chan, K.-S. and Tsai, H. (2012). Inference of seasonal long-memory aggregate time series. *Bernoulli*, 18(4):1448–1464.

106. Chesneau, C., Doosti, H., and Stone, L. (2019). Adaptive wavelet estimation of a function from an m-dependent process with possibly unbounded m. *Communications in Statistics-Theory and Methods*, 48(5):1123–1135.

107. Chhabra, A. B., Meneveau, C., Jensen, R. V., and Sreenivasan, K. R. (1989). Direct determination of the f(α) singularity spectrum and its application to fully developed turbulence. *Physical Review A*, 40(9):5284.

108. Chiann, C. and Morettin, P. A. (1998). A wavelet analysis for time series. *Journal of Nonparametric Statistics*, 10(1):1–46.

109. Chingangbam, P., Park, C., Yogendran, K. P., and Van de Weygaert, R. (2012). Hot and cold spot counts as probes of non-Gaussianity in the cosmic microwave background. *The Astrophysical Journal*, 755(2):122.

110. Christakos, G. (2017). *Spatiotemporal Random Fields: Theory and Applications*. Elsevier, Amsterdam.

111. Chung, C.-F. (1996a). Estimating a generalized long-memory process. *Journal of Econometrics*, 73(1):237–259.

112. Chung, C.-F. (1996b). A generalized fractionally integrated autoregressive moving-average process. *Journal of Time Series Analysis*, 17(2):111–140.

113. Coeurjolly, J.-F. (2001). Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths. *Statistical Inference for Stochastic Processes*, 4(2):199–227.

114. Coeurjolly, J.-F. (2005). Identification of multifractional Brownian motion. *Bernoulli*, 11(6):987–1008.

115. Cohen, S. and Istas, J. (2013). *Fractional Fields and Applications*. Springer, Berlin, Heidelberg.

116. Coleman, P. H. and Pietronero, L. (1992). The fractal structure of the universe. *Physics Reports*, 213(6):311–389.

117. Collaboration, T. C., Armitage-Caplan, C., Avillez, M., Barbosa, D., Banday, A.,

Bartolo, N., et al. (2011). COrE (Cosmic Origins Explorer) A white paper. https://arxiv.org/abs/1102.2181v1.

118. Constantine, W. and Percival, D. (2017). *fractal: A Fractal Time Series Modeling and Analysis Package*. R package version 2.0-4. Retrieved from https://CRAN.R-project.org/package=fractal.
119. Copi, C. J., Huterer, D., Schwarz, D. J., and Starkman, G. D. (2010). Large-angle anomalies in the CMB. *Advances in Astronomy*, 2010(847541):1–17.
120. Cornish, C. R., Bretherton, C. S., and Percival, D. B. (2006). Maximal overlap wavelet statistical analysis with application to atmospheric turbulence. *Boundary-Layer Meteorology*, 119(2):339–374.
121. Cox, D. R. and Miller, H. D. (1977). *The Theory of Stochastic Processes*. Chapman & Hall/CRC, Boca Raton.
122. Cressie, N. (1989). Geostatistics. *The American Statistician*, 43(4):197–202.
123. Cressie, N. (1993). *Statistics for Spatial Data*. John Wiley & Sons, New York.
124. Cruz, M., Martínez-González, E., Vielva, P., and Cayón, L. (2005). Detection of a non-Gaussian spot in WMAP. *Monthly Notices of the Royal Astronomical Society*, 356(1):29–40.
125. Csörgo, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York.
126. Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *The Annals of Statistics*, 17(4):1749–1766.
127. Daoudi, K., Véhel, J. L., and Meyer, Y. (1998). Construction of continuous functions with prescribed local regularity. *Constructive Approximation*, 14(3):349–385.
128. de Bernardis, P., Ade, P. A., Bock, J. J., Bond, J., Borrill, J., Boscaleri, A., et al. (2000). A flat universe from high-resolution maps of the cosmic microwave background radiation. *Nature*, 404(6781):955–959.
129. Dedecker, J., Doukhan, P., Lang, G., Leon, J. R., Louhichi, S., and Prieur, C. (2007). *Weak Dependence: With Examples and Applications*. Springer-Verlag, New York.
130. Dehling, H. and Philipp, W. (2002). Empirical process techniques for dependent data. In Dehling, H., Mikosch, T., and Sorensen, M., editors, *Empirical Process Techniques for Dependent Data*, pages 3–113. Birkhäuser, Boston.
131. del Barrio Castro, T. and Rachinger, H. (2021). Aggregation of seasonal long-

memory processes. *Econometrics and Statistics*, 17:95–106.

132. Denisov, D. and Leonenko, N. (2016). Limit theorems for multifractal products of geometric stationary processes. *Bernoulli*, 22(4):2579–2608.

133. Diego, J. M., Martínez-González, E., Sanz, J. L., Mollerach, S., and Martínez, V. J. (1999). Partition function based analysis of cosmic microwave background maps. *Monthly Notices of the Royal Astronomical Society*, 306(2):427–436.

134. Dissanayake, G. S., Peiris, M. S., and Proietti, T. (2016). State space modeling of Gegenbauer processes with long-memory. *Computational Statistics & Data Analysis*, 100:115–130.

135. Doob, J. L. (1937). Stochastic processes depending on a continuous parameter. *Transactions of the American Mathematical Society*, 42(1):107–140.

136. Doob, J. L. (1953). *Stochastic Processes*. Wiley, New York.

137. Doukhan, P. (1994). *Mixing: Properties and Examples, Volume 85 of Lecture Notes in Statistics*. Springer-Verlag, New York.

138. Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, 84(2):313–342.

139. Doukhan, P., Mayo, N., and Truquet, L. (2009). Weak dependence, models and some applications. *Metrika*, 69(2):199–225.

140. Doukhan, P., Oppenheim, G., and Taqqu, M. (2003). *Theory and Applications of Long-Range Dependence*. Birkhäuser, Boston.

141. D'Ovidio, M., Leonenko, N., and Orsingher, E. (2016). Fractional spherical random fields. *Statistics & Probability Letters*, 116:146–156.

142. Du, P., Kibbe, W. A., and Lin, S. M. (2006). Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics*, 22(17):2059–2065.

143. Durastanti, C., Lan, X., and Marinucci, D. (2014). Gaussian semiparametric estimates on the unit sphere. *Bernoulli*, 20(1):28–77.

144. Durrer, R. (2015). The cosmic microwave background: the history of its experimental investigation and its significance for cosmology. *Classical and Quantum Gravity*, 32(12):124007.

145. Efstathiou, G., Bond, J., and White, S. D. (1992). COBE background radiation anisotropies and large-scale structure in the Universe. *Monthly Notices of the Royal*

Astronomical Society, 258(1):1P–6P.

146. Emery, X. and Porcu, E. (2019). Simulating isotropic vector-valued Gaussian random fields on the sphere through finite harmonics approximations. *Stochastic Environmental Research and Risk Assessment*, 33(8):1659–1667.

147. Emery, X., Porcu, E., and Bissiri, P. G. (2019). A semiparametric class of axially symmetric random fields on the sphere. *Stochastic Environmental Research and Risk Assessment*, 33:1863–1874.

148. Espejo, R. M., Leonenko, N., Olenko, A., and Ruiz-Medina, M. D. (2015). On a class of minimum contrast estimators for Gegenbauer random fields. *TEST*, 24(4):657–680.

149. Espejo, R. M., Leonenko, N. N., and Ruiz-Medina, M. D. (2014). Gegenbauer random fields. *Random Operators and Stochastic Equations*, 22(1):1–16.

150. European Space Agency (2021a). Euclid. https://sci.esa.int/web/euclid. accessed 21 September 2021.

151. European Space Agency (2021b). Planck and the cosmic microwave background. https://www.esa.int/Science_Exploration/Space_Science/Planck/Planck_and_the_cosmic_microwave_background. accessed 21 September 2021.

152. European Space Agency (2021c). Planck finds no new evidence for cosmic anomalies. http://www.esa.int/Science_Exploration/Space_Science/Planck/Planck_finds_no_new_evidence_for_cosmic_anomalies. accessed 21 September 2021.

153. European Space Agency and the Planck Collaboration (2021). Planck's Power spectrum of temperature fluctuations in the Cosmic Microwave Background. https://sci.esa.int/web/planck/-/51555-planck-power-spectrum-of-temperature-fluctuations-in-the-cosmic-microwave-background. accessed 21 September 2021.

154. Evertsz, C. J. G. (1995). Fractal geometry of financial time series. *Fractals*, 3(3):609–616.

155. Evertsz, C. J. G. and Mandelbrot, B. B. (1992). Multifractal measures. *Chaos and Fractals*, 1092.

156. Falconer, K. J. (1990). *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons, Chichester.

157. Falconer, K. J. (1994). The multifractal spectrum of statistically self-similar measures. *Journal of Theoretical Probability*, 7(3):681–702.

158. Falconer, K. J. (1997). *Techniques in Fractal Geometry*. John Wiley & Sons, Chichester.

159. Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, 24(1):395–458.

160. Feller, W. (1950). *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, New York.

161. Ferrara, L. and Guégan, D. (2001). Comparison of parameter estimation methods in cyclical long memory time series. In Dunis, C. L. and Moody, J. and Timmermann, A., editor, *Developments in Forecast Combination and Portfolio Choice*, pages 179–195. Wiley, New York.

162. Ferreira, P. G., Magueijo, J., and Górski, K. M. (1998). Evidence for non-Gaussianity in the COBE DMR 4 year sky maps. *The Astrophysical Journal*, 503(1):L1.

163. Fisher, A. J., Calvet, L. E., and Mandelbrot, B. B. (1997). Multifractality of deutschemark/us dollar exchange rates. Cowles Foundation Discussion Paper No. 1166, Sauder School of Business Working Paper.

164. Fisher, N. I., Lewis, T., and Embleton, B. J. J. (1993). *Statistical Analysis of Spherical Data*. Cambridge University Press, Cambridge.

165. Florescu, I. (2014). *Probability and Stochastic Processes*. John Wiley & Sons, New Jersey.

166. Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 14(2):517–532.

167. Franzke, C. (2010). Long-range dependence and climate noise characteristics of Antarctic temperature data. *Journal of Climate*, 23(22):6074–6081.

168. Fryer, D., Li, M., and Olenko, A. (2020). rcosmo: R Package for Analysis of Spherical, HEALPix and Cosmological Data. *The R Journal*, 12(1):206–225.

169. Fryer, D. and Olenko, A. (2019). Spherical data handling and analysis with R package rcosmo. In Nguyen, H., editor, *Statistics and Data Science. RSSDS 2019*, pages 211–225. Springer, Singapore.

170. Fryer, D., Olenko, A., Li, M., and Wang, Y. G. (2019). *rcosmo: Cosmic Microwave Background Data Analysis*. R package version 1.1.0. Retrieved from https://CRAN.R-project.org/package=rcosmo.

171. Gallager, R. G. (2013). *Stochastic Processes: Theory for Applications*. Cambridge University Press, Cambridge.

172. Gao, J., Anh, V., Heyde, C., and Tieng, Q. (2001). Parameter estimation of stochastic processes with long-range dependence and intermittency. *Journal of Time Series Analysis*, 22(5):517–535.

173. Gao, J. T. and Anh, V. V. (1999). Semiparametric regression under long-range dependent errors. *Journal of Statistical Planning and Inference*, 80(1-2):37–57.

174. García-Ancona, R., Morais, J., and Michael Porter, R. (2020). Relations among spheroidal and spherical harmonics. *Applied Mathematics and Computation*, 384:125147.

175. Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long-memory time series models. *Journal of Time Series Analysis*, 4(4):221–238.

176. Gikhman, I. I. and Skorokhod, A. V. (2004a). *The Theory of Stochastic Processes I*. Springer-Verlag, Berlin.

177. Gikhman, I. I. and Skorokhod, A. V. (2004b). *The Theory of Stochastic Processes II*. Springer-Verlag, Berlin.

178. Gikhman, I. I. and Skorokhod, A. V. (2007). *The Theory of Stochastic Processes III*. Springer-Verlag, Berlin.

179. Giraitis, L., Hidalgo, J., and Robinson, P. M. (2001). Gaussian estimation of parametric spectral density with unknown pole. *The Annals of Statistics*, 29(4):987–1023.

180. Giraitis, L., Koul, H. L., and Surgailis, D. (2012). *Large Sample Inference For Long Memory Processes*. Imperial College Press, London.

181. Giraitis, L. and Leipus, R. (1995). A generalized fractionally differencing approach in long-memory modeling. *Lithuanian Mathematical Journal*, 35(1):53–65.

182. Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probability Theory and Related Fields*, 86(1):87–104.

183. Gneiting, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli*, 19(4):1327–1349.

184. Gonçalvés, P. and Riedi, R. (2005). Diverging moments and parameter estimation. *Journal of the American Statistical Association*, 100(472):1382–1393.

185. Gorski, K. M., Hivon, E., Banday, A., Wandelt, B. D., Hansen, F. K., Reinecke, M., et al. (2005). HEALPix: A framework for high-resolution discretization

and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759–771.

186. Grahovac, D. and Leonenko, N. (2014). Detecting multifractal stochastic processes under heavy-tailed effects. *Chaos, Solitons & Fractals*, 65:78–89.

187. Grahovac, D. and Leonenko, N. (2018). Bounds on the support of the multifractal spectrum of stochastic processes. *Fractals*, 26(4):1850055.

188. Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29.

189. Graves, T., Gramacy, R., Watkins, N., and Franzke, C. (2017). A brief history of long-memory: Hurst, Mandelbrot and the road to ARFIMA, 1951-1980. *Entropy*, 19(9):437.

190. Gray, H. L., Zhang, N.-F., and Woodward, W. A. (1989). On generalized fractional processes. *Journal of Time Series Analysis*, 10(3):233–257.

191. Guerrier, S., Molinari, R., Victoria-Feser, M.-P., and Xu, H. (2021). Robust two-step wavelet-based inference for time series models. *To appear in Journal of the American Statistical Association*, pages 1–18.

192. Gupta, V. K. and Waymire, E. C. (1993). A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology and Climatology*, 32(2):251–267.

193. Gustafson, J. L. (1988). Reevaluating Amdahl's law. *Communications of the ACM*, 31(5):532–533.

194. Hamann, J., Gia, Q. T. L., Sloan, I. H., Wang, Y. G., and Womersley, R. S. (2021). A new probe of Gaussianity and isotropy with application to cosmic microwave background maps. *International Journal of Modern Physics C*, 32(6):2150084.

195. Harte, D. (2001). *Multifractals: Theory and Applications*. Chapman and Hall/CRC, Boca Raton.

196. Hassler, U. (1994). (mis)specification of long-memory in seasonal time series. *Journal of Time Series Analysis*, 15(1):19–30.

197. Hassler, U. and Wolters, J. (1995). Long memory in inflation rates: International evidence. *Journal of Business & Economic Statistics*, 13(1):37–45.

198. Hawking, S. W. and Ellis, G. F. R. (1968). The cosmic black-body radiation and the existence of singularities in our universe. *The Astrophysical Journal*, 152:25.

199. Herbin, E. (2006). From N parameter fractional Brownian motions to N parameter multifractional Brownian motions. *The Rocky Mountain Journal of Mathematics*,

36(4):1249–1284.

200. Hernàndez, D. B. (1995). *Lectures on Probability and Second Order Random Fields*. World Scientific, Singapore.

201. Hesse, K., Sloan, I. H., and Womersley, R. S. (2017). Radial basis function approximation of noisy scattered data on the sphere. *Numerische Mathematik*, 137(3):579–605.

202. Hesse, K., Sloan, I. H., and Womersley, R. S. (2021). Local rbf-based penalized least-squares approximation on the sphere with noisy scattered data. *Journal of Computational and Applied Mathematics*, 382:113061.

203. Heyde, C. C. and Gay, R. (1989). On asymptotic quasi-likelihood estimation. *Stochastic Processes and their Applications*, 31(2):223–236.

204. Heyde, C. C. and Gay, R. (1993). Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. *Stochastic Processes and their Applications*, 45(1):169–182.

205. Hidalgo, J. (2005). Semiparametric estimation for stationary processes whose spectra have an unknown pole. *The Annals of Statistics*, 33(4):1843–1889.

206. Hidalgo, J. and Soulier, P. (2004). Estimation of the location and exponent of the spectral singularity of a long-memory process. *Journal of Time Series Analysis*, 25(1):55–81.

207. Hikage, C., Komatsu, E., and Matsubara, T. (2006). Primordial non-Gaussianity and analytical formula for Minkowski functionals of the cosmic microwave background and large-scale structure. *The Astrophysical Journal*, 653(1):11.

208. Hikage, C., Koyama, K., Matsubara, T., Takahashi, T., and Yamaguchi, M. (2009). Limits on isocurvature perturbations from non-gaussianity in wmap temperature anisotropy. *Monthly Notices of the Royal Astronomical Society*, 398(4):2188–2198.

209. Hikage, C., Matsubara, T., Coles, P., Liguori, M., Hansen, F. K., and Matarrese, S. (2008). Limits on primordial non-Gaussianity from Minkowski functionals of the WMAP temperature anisotropies. *Monthly Notices of the Royal Astronomical Society*, 389(3):1439–1446.

210. Hill, J. C. (2018). Foreground biases on primordial non-Gaussianity measurements from the CMB temperature bispectrum: Implications for Planck and beyond. *Physical Review D*, 98(8):083542.

211. Hivon, E. (2021). Geometric and Algebraic Properties of HEALPix. https://healpix.jpl.nasa.gov/html/intronode4.htm. accessed 21 September 2021.

212. Hobson, M. P., Jones, A. W., and Lasenby, A. N. (1999). Wavelet analysis and the detection of non-Gaussianity in the cosmic microwave background. *Monthly Notices of the Royal Astronomical Society*, 309(1):125–140.

213. Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.

214. Hosking, J. R. M. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research*, 20(12):1898–1908.

215. Hosoya, Y. (1997). A limit theory for long–range dependence and statistical inference on related models. *The Annals of Statistics*, 25(1):105–137.

216. Hunt, R., Peiris, S., and Weber, N. (2021). A general frequency domain estimation method for Gegenbauer processes. *Journal of Time Series Econometrics*, 13(2):119–144.

217. Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1):770–799.

218. Hurvich, C. M. and Beltrao, K. I. (1993). Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *Journal of Time Series Analysis*, 14(5):455–472.

219. IRSA (2021). NASA/IPAC infrared science archive. https://irsa.ipac.caltech.edu/data/Planck/release_2/all-sky-maps/maps/component-maps/cmb/. accessed 21 September 2021.

220. Istas, J. and Lang, G. (1997a). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 33(4):407–436.

221. Istas, J. and Lang, G. (1997b). Quadratic variations and estimation of the local hölder index of a Gaussian process. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 33(4):407–436.

222. Itô, K. (1944). Stochastic integral. *Proceedings of the Imperial Academy Tokyo*, 20(8):519–524.

223. Ivanov, A., Leonenko, N., Ruiz-Medina, M. D., and Savich, I. N. (2013). Limit theorems for weighted nonlinear transformations of Gaussian stationary processes with singular spectra. *The Annals of Probability*, 41(2):1088–1114.

224. Ivanov, A. V. and Leonenko, N. N. (1989). *Statistical Analysis of Random Fields*. Springer, Dordrecht.

225. Jaffard, S. (1995). Functions with prescribed Hölder exponent. *Applied and Com-*

putational Harmonic Analysis, 2(4):400–401.

226. Jaffard, S. (1997a). Multifractal formalism for functions part i: results valid for all functions. *SIAM Journal on Mathematical Analysis*, 28(4):944–970.

227. Jaffard, S. (1997b). Old friends revisited: the multifractal nature of some classical functions. *Journal of Fourier Analysis and Applications*, 3(1):1–22.

228. Jaffard, S. (1999). The multifractal nature of Lévy processes. *Probability Theory and Related Fields*, 114(2):207–227.

229. Jaffard, S. (2004). Wavelet techniques in multifractal analysis. In Lapidus, M. L. and Frankenhuysen, M. V., editors, *Fractal Geometry and Applications: A Jubilee of Benoit Mandelbrot*, volume 72 of *Proceedings of Symposia in Pure Mathematics*, pages 91–152. American Mathematical Society, Providence.

230. Jain, P. L., Singh, G., and Mukhopadhyay, A. (1992). Multifractals at relativistic energies. *Physical Review C*, 46(2):721.

231. Janson, S. (1997). *Gaussian Hilbert Spaces*. Cambridge University Press, Cambridge.

232. Jeffreys, H. (1939). *Theory of Probability*. Clarendon Press, Oxford.

233. Jeong, J., Jun, M., and Genton, M. G. (2017). Spherical process models for global spatial statistics. *Statistical Science*, 32(4):501–513.

234. Jizba, P. and Arimitsu, T. (2004). The world according to Rényi: thermodynamics of multifractal systems. *Annals of Physics*, 312(1):17–59.

235. Kahane, J.-P. (1985). Sur le chaos multiplicatif. *Annales des sciences mathématiques du Québec*, 9(2):105–150.

236. Kahane, J.-P. (1987). Positive martingales and random measures. *Chinese Annals of Mathematics, Series B*, 8(1):1–12.

237. Karagiannis, T., Molle, M., and Faloutsos, M. (2004). Long-range dependence ten years of internet traffic modeling. *IEEE Internet Computing*, 8(5):57–64.

238. Karamata, J. (1930). Sur un mode de croissance reguliere des fonctions. *Mathematica (Cluj)*, 4:38–53.

239. Karhunen, K. (1947). *About linear methods in probability theory*. University of Helsinki, Helsinki.

240. Kent, J. T. and Wood, A. T. A. (1997). Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *Journal of the Royal Statistical Society. Series B (Methodological)*, 59(3):679–699.

241. Khatri, R. and Sunyaev, R. A. (2012). Creation of the CMB spectrum: precise

analytic solutions for the blackbody photosphere. *Journal of Cosmology and Astroparticle Physics*, 2012(6):038.

242. Khintchine, A. (1934). Correlation theory of stationary stochastic processes. *Mathematical annals*, 109(1):604–615.

243. Klimontovich, Y. L. (1967). *The Statistical Theory of Non-Equilibrium Processes in a Plasma.* Pergamon Press, Oxford.

244. Klykavka, B., Olenko, A., and Vicendese, M. (2012). Asymptotic behavior of functionals of cyclic long-range dependent random fields. *Journal of Mathematical Sciences*, 187(1):35–48.

245. Koenig, E. and Chainais, P. (2008). Multifractal analysis on the sphere. In Elmoataz, A., Lezoray, O., Nouboud, F., and Mammass, D., editors, *International Conference on Image and Signal Processing*, volume 5099, pages 613–621. Springer, Berlin, Heidelberg.

246. Kogut, A., Banday, A. J., Bennett, C. L., Górski, K. M., Hinshaw, G., Smoot, G. F., et al. (1996). Tests for non-Gaussian statistics in the DMR four-year sky maps. *The Astrophysical Journal Letters*, 464(1):L29–L33.

247. Kolmogorov, A. (1931). About the analytical methods in probability theory. *Mathematical annals*, 104(1):415–458.

248. Kolmogorov, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer-Verlag, Berlin.

249. Kolmogorov, A. N. (1940). Wienersche spiralen und einige andere interessante kurven in hilbertscen raum, comptes rendus (doklady). *Académie des Sciences de l'URSS (N.S.)*, 26:115–118.

250. Komatsu, E., Dunkley, J., Nolta, M. R., Bennett, C. L., Gold, B., Hinshaw, G., et al. (2009). Five-year Wilkinson microwave anisotropy probe observations: cosmological interpretation. *The Astrophysical Journal Supplement Series*, 180(2):330–376.

251. Laing, C. and Lord, G. J. (2010). *Stochastic Methods in Neuroscience.* Oxford University Press, Oxford.

252. Lang, A. and Schwab, C. (2015). Isotropic Gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations. *Annals of Applied Probability*, 25(6):3047–3094.

253. Larson, D. L. and Wandelt, B. D. (2004). The hot and cold spots in the Wilkinson microwave anisotropy probe data are not hot and cold enough. *The Astrophysical*

Journal, 613(2):L85.

254. Lawrance, A. J. and Kottegoda, N. T. (1977). Stochastic modelling of riverflow time series. *Journal of the Royal Statistical Society: Series A (General)*, 140(1):1–31.

255. Le Gia, Q. T., Sloan, I. H., Wang, Y. G., and Womersley, R. S. (2017). Needlet approximation for isotropic random fields on the sphere. *Journal of Approximation Theory*, 216:86–116.

256. Le Gia, Q. T., Sloan, I. H., Womersley, R. S., and Wang, Y. G. (2020). Isotropic sparse regularization for spherical harmonic representations of random fields on the sphere. *Applied and Computational Harmonic Analysis*, 49(1):257–278.

257. Lebovits, J. and Podolskij, M. (2017). Estimation of the global regularity of a multifractional Brownian motion. *Electronic Journal of Statistics*, 11(1):78–98.

258. Leclercq, F., Pisani, A., and Wandelt, B. D. (2014). Cosmology: From theory to data, from data to theory. In Cooray, A., Komatsu, E., and Melchiorri, A., editors, *New Horizons for Observational Cosmology*, pages 189–233. IOS Press, Amsterdam.

259. Leonenko, N. (1999). *Limit Theorems for Random Fields with Singular Spectrum.* Springer, Dordrecht.

260. Leonenko, N., Nanayakkara, R., and Olenko, A. (2021). Analysis of spherical monofractal and multifractal random fields. *Stochastic Environmental Research and Risk Assessment*, 35(3):681–701.

261. Leonenko, N. and Olenko, A. (2013). Tauberian and Abelian theorems for long-range dependent random fields. *Methodology and Computing in Applied Probability*, 15(4):715–742.

262. Leonenko, N. and Olenko, A. (2014). Sojourn measures of Student and Fisher-Snedecor random fields. *Bernoulli*, 20(3):1454–1483.

263. Leonenko, N. and Sakhno, L. (2012). On spectral representations of tensor random fields on the sphere. *Stochastic Analysis and Applications*, 30(1):44–66.

264. Leonenko, N. and Shieh, N.-R. (2013). Rényi function for multifractal random fields. *Fractals*, 21(2):1350009.

265. Leonenko, N. N. and Ruiz-Medina, M. D. (2018). Increasing domain asymptotics for the first Minkowski functional of spherical random fields. *Theory of Probability and Mathematical Statistics*, 97:127–149.

266. Li, C.-K. and Zhang, F. (2019). Eigenvalue continuity and Gersgorin's theorem.

Electronic Journal of Linear Algebra, 35:619–625.

267. Li, D., Robinson, P. M., and Shang, H. L. (2020a). Local Whittle estimation of long-range dependence for functional time series. *To appear in Journal of Time Series Analysis*.

268. Li, D., Robinson, P. M., and Shang, H. L. (2020b). Long-range dependent curve time series. *Journal of the American Statistical Association*, 115(530):957–971.

269. Liu, C.-L. (2010). A tutorial of the wavelet transform. Taipei, Taiwan: NTUEE. Retrieved from http://disp.ee.ntu.edu.tw/tutorial/WaveletTutorial.pdf.

270. Loboda, D., Mies, F., and Steland, A. (2021). Regularity of multifractional moving average processes with random Hurst exponent. *Stochastic Processes and their Applications*, 140:21–48.

271. Loéve, M. (1977). *Probability Theory I*. Springer, New York.

272. Lu, Z. and Guegan, D. (2011). Estimation of time-varying long memory parameter using wavelet method. *Communications in Statistics - Simulation and Computation*, 40(4):596–613.

273. Ma, C. (2016). Stochastic representations of isotropic vector random fields on spheres. *Stochastic Analysis and Applications*, 34(3):389–403.

274. Ma, C. (2017). Time varying isotropic vector random fields on spheres. *Journal of Theoretical Probability*, 30(4):1763–1785.

275. Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. Academic Press, San Diego.

276. Malyarenko, A. (2012). *Invariant Random Fields on Spaces with a Group Action*. Springer-Verlag, Berlin.

277. Malyarenko, A. and Ostoja-Starzewski, M. (2019). *Tensor-Valued Random Fields for Continuum Physics*. Cambridge University Press, Cambridge.

278. Malyarenko, A., Ostoja-Starzewski, M., and Amiri-Hezaveh, A. (2020). *Random Fields of Piezoelectricity and Piezomagnetism: Correlation Structures*. Springer, Cham.

279. Mandelbrot, B. B. (1965). Une classe de processus stochastiques homothétiques á soi; application á la loi climatologique de H.E. Hurst. *Comptes Rendus de l'Académie des Sciences de Paris*, 260:3274–3277.

280. Mandelbrot, B. B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, 156(3775):636–638.

281. Mandelbrot, B. B. (1972). Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. In Rosenblatt,

M. and Van Atta, C., editors, *Statistical Models and Turbulence*, pages 333–351. Springer, Berlin.

282. Mandelbrot, B. B. (1974). Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics*, 62(2):331–358.

283. Mandelbrot, B. B. (1977). *Fractals: Form, Chance, and Dimension.* W. H. Freeman and Company, San Francisco.

284. Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature.* W. H. Freeman and Company, New York.

285. Mandelbrot, B. B. (1989). Multifractal measures, especially for the geophysicist. In Scholz, C. H. and Mandelbrot, B. B., editors, *Fractals in Geophysics*, pages 5–42. Birkhäuser, Basel.

286. Mandelbrot, B. B. (1997). *Fractals and Scaling in Finance.* Springer, New York.

287. Mandelbrot, B. B., Fisher, A. J., and Calvet, L. E. (1997). A multifractal model of asset returns. Cowles Foundation Discussion Paper No. 1164, Sauder School of Business Working Paper.

288. Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437.

289. Mandelbrot, B. B. and Wallis, J. R. (1968). Noah, Joseph, and operational hydrology. *Water Resources Research*, 4(5):909–918.

290. Mannersalo, P., Norros, I., and Riedi, R. H. (2002). Multifractal products of stochastic processes: construction and some basic properties. *Advances in Applied Probability*, 34(4):888–903.

291. Marinucci, D. (2004). Testing for non-Gaussianity on cosmic microwave background radiation: A review. *Statistical Science*, 19(2):294–307.

292. Marinucci, D. (2005). Regression methods for testing gaussianity on a spherical random field. *Random Operators and Stochastic Equations*, 13(4):313–324.

293. Marinucci, D. (2006). High-resolution asymptotics for the angular bispectrum of spherical random fields. *The Annals of Statistics*, 34(1):1–41.

294. Marinucci, D. (2008). A central limit theorem and higher order results for the angular bispectrum. *Probability Theory and Related Fields*, 141(3):389–409.

295. Marinucci, D. and Peccati, G. (2010a). Ergodicity and gaussianity for spherical random fields. *Journal of Mathematical Physics*, 51(4):043301.

296. Marinucci, D. and Peccati, G. (2010b). Representations of SO(3) and angular polyspectra. *Journal of Multivariate Analysis*, 101(1):77–100.

297. Marinucci, D. and Peccati, G. (2011). *Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications*. Cambridge University Press, New York.

298. Martìnez, V. (1990). Fractals and multifractals in the description of the cosmic structure. *Vistas in Astronomy*, 33(3):337–356.

299. Mathias, A. C., Viana, R. L., Kroetz, T., and Caldas, I. L. (2017). Fractal structures in the chaotic motion of charged particles in a magnetized plasma under the influence of drift waves. *Physica A: Statistical Mechanics and its Applications*, 469:681–694.

300. Matsumura, T., Akiba, Y., Borrill, J., Chinone, Y., Dobbs, M., Fuke, H., et al. (2014). Mission design of LiteBIRD. *Journal of Low Temperature Physics*, 176(5):733–740.

301. McEwen, J. D., Hobson, M. P., Lasenby, A. N., and Mortlock, D. J. (2005). A high-significance detection of non-Gaussianity in the Wilkinson microwave anisotropy probe 1-yr data using directional spherical wavelets. *Monthly Notices of the Royal Astronomical Society*, 359(4):1583–1596.

302. McEwen, J. D., Hobson, M. P., Lasenby, A. N., and Mortlock, D. J. (2006). Non-Gaussianity detections in the Bianchi VII_h corrected WMAP one-year data made with directional spherical wavelets. *Monthly Notices of the Royal Astronomical Society*, 369(4):1858–1868.

303. Meyer, Y. (1992). *Wavelets and Operators: Volume 1*. Cambridge University Press, Cambridge.

304. Meyer, Y. (1993). *Wavelets: Algorithms and Applications*. SIAM, Philadelphia.

305. Microsoft and Weston, S. (2020). *foreach: Provides Foreach Looping Construct*. R package version 1.5.1. Retrieved from https://CRAN.R-project.org/package=foreach.

306. Microsoft Corporation and Weston, S. (2020). *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*. R package version 1.0.16. Retrieved from https://CRAN.R-project.org/package=doParallel.

307. Minkov, M., Pinkwart, M., and Schupp, P. (2019). Entropy methods for CMB analysis of anisotropy and non-Gaussianity. *Physical Review D*, 99(10):103501.

308. Molchan, G. M. (1996). Scaling exponents and multifractal dimensions for indepen-

dent random cascades. *Communications in Mathematical Physics*, 179(3):681–702.

309. Montanari, A., Rosso, R., and Taqqu, M. S. (2000). A seasonal fractional ARIMA model applied to the nile river monthly flows at aswan. *Water Resources Research*, 36(5):1249–1259.

310. Montillet, J.-P. and Yu, K. (2015). Modeling geodetic processes with Levy α-stable distribution and FARIMA. *Mathematical Geosciences*, 47(6):627–646.

311. Moulines, E., Roueff, F., and Taqqu, M. S. (2007). On the spectral density of the wavelet coefficients of long-memory time series with application to the log-regression estimation of the memory parameter. *Journal of Time Series Analysis*, 28(2):155–187.

312. Moulines, E., Roueff, F., and Taqqu, M. S. (2008). A wavelet whittle estimator of the memory parameter of a nonstationary Gaussian time series. *The Annals of Statistics*, 36(4):1925–1956.

313. Muir, J., Adhikari, S., and Huterer, D. (2018). Covariance of CMB anomalies. *Physical Review D*, 98(2):023521.

314. Munshi, D., Heavens, A., Cooray, A., Smidt, J., Coles, P., and Serra, P. (2011). New optimized estimators for the primordial trispectrum. *Monthly Notices of the Royal Astronomical Society*, 412(3):1993–2016.

315. Muzy, J. F., Bacry, E., and Arneodo, A. (1993). Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method. *Physical Review E*, 47(2):875.

316. Muzy, J.-F., Bacry, E., and Arneodo, A. (1994). The multifractal formalism revisited with wavelets. *International Journal of Bifurcation and Chaos*, 4(2):245–302.

317. Muzy, J.-F., Sornette, D., Delour, J., and Arneodo, A. (2001). Multifractal returns and hierarchical portfolio theory. *Quantitative Finance*, 1(1).

318. NASA (2021). Tests of Big Bang: The CMB. https://map.gsfc.nasa.gov/universe/bb_tests_cmb.html. accessed 21 September 2021.

319. NASA Science (2021). Dark Energy, Dark Matter. https://science.nasa.gov/astrophysics/focus-areas/what-is-dark-energy/. accessed 21 September 2021.

320. Naselsky, P., Novikov, D., and Novikov, I. (2006). *The Physics of the Cosmic Microwave Background*. Cambridge University Press, Cambridge.

321. Nason, G. (2016). *wavethresh: Wavelets Statistics and Transforms*. R package version 4.6.8. Retrieved from https://CRAN.R-project.org/package=wavethresh.

322. Nason, G. P. and Sachs, R. v. (1999). Wavelets in time-series analysis. *Philosophical Transactions of the Royal Society A*, 357(1760):2511–2526.

323. Novikov, D., Schmalzing, J., and Mukhanov, V. (2000). On non-Gaussianity in the cosmic microwave background. *Astronomy & Astrophysics*, 364(1):17–25.

324. Nze, P. A. and Doukhan, P. (2004). Weak dependence: Models and applications to econometrics. *Econometric Theory*, 20(6):995–1045.

325. Ogden, T. (1997). *Essential Wavelets for Statistical Applications and Data Analysis*. Birkhäuser, Boston.

326. Olenko, A. (2006). Tauberian theorems for random fields with an OR spectrum. I. *Theory of Probability and Mathematical Statistics*, 73:135–149.

327. Olenko, A. (2013). Limit theorems for weighted functionals of cyclical long-range dependent random fields. *Stochastic Analysis and Applications*, 31(2):199–213.

328. Olsen, L. (1995). A multifractal formalism. *Advances in Mathematics*, 116(1):82–196.

329. Olsen, L. (2002). Empirical multifractal moment measures and moment scaling functions of self-similar multifractals. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 133(3), pages 459–485. Cambridge University Press, Cambridge.

330. Palma, W. and Chan, N. H. (2005). Efficient estimation of seasonal long-range-dependent processes. *Journal of Time Series Analysis*, 26(6):863–892.

331. Pando, J., Valls-Gabaud, D., and Fang, L.-Z. (1998). Evidence for scale-scale correlations in the cosmic microwave background radiation. *Physical Review Letters*, 81(21):4568.

332. Parisi, G. and Frisch, U. (1985). On the singularity structure of fully developed turbulence. In Ghil, M., Benzi, R., and Parisi, G., editors, *Turbulence and Predictability in Geophysical Fluid Dynamics*, pages 84–88. Elsevier, New York.

333. Paul, W. and Baschnagel, J. (1999). *Stochastic Processes: From Physics to Finance*. Springer-Verlag Berlin.

334. Pearson, K. (1902). On the mathematical theory of errors of judgement, with special reference to the personal equation. *Philosophical Transactions of the Royal Society A*, 198(300-311):235–299.

335. Peccati, G. and Taqqu, M. S. (2011). *Wiener Chaos: Moments, Cumulants and Diagrams: A Survey with Computer Implementation*. Springer-Verlag, Mailand.

336. Peltier, R. and Véhel, J. (1995). Multifractional Brownian motion: definition and

preliminary results. Technical Report 2645, Institut National de Recherche en Informatique et en Automatique, Le Chesnay Cedex, France.

337. Penzias, A. A. and Wilson, R. W. (1965). A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, 142:419–421.

338. Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.

339. Pietroni, M. (2009). Non-equilibrium in cosmology. *The European Physical Journal Special Topics*, 168:149–177.

340. Pipiras, V. and Taqqu, M. S. (2017). *Long-Range Dependence and Self-Similarity*. Cambridge University Press, Cambridge.

341. Planck Collaboration, Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., et al. (2020). Planck 2018 results - VII. Isotropy and statistics of the CMB. *Astronomy & Astrophysics*, 641(A7):61.

342. Planck Satellite (2021). The Cosmic Microwave Background. http://planck.cf.ac.uk/science/cmb. accessed 21 September 2021.

343. Planck Science Team (2021). Planck. https://www.cosmos.esa.int/web/planck. accessed 21 September 2021.

344. Porcu, E., Alegria, A., and Furrer, R. (2018). Modeling temporally evolving and spatially globally dependent data. *International Statistical Review*, 86(2):344–377.

345. Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association*, 85(410):338–344.

346. Price-Whelan, A. M., SipHocz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., et al. (2018). The astropy project: Building an open-science project and status of the v2.0 core package. *The Astronomical Journal*, 156(3):123.

347. Priestley, M. B. (1996). Wavelets and time-dependent spectral analysis. *Journal of Time Series Analysis*, 17(1):85–103.

348. R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

349. Racca, G. D., Laureijs, R., Stagnaro, L., Salvignol, J.-C., Alvarez, J. L., Criado, G. S., et al. (2016). The Euclid mission design. In MacEwen, H. A., Fazio, G. G., Lystrup, M., Batalha, N., Siegler, N., and Tong, E. C., editors, *Space Telescopes and*

Instrumentation 2016: Optical, Infrared, and Millimeter Wave, volume 9904, pages 235–257. International Society for Optics and Photonics, SPIE Digital Library.

350. Ray, B. K. (1993). Long-range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model. *International Journal of Forecasting*, 9(2):255–269.

351. Regan, D. and Munshi, D. (2015). Principal components of CMB non-Gaussianity. *Monthly Notices of the Royal Astronomical Society*, 448(3):2232–2244.

352. Rényi, A. (1959). On the dimension and entropy of probability distributions. *Acta Mathematica Academiae Scientiarum Hungarica*, 10(1-2):193–215.

353. Rényi, A. (1961). On measures of entropy and information. In Neyman, J., editor, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, volume 4.1, pages 547–561. Statistical Laboratory of the University of California, Berkeley.

354. Rényi, A. (1965). On the foundations of information theory. *Review of the International Statistical Institute*, 33(1):1–14.

355. Rényi, A. (1970). *Probability Theory*. Elsevier Science Publishers B.V. (North-Holland), Amsterdam.

356. Restrepo, J. E., Ruzhansky, M., and Suragan, D. (2021). Explicit solutions for linear variable-coefficient fractional differential equations with respect to functions. *Applied Mathematics and Computation*, 403:126177.

357. Rhodes, R. and Vargas, V. (2010). Multidimensional multifractal random measures. *Electronic Journal of Probability*, 15:241–258.

358. Ribeiro Jr, P. J., Diggle, P. J., Schlather, M., Bivand, R., and Ripley, B. (2020). *geoR: Analysis of Geostatistical Data*. R package version 1.8-1. Retrieved from https://CRAN.R-project.org/package=geoR.

359. Richard, F. J. P. (2018). Anisotropy of hölder gaussian random fields: characterization, estimation, and application to image textures. *Statistics and Computing*, 28(6):1155–1168.

360. Riedi, R. (1995). An improved multifractal formalism and self-similar measures. *Journal of Mathematical Analysis and Applications*, 189(2):462–490.

361. Riedi, R. H. (2002). Multifractal processes. In Doukhan, P., Oppenheim, G., and Taqqu, M., editors, *Theory and Applications of Long-Range Dependence*, pages 625–716. Birkhäuser, Basel.

362. Riedi, R. H. and Véhel, J. L. (1997). Multifractal properties of TCP traffic: a numerical study. Research report no. 3129, INRIA.

363. Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., and Scolnic, D. (2019). Large Magellanic Cloud Cepheid standards provide a 1% foundation for the determination of the Hubble constant and stronger evidence for physics beyond ΛCDM. *The Astrophysical Journal*, 876(1):85.

364. Rio, E. (2017). *Asymptotic Theory of Weakly Dependent Random Processes*. Springer, Berlin.

365. Robinson, P. M. (1994a). Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association*, 89(428):1420–1437.

366. Robinson, P. M. (1994b). Semiparametric analysis of long-memory time series. *The Annals of Statistics*, 22(1):515–539.

367. Robinson, P. M. (1994c). Time series with strong dependence. In Sims, C. A., editor, *Advances in Econometrics, Sixth World Congress*, pages 47–95. Cambridge University Press, Cambridge.

368. Robinson, P. M. (1995a). Gaussian semiparametric estimation of long-range dependence. *The Annals of Statistics*, 23(5):1630–1661.

369. Robinson, P. M. (1995b). Log-periodogram regression of time series with long-range dependence. *The Annals of Statistics*, 23(3):1048–1072.

370. Robinson, P. M. (2020). Spatial long-memory. *Japanese Journal of Statistics and Data Science*, 3:243–256.

371. Robson, B. A. (2019). Introductory chapter: Standard model of cosmology. In Robson, B. A., editor, *Redefining Standard Model Cosmology*. IntechOpen, London.

372. Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America*, 42(1):43–47.

373. Rosenblatt, M. (2015). Short Range and Long Range Dependence. In Dawson, D., R., K., Ould Haye, M., B., S., and Zhao, Y., editors, *Asymptotic Laws and Methods in Stochastics*, pages 283–294. Springer, New York.

374. Ruiz-Medina, M. D., Angulo, J. M., and Anh, V. V. (2008). Multifractality in space-time statistical models. *Stochastic Environmental Research and Risk Assessment*, 22:81–86.

375. Salat, H., Murcio, R., and Arcaute, E. (2017). Multifractal methodology. *Physica A: Statistical Mechanics and its Applications*, 473:467–487.

376. Samorodnitsky, G. (2007). *Long Range Dependence*. Now Publishers Inc, Hanover.

377. Schlather, M., Malinowski, A., Oesting, M., Boecker, D., Strokorb, K., Engelke, S., et al. (2019). *RandomFields: Simulation and Analysis of Random Fields*. R package version 3.3.6. Retrieved from https://cran.r-project.org/package=RandomFields.

378. Seneta, E. (1976). *Regularly Varying Functions*. Springer, Berlin.

379. Serroukh, A., Walden, A. T., and Percival, D. B. (2000). Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series. *Journal of the American Statistical Association*, 95(449):184–196.

380. Sheng, H., Chen, Y., and Qiu, T. (2011). *Fractional Processes and Fractional-Order Signal Processing: Techniques and Applications*. Springer, London.

381. Shimotsu, K. and Phillips, P. C. B. (2005). Exact local Whittle estimation of fractional integration. *The Annals of Statistics*, 33(4):1890–1933.

382. Shirazi, E. and Doosti, H. (2015). Multivariate wavelet-based density estimation with size-biased data. *Statistical Methodology*, 27:12–19.

383. Sloan, I. H. and Wendland, H. (2009). Inf-sup condition for spherical polynomials and radial basis functions on spheres. *Mathematics of Computation*, 78(267):1319–1331.

384. Sloan, I. H. and Womersley, R. S. (2000). Constructive polynomial approximation on the sphere. *Journal of Approximation Theory*, 103(1):91–118.

385. Smidt, J., Amblard, A., Cooray, A., Heavens, A., Munshi, D., and Serra, P. (2010). A measurement of cubic-order primordial non-Gaussianity (g_{NL} and τ_{NL}) with WMAP 5-year data. https://arxiv.org/abs/1001.5026.

386. Smith, H. F. (1938). An empirical law describing heterogeneity in the yields of agricultural crops. *The Journal of Agricultural Science*, 28(1):1–23.

387. Smoot, G. F. (2007). Nobel lecture: Cosmic microwave background radiation anisotropies: Their discovery and utilization. *Reviews of Modern Physics*, 79(4):1349.

388. Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., et al. (2003). First-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: determination of cosmological parameters. *The Astrophysical Journal Supplement Series*, 148(1):175.

389. Spodarev, E. (2013). *Stochastic Geometry, Spatial Statistics and Random Fields: Asymptotic Methods*. Springer-Verlag, Berlin.

390. Starck, J. L., Aghanim, N., and Forni, O. (2004). Detection and discrimination

of cosmological non-gaussian signatures by multi-scale methods. *Astronomy & Astrophysics*, 416(1):9–17.

391. Stoev, S. A. and Taqqu, M. S. (2006). How rich is the class of multifractional Brownian motions? *Stochastic Processes and their Applications*, 116(2):200–221.

392. Struzik, Z. (2000). Determining local singularity strengths and their spectra with the wavelet transform. *Fractals*, 8(2):163–179.

393. Taqqu, M. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31(4):287–302.

394. Taqqu, M. S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Zeitschrift Fur Wahrscheinlichkeitstheorie Und Verwandte Gebiete*, 50(1):53–83.

395. Telesca, L., Lovallo, M., Mammadov, S., Kadirov, F., and Babayev, G. (2015). Power spectrum analysis and multifractal detrended fluctuation analysis of earth's gravity time series. *Physica A: Statistical Mechanics and its Applications*, 428:426–434.

396. Todorovic, P. (2012). *An Introduction to Stochastic Processes and Their Applications*. Springer-Verlag, New York.

397. Tsai, H., Rachinger, H., and Lin, E. M. (2015). Inference of seasonal long-memory time series with measurement error. *Scandinavian Journal of Statistics*, 42(1):137–154.

398. Turiel, A. and Parga, N. (2000). Multifractal wavelet filter of natural images. *Physical Review Letters*, 85(15):3325.

399. Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

400. Van Kampen, N. (1992). *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam.

401. Véhel, J. L. and Mignot, P. (1994). Multifractal segmentation of images. *Fractals*, 2(3):371–377.

402. Véhel, J. L. and Riedi, R. (1997). Fractional Brownian motion and data traffic modeling: The other end of the spectrum. In Véhel, J. L., Lutton, E., and C., T., editors, *Fractals in Engineering*, pages 185–202. Springer, London.

403. Viana, R. L., Da Silva, E. C., Kroetz, T., Caldas, I. L., Roberto, M., and Sanjuán, M. A. F. (2011). Fractal structures in nonlinear plasma physics. *Philosophical Transactions of the Royal Society A*, 369(1935):371–395.

404. Wang, H. and Sloan, I. H. (2017). On filtered polynomial approximation on the

sphere. *Journal of Fourier Analysis and Applications*, 23(4):863–876.

405. Wang, Y. G. (2021). CMBProbe, the Python package for generating AC discrepancy maps. https://github.com/wangyg19/CMBProbe. accessed 21 September 2021.

406. Wang, Y. G., Le Gia, Q. T., Sloan, I. H., and Womersley, R. S. (2017). Fully discrete needlet approximation on the sphere. *Applied and Computational Harmonic Analysis*, 43(2):292–316.

407. Wang, Y. G., Sloan, I. H., and Womersley, R. S. (2018). Riemann localisation on the sphere. *Journal of Fourier Analysis and Applications*, 24(1):141–183.

408. Weinberg, S. (2008). *Cosmology.* Oxford University Press, Oxford.

409. Whitcher, B. (2004). Wavelet-based estimation for seasonal long-memory processes. *Technometrics*, 46(2):225–238.

410. Whitcher, B. (2020). *waveslim: Basic Wavelet Routines for One-, Two-, and Three-Dimensional Signal Processing.* R package version 1.8.2. Retrieved from https://CRAN.R-project.org/package=waveslim.

411. Wishart, J. R. (2019). *mwaved: Multichannel wavelet deconvolution with additive long memory noise.* R package version 1.1.7. Retrieved from https://CRAN.R-project.org/package=mwaved.

412. Woodward, W. A., Cheng, Q. C., and Gray, H. L. (1998). A k-factor GARMA long-memory model. *Journal of Time Series Analysis*, 19(4):485–504.

413. Yadav, A. P. S., Komatsu, E., and Wandelt, B. D. (2007). Fast estimator of primordial non-Gaussianity from temperature and polarization anisotropies in the cosmic microwave background. *The Astrophysical Journal*, 664(2):680.

414. Yadrenko, M. I. (1983). *Spectral Theory of Random Fields.* Optimization Software, New York.

415. Yajima, Y. (1985). On estimation of long-memory time series models. *Australian Journal of Statistics*, 27(3):303–320.

416. Yajima, Y. (1996). Estimation of the frequency of unbounded spectral densities. In *Proceedings of the Business and Economic Statistical Section*, pages 4–7. American Statistical Association.

417. Yu, Z.-G., Anh, V., and Lau, K.-S. (2001). Measure representation and multifractal analysis of complete genomes. *Physical Review E*, 64(3):031903.

418. Zonca, A., Singer, L., Lenz, D., Reinecke, M., Rosset, C., Hivon, E., et al. (2019). healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python. *Journal of Open Source Software*, 4(35):1298.