An analysis of non-Gaussian models predicting anomalies in the cosmic microwave background

Markus Bjørklund

Thesis submitted for the degree of
Master of Science in Astronomy

Institute of Theoretical Astrophysics
University of Oslo
15.06.2021

Abstract

The current paradigm of a 6 parameter ΛCDM cosmology with single field slow-roll inflation predicts a perfectly Gaussian and isotropic distribution for the Cosmic Microwave Background (CMB) temperature field. However, several statistical anomalies have been reported in the CMB, calling into question whether other models could more accurately describe the statistical properties of the CMB.

In this thesis, we investigate a model proposed by Hansen et al. [1] reproducing the aforementioned anomalies. We work in the basis of wavelets, Spherical Mexican Hat Wavelets (SMHW) and spherical standard needlets, yielding several advantages over the traditional spherical harmonics in the presence of incomplete sky coverage.

We simulate a large number of CMB maps, both Gaussian and non-Gaussian realizations as per the model, and estimate the wavelet 4-point correlation function, the trispectrum. The use of the trispectrum is motivated by the presence of scale-dependent, non-Gaussian, g_{NL}-like terms in the model.

We propose a blind test of non-Gaussianity, a χ^2 statistic testing the Gaussian null hypothesis on the simulated maps. We also propose a method for testing the model hypothesis, an estimator $\hat{\alpha}$ quantifying the amount of contribution from a non-Gaussian term in the CMB maps preferred by the data.

In addition, we compare our results with those of the Planck experiment, by applying the same framework to the 2015 Planck FFP8.1 Monte Carlo simulation data. For the spherical standard needlets, the results are consistent with the Gaussian null hypothesis. However, a detection is seen in the Spherical Mexican Hat Wavelets maps, disfavoring the Gaussian null hypothesis at 0.63% significance.

Acknowledgments

First and foremost, I would like to thank my supervisor Frode K. Hansen, for being more helpful and available than any student could possibly hope for. He has probably answered a Fermi estimate of about 1000 emails of questions and without his help, this thesis would not be remotely possible. I would also like to thank my partner Elin and my family for being incredibly supportive, and always offering to help in any way they could. Lastly, a big thanks to all my fellow master's students at the institute, for always keeping spirits and morale high even during the Covid pandemic, and for valuable insight, academic or otherwise.

I acknowledge the use of the computing clusters at Sigma2, the Norwegian e-infrastructure for research and education, as well as the use of the HEALPix software [2], available at `http://healpix.sourceforge.net`.

Contents

List of Figures

Chapter 1

(A brief) History of cosmology

1.1 From cave paintings to celestial mechanics

For as long as mankind has been able to look up at the twinkling stars in the night sky, people have contemplated their place in the Universe and how it is all connected. There are depictions of constellations, presumably to record dates, in cave paintings from as early on as around 40 000 years ago [3], showing that even the early humans observed the sky in an attempt to understand the world around them. There have been many attempts to describe the Universe through the ages, often intertwined with astrology or religion, with each model making a small step towards modern astronomy as we know it today.

Modern astronomy took flight with the introduction of the heliocentric system, sparking the Copernican Revolution, named after Nicolaus Copernicus who proposed the model. The heliocentric system was later firmly cemented in theory by another important figure, namely Isaac Newton. Newton provided us with a universal force of gravitation, a relation between force and acceleration, among many other contributions.

Some problems, however, remained which could not be explained by Newtonian gravity and mechanics, among them the perihelion precession of Mercury. Mercury's peculiar orbit did not seem to fit with the Newtonian model of the movement of the stars and planets. Almost 200 years after Newton's death, Mercury's strange orbit would be explained by arguably the most famous scientist of all time, Albert Einstein.

1.2 Modern cosmology

Cosmology is "the scientific study of the large scale properties of the Universe as a whole." [4]. In other words, cosmology defines the study of the very, very large in the Universe. A foundational principle for all modern cosmology is what we call "The cosmological principle". This principle states that overall, wherever you look, the Universe will be pretty much the same. It will look the same in every direction, a property called isotropy, and it will look the same at every point, homogeneity. This principle can not be proven per se but has enormous empirical backing, and we have no

reason to doubt its validity. Subsequently, something differing based on the direction can be called an anisotropy, or an inhomogeneity if it differs based on distance.

If one were to determine a point of origin for this modern branch of astronomy, many would look to Albert Einstein and his theory of special [5] and general [6] relativity. In the past, it was "common knowledge" that waves had to propagate through a medium, and if space was empty, how could light move and gravity act? One proposed solution was the Aether, a backdrop of the Universe on which everything propagated. Einstein, however, introduced the concept of space-time. Space and time were intertwined in a wobbly fabric, and gravity was no longer a classical force. Matter and light were simply falling in straight lines "on" a curved space, and matter, or more generally energy, would curve space. Much like marbles rolling in circles on a trampoline that has a bowling ball in the center, although the analogy leaves us two dimensions short. So while Einstein's field equations may look complicated, this is the simple idea they portray: energy content tells space how to curve, and the curvature of space tells matter (or light) how to move.

1.2.1 Expansion and the ridicule that stuck

Einstein's theory of relativity found tremendous experimental success and made him world-famous overnight as his predictions were corroborated to a very high degree of precision. There was now a theoretical framework in place for the start of modern cosmology. And then another stunning observation was made, by Edwin P. Hubble. Previously, one thought that our own Milky Way was an "island universe", and that the so-called "spiral nebulae" observed through telescopes were within our own galaxy. Hubble's findings were formally published in 1926 [7] and 1929 [8], showing that one such spiral nebula, our neighboring galaxy Andromeda, was indeed a galaxy in its own right outside the boundaries of the Milky Way, among other discoveries.

Now, with the framework of Einstein's theory of relativity, and the ability to observe galaxies and other objects outside our own Milky Way, the stage was set for tackling the biggest problem of them all: the evolutionary history of the Universe. And one such solution, with an expanding universe, was proposed by a Belgian priest by the name of George Lemaître. To explain the red-shift of the observed spiral nebulae, now known as galaxies, he proposed the expanding universe solution and working backward to a logical conclusion, the beginning of the Universe as a single point. He called this theory the "Primeval Atom" [9]. One might think this seems perfectly reasonable, but at the time, most people believed in an unchanging, eternal universe. In actuality, the term "Big-Bang" was coined, arguably with derision, by Fred Hoyle, another astronomer who starkly opposed the theory until his death.

The Big-Bang theory however was overwhelmingly supported by observational data, especially from comprehensive observations by Hubble. The Universe was indeed expanding, and should therefore have a beginning.

1.2.2 Precision cosmology and the CMB

The nail in the coffin for the competing theory to the Big-Bang theory, the steady-state theory to which aforementioned Fred Hoyle subscribed, was to come in the 1960s. At Bell Laboratories in New Jersey, two American radio astronomers named Robert Wilson and Arno Penzias were originally studying microwave signals from the Milky Way, when they stumbled upon a faint background noise. They could not deduce the origins of the signal, after exhausting nearly every possible source, including scrubbing pigeon droppings off the antenna and relocating the pigeons that nested there. The only remaining conclusion was that the observed signal was a real one. Concurrently, at Princeton University, the three cosmologists Robert H. Dicke, Jim Peebles, and David Wilkinson were theorizing that just such radiation should be detectable as a remnant from the Big-Bang event. Wilson and Penzias learned of this fact and made contact. They published their findings in two companion papers, Wilson and Penzias on the discovery itself [10], and Dicke, Peebles, Wilkinson and Roll on the implications for the history of the Universe [11]. Together, they had discovered Cosmic Microwave Background (CMB) radiation, the most ancient light in the Universe and the key to understanding the birth of the cosmos.

1.2.3 The birth certificate of the Universe

Why is the CMB important, you might ask? In the very early Universe, before the CMB was released, the Universe was a "very hot and dense soup". Matter and radiation were so tightly coupled that no light could travel very far, and the whole Universe was opaque. Then, 380 000 years after the Big Bang, the matter and radiation decoupled, an event we call "The surface of last scattering", as it was the last time this light scattered off the matter in the hot, dense primordial plasma. And unlike many other types of information, it has streamed towards us relatively unaltered since. Therefore, the light that could start its journey towards us also traces the distribution of matter at that time. Now, 380 000 years may seem like a long time, but if the Universe today was an 80-year-old person, the CMB is a picture of an 18-hour-old baby. The infant picture lets us deduce the anatomy of the baby.

1.2.4 Inflation

While the detection of the CMB fortified the Big-Bang theory as our best model of the Universe, the new observations did not come without their own set of mysteries. Especially two problems were concerning. First off, wherever you looked, the CMB seemed to be pretty much the same. The Universe looked the same everywhere, in every direction. But how could this be, if the Universe underwent superluminal, "faster than light", expansion, such that distant regions could never "speak" to each other? We say that the regions could not have been "causally connected" and the problem was dubbed the "Horizon problem".

Secondly, the Universe seemed to be immensely flat; that is, the curvature of the Universe was extremely small. This was concerning because it was a so-called fine-

tuning problem: a small curvature at the beginning would only amplify over time, such that if the Universe is extremely flat today, it must have been EXTREMELY flat at the beginning. Cosmologists are rarely comfortable with relying on such unlikely coincidences, and the problem was named the "Flatness problem". A solution was needed and a solution was posed, called inflation [12]. Inflation theory postulates a period of exponential expansion of space, around $10^{-34}s$ after the Big Bang, ending almost instantaneously. In that microscopic period of time, the Universe expanded in size by at least 60 e-folds, meaning it increased by a factor of at least e^{60}[13]. To fully grasp the staggering size of this number, it is a 10 with 25 trailing zeros, or coincidentally about 10 times the upper estimate on the number of stars in the Universe [14]. Then, causally connected patches could be in thermal equilibrium before inflation and become causally disconnected and very, very distant to each other in an instant, and the Universe would look very similar wherever you looked. Not unlike the fact that one can find similarities in languages from different corners of the world because they share a common origin, although to accurately portray the uniformity of the CMB, only one word in a hundred thousand could differ. Inflation also solved the flatness problem, as any initial curvature of space would be blown up to such a size as to appear flat. Very much like the earth appearing flat when you stand on it, while really being curved (REWRITE). The theory of inflation has been a hugely successful theory to this day.

1.2.5 From quantum seeds to galaxies

The theory of inflation was originally postulated as a way out of the Horizon and Flatness Problems, among others, but it was soon revealed that the theory also bore additional, unexpected fruit. While the Universe was very homogeneous, isotropic, and flat in the grand scheme of things, there still obviously existed anisotropies; different places in the sky had differing densities. We were, after all, living in such a place; a galaxy whose density is far higher than the surrounding void. Inflation provides a mechanism for the tiniest quantum fluctuations, in a field called the "Inflaton", to be "stretched" into astronomical scales, laying the foundation for the formation of the planets, stars, and galaxies we see today. These fluctuations are the primordial seeds of structure in the Universe and give rise to the temperature fluctuations in the Cosmic Microwave Background radiation.

1.2.6 The importance of measurement

It soon became apparent that quantifying the CMB radiation could provide us with answers to major outstanding questions about the Universe. Different models of the Universe predict specific power spectra of the temperature fluctuations in the CMB. This spectrum, as measured by the Planck satellite, can be seen in figure 1.1, where the data points are shown as red dots with error bars, the predicted spectrum from the current best-fit model is the solid green line, and the shaded green area represents

[2] https://www.esa.int/ESA_Multimedia/Images/2013/03/Planck_Power_Spectrum#.X4m9sRFiOSk.link

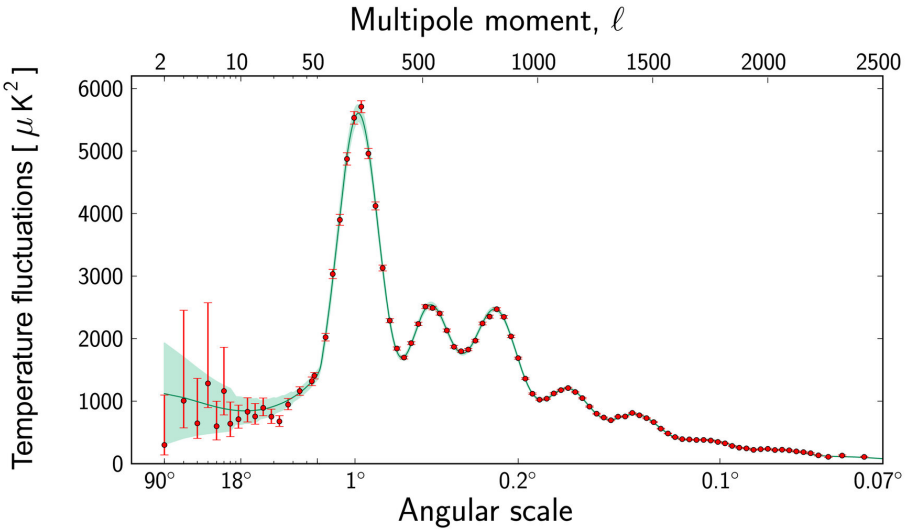


Figure 1.1: The 2013 Planck CMB temperature angular power spectrum. Original image credit [15], direct access from ESA[2]

the standard deviation. These spectra in turn depend on some important cosmological parameters, and the current paradigm consists of a 6-parameter model [16]. For example, the location of the first peak is highly sensitive to the curvature of the Universe, and the measurement shows that the Universe is extremely flat. The ratio between the first and the second peak gives us information about how much of the Universe is made up of baryons, in the form of the parameter Ω_b, while the other peaks contain information about dark energy and dark matter. The tilt of the spectrum is governed by the parameter n_s, known as the spectral index, and the primordial amplitude A_s governs the height, which tells us among other things about the optical depth to "reionization", a period when the Universe became ionized again. From the power spectrum, one can also measure the rate of expansion, the Hubble constant H_0, in turn letting us estimate the age of the Universe. Many more parameters can be derived from the original 6. In short, the shape and size of the power spectrum contain a wealth of information about the Universe.

1.2.7 Telescopes in space

It was clear that precise measurements had to be made of the CMB, and ground-based observations were coming up short. We would need a telescope in space to rid the results of interference from the atmosphere and other environmental factors. The first of these endeavors would be the Cosmic Background Explorer or COBE.

COBE

NASA took the first step towards precision measurements of the CMB, and COBE was launched on November 18, 1989. COBE was a space-based telescope comprised of three main instruments:

FIRAS (Far InfraRed Absolute Spectrophotometer), whose tasks were to measure the spectrum of the electromagnetic radiation of the CMB and to observe dust and line emission from the galaxy.

DIRBE (Diffuse InfraRed Background Explorer), whose main objective was to measure Cosmic Infrared Background (CIB) radiation as well as other sources of radiation contributing to the received signal.

DMR (Differential Microwave Radiometer), which was searching for fluctuations in the temperature in the CMB, the so-called anisotropies.

COBE mapped the spectrum and the anisotropies of the CMB at an angular resolution of 7 degrees, while the diffuse galactic emission was observed with a resolution of 0.7 degrees [17]. After only two years of observation, the results from COBE had already revealed the CMB as the most perfect blackbody radiation ever measured, with a temperature of 2.726 ± 0.010 K [18]. The mission was a striking success, and the Big-Bang theory was now firmly backed up by observations. During its total four years of observation, COBE had yielded a veritable gold mine of cosmological information and constraints on important parameters in cosmological models.

WMAP

The next step was to increase the precision of measurement even further, and in 1995 a new telescope was proposed; the Wilkinson Microwave Anisotropy Probe(WMAP). WMAP's mission, again led by NASA, was to improve on the measurements of COBE and achieve high precision measurements of the temperature fluctuations in the CMB. WMAP would improve tremendously on the resolution of its predecessor COBE, with sensitivity and angular resolution improving by a factor of 45 and 33 respectively, now boasting an angular resolution of around 0.2 degrees and a sensitivity to polarization [19]. WMAP began observations in 2001, and the results were truly revolutionary. During its 9 year period of observation it had, among many other results, derived the age of the Universe to be 13.772 ± 0.059 billion years, determined the curvature to be completely flat within 0.95 percent, and improved constraints on the standard 6 parameter cosmology by a factor of 68 000 [20]. WMAP truly brought us into the age of precision cosmology.

Planck

Finally, the most ambitious CMB experiment to date was to take place. This time, the European Space Agency (ESA) took the lead in creating a new telescope with unprecedented precision. The telescope was named Planck, after the famous German physicist. The Planck satellite launched on 14 May 2009, with the goal of surveying at such high precision that the resulting images would only be limited by cosmic variance,

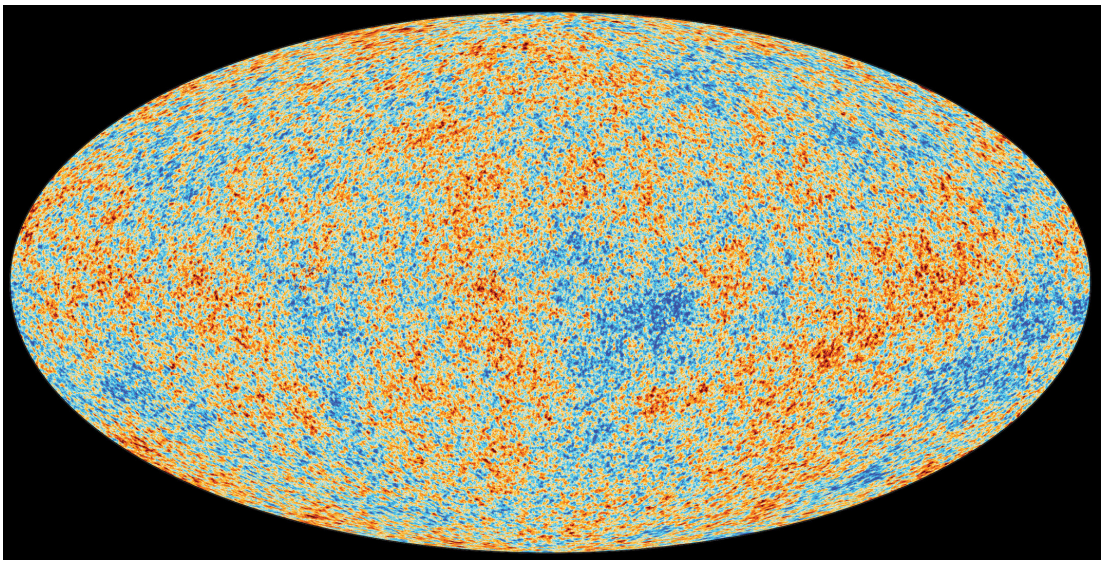


Figure 1.2: The full CMB SMICA sky map, with 3% of the sky filled in. Original image credit [22], direct access from ESA[4]

the most accurate measurement achievable. Planck contained two instruments, the Low Frequency Instrument (LFI), whose detectors operated in the bandwidth 30-70 GHz, and the High Frequency Instrument (HFI), with detectors in the range 100-857 GHz. The Planck satellite was situated at the second Lagrange point, 1.5 million km from Earth [21]. Planck remains one of the principal sources of cosmological information today and provided us with the two famous pictures, the power spectrum previously shown in figure 1.1, and the full-sky map of the anisotropies of the CMB shown in figure 1.2.

1.3 Where we are now

All the information about the Universe in the following sections has been inferred from the analysis of the CMB, and especially from the latest results from Planck [16, 23, 24, 25, 26]. The fact that the following is only a fraction of the cosmological information contained in the CMB shows just how important this ancient radiation is for our understanding of the Universe.

[4] https://www.esa.int/ESA_Multimedia/Images/2018/07/Planck_s_view_of_the_cosmic_microwave_background#.X4m9yiXQx9w.link

[6] https://www.esa.int/ESA_Multimedia/Images/2013/03/Planck_history_of_Universe

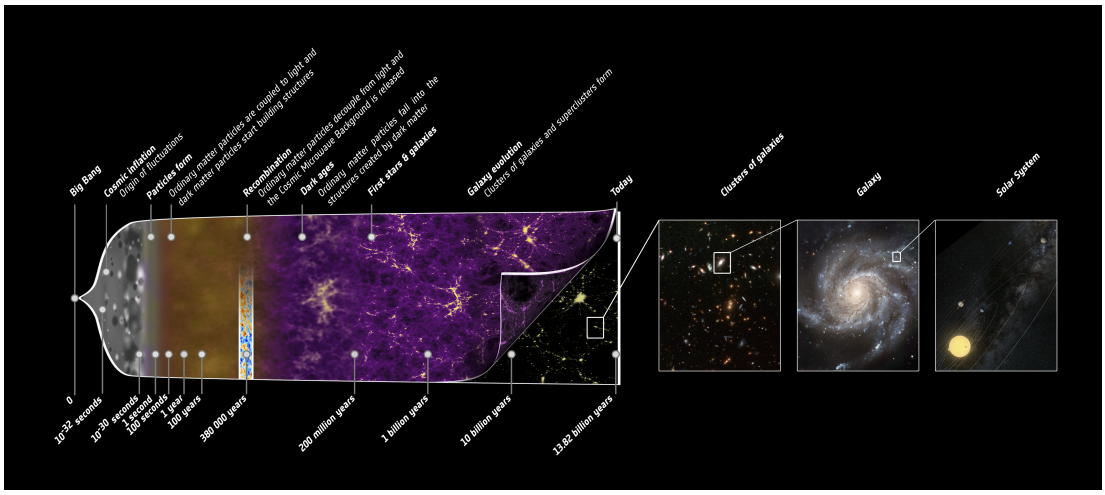


Figure 1.3: The schematic evolution of the Universe. Image credit: ESA - C. Carreau[6].

1.3.1 The stuff in the Universe

The current widely accepted model of the Universe, in accordance with the Big-Bang theory (WITHIN THE FRAMEWORK OF?), is called ΛCDM. In the ΛCDM model, the Universe is comprised of three components: Dark energy Λ (originally Einstein's "cosmological constant"), matter (Cold Dark Matter (CDM) and "normal" matter) and radiation. The distribution of these three components is around 31.1% matter, of which only around 4.9% is the visible matter familiar to us, which is somewhat haphazardly referred to as baryons in cosmology. The remaining 26.2 % of matter is what we call Dark Matter. Dark matter does not strongly interact electromagnetically, so it is hard or maybe even impossible to observe directly. However, we can infer its existence by observing its gravitational pull on its surroundings. The dominating energy contribution in today's universe is Dark energy, which contributes around 68.9 %. This is a mysterious energy component of the Universe we know very little about, responsible for driving the accelerated expansion of the Universe. Observant readers might by now have realized that this all adds up to 100% , so what about radiation? This is due to the fact that radiation today contributes only a tiny amount to the total energy content, below 0.001%.

1.3.2 A spanner in the works

While the ΛCDM model has been very successful, there is still debate over just how well the observed CMB fits into the predicted picture of a perfectly isotropic, homogeneous distribution. Several statistical anomalies in the fluctuations were detected, which will be described in more detail later in this thesis. The anomalies were first detected by WMAP, and later confirmed by Planck, thus eliminating instrumental properties as the main cause. Significant effort has been put into analyzing these anomalies, and

most anomalies appear similar at very different frequencies, which makes them hard to explain in terms of foreground radiation. The question then remains, are these anomalies a statistical fluke, or are they a manifestation of some unknown physics at work? Do we need a new model of the Universe which more naturally explains the departures from statistical isotropy in the CMB? The core of this thesis can thus be summed up in the following question: What kind of non-standard model can reproduce the anomalies currently seen in the CMB?

Chapter 2

Theory

To properly connect trace the observed CMB fluctuations to a model of the cosmos, one must first know the time evolution of the universe. The framework which enables us to do this is Einstein's general theory of relativity, which will be our starting point. Unless otherwise stated, the information in this chapter is taken from the two books "Spacetime and geometry" by Sean Carroll [27] and "Modern Cosmology" by Scott Dodelson [28].

2.1 Einstein's field equations

Einstein provided the tensor equations for the evolution of the universe, which read

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{2.1}$$

with the Einstein tensor $G_{\mu\nu}$ given as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \tag{2.2}$$

where $R_{\mu\nu}$ is the Ricci tensor, R the Ricci scalar, Λ the cosmological constant, $T_{\mu\nu}$ the energy-momentum tensor, G is the gravitational constant and $g_{\mu\nu}$ is the metric tensor.

2.2 Solutions to Einstein's field equations

Under the assumption of the cosmological principle, we obtain the Friedmann-Lemaître-Robertson-Walker (FLRW) line-element

$$ds^2 = -dt^2 + a^2(t) \left[\frac{dr^2}{1 - \kappa r^2} + r^2 d\Omega^2 \right], \tag{2.3}$$

where $a(t)$ is the scale factor and κ is proportional to the curvature parameter $k \in \{-1, 0, 1\}$ for open, flat or closed geometry respectively. Now under the assumption that the energy-momentum tensor is given as a perfect fluid

$$T_{\mu\nu} = (p + \rho)\, U_\mu U_\nu + p g_{\mu\nu}, \tag{2.4}$$

we have analytical solutions. This is exactly what Alexander Friedmann did, and solving for the 00-component along with the trace of the field equations gives us the two Friedmann equations, in natural units $\hbar = c = 1$ and $\rho_{\text{tot}} = \rho + \rho_\Lambda = \rho + \frac{\Lambda}{8\pi G}$:

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho_{\text{tot}} - \frac{\kappa}{a^2} \tag{2.5}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho_{\text{tot}} + 3p_{\text{tot}}\right), \tag{2.6}$$

where H is the Hubble function and a dot denotes a time derivative. Along with the equation of state

$$w = \frac{p}{\rho}, \tag{2.7}$$

this system is solvable for the time evolution of the universe, in the form of the scale factor a and the critical densities Ω.

2.3 (Single-field) Inflation

One of the simplest solutions to the Horizon and Flatness problems is a single field inflation model. An exponentially expanding solution to the Friedmann equations, $a \propto e^{Ht}$, is obtained through vacuum energy, with equation of state parameter $w = -1$. As inflation needs a mechanism to start and end, we model this driving mechanism, not as a cosmological constant, but a scalar field Φ, the Inflaton. Introduce a classical Inflaton field with a Lagrangian

$$\mathcal{L}_\Phi = \frac{1}{2}\partial^\mu\Phi\partial_\mu\Phi - V(\Phi), \tag{2.8}$$

and the energy momentum tensor is given as

$$T^{\mu\nu} = \frac{\delta}{\delta g^{\mu\nu}}\left(\sqrt{-g}\mathcal{L}_\Phi\right). \tag{2.9}$$

With a FLRW metric, we obtain

$$\rho_\Phi = T^{00} = \frac{1}{2}\dot{\Phi}^2 + V(\Phi) + \frac{(\nabla\Phi)^2}{2a^2} \tag{2.10}$$

$$p_\Phi = \frac{1}{3}T^i_i = \frac{1}{2}\dot{\Phi}^2 - V(\Phi) - \frac{(\nabla\Phi)^2}{6a} \tag{2.11}$$

where the last terms in each equation can be neglected due to homogeneity. Now if we have $\frac{1}{2}\dot{\Phi}^2 \ll |V(\Phi)|$, we obtain $\rho_\Phi \approx -p_\Phi$, or $w \approx -1$, as desired. This phase, where the potential is much larger than the kinetic term of the Inflaton field, is called "slow roll", shown in figure 2.1. This is analogous to a regular ball slowly rolling down a hill, and eventually obtaining more kinetic energy than potential gravitational energy,

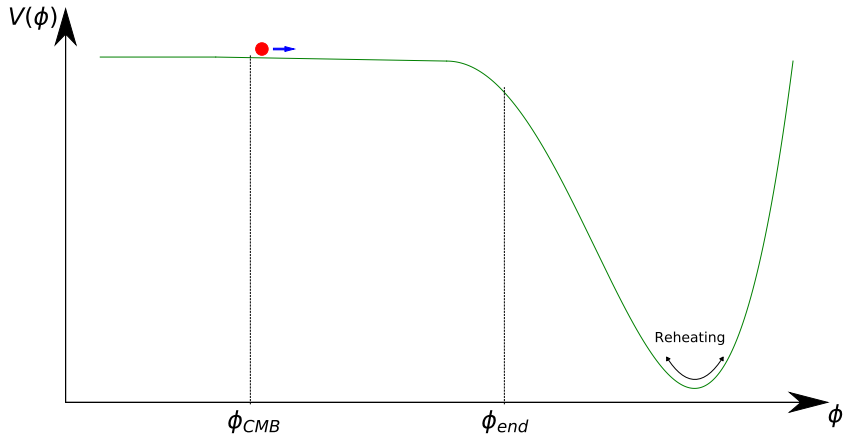


Figure 2.1: The Inflaton field, "slowly rolling" down its potential. The start and end of inflation are marked, the period where the quantum fluctuations responsible for the CMB temperature fluctuations are blown up. Reheating is shown schematically, as the Inflaton decays away at the bottom of the potential well.

hence the name. The slow-roll conditions are $\epsilon \ll 1$ and $|\eta| \ll 1$, with the slow-roll parameters given as

$$\epsilon = \frac{M_{\mathrm{Pl}}^2 V_\Phi^2}{2V^2} \tag{2.12}$$

$$\eta = \frac{M_{\mathrm{Pl}}^2 V_{\Phi\Phi}}{V}, \tag{2.13}$$

where $M_{\mathrm{Pl}} = \frac{1}{\sqrt{8\pi G}}$ is the reduced Planck mass, V_Φ is the potential and the subscript Φ denotes the first derivative (and $\Phi\Phi$ the second derivative) with respect to Φ. The CMB fluctuations are created by quantum fluctuations in the Inflaton field, which are blown up during inflation until inflation ends when the slow-roll condition is violated. The Inflaton decays into radiation during reheating, at the bottom of the potential well [29].

2.4 From inflation to the CMB

According to quantum field theory, the Inflaton field will experience quantum fluctuations during inflation. We can split the field into a zero-order homogeneous component ϕ_0 responsible for driving inflation, and perturbations around this value

$$\phi(t) = \phi_0(t) + \delta\phi(\mathbf{x}, t). \tag{2.14}$$

Since the Inflaton field is the dominating energy content in the universe during inflation, such a perturbation in the inflaton will result in perturbation in the energy-momentum tensor, $\delta T_{\mu\nu}$, and by Einstein's equations a perturbation in the metric, $\delta g_{\mu\nu}$. The wavelengths of these perturbations will experience exponential growth during inflation and may cross the Hubble horizon. The Hubble horizon differs from the concept of a particle horizon in the way that while the particle horizon describes the limit outside which two particles never could have or will be causally connected, the Hubble horizon denotes the limit where two particles can not be causally connected in the future (granted the Hubble horizon growth stays constant). If the perturbations do cross the Hubble horizon, i.e. the wavelength of the perturbations are larger than the Hubble horizon, they will be "frozen in", meaning the perturbation can not be affected by causal physics and the amplitude of the perturbation is preserved as it was at the time of crossing, thus the name "frozen". After the end of inflation, when the Hubble horizon grows faster than the expansion of space again, the frozen fluctuations may enter the Hubble horizon again, giving rise to primordial gravitational potential fluctuations, Φ_P. These fluctuations in the potential can be expanded in terms of a potential with Gaussian distributed fluctuations, Φ_G, plus higher order terms. For example, keeping up to third order terms we can write

$$\Phi_P(\mathbf{x}) = \Phi_G(\mathbf{x}) + f_{NL}\left(\Phi_G^2(\mathbf{x}) - \langle\Phi_G^2(\mathbf{x})\rangle\right) + g_{NL}\Phi_G^3(\mathbf{x}), \tag{2.15}$$

(and higher order terms) where $\Phi_G(\mathbf{x})$ is the linear Gaussian part of the primordial gravitational potential [30, 31, 32, 33, 34]. A perturbation in the gravitational potential will lead to perturbations in the matter density, through the Poisson equation

$$\nabla^2\Phi = 4\pi G\rho. \tag{2.16}$$

Furthermore, the perturbations in the gravitational potential will also create perturbations Θ in the photon distribution. Through solving the Boltzmann equation for photons with a perturbed metric, and using Line Of Sight (LOS) integration by Zeljak and Zaldarriaga [35], it can be shown that the harmonic transform of the photon perturbation, Θ_l, can then be related to the CMB fluctuations via the angular power spectrum,

$$C_l \propto \frac{2}{\pi}\int_0^\infty dk\, P(k)\Theta_l(k)^2, \tag{2.17}$$

where C_l is the angular power spectrum of the temperature fluctuations in the CMB, $P(k)$ is the primordial power spectrum of the temperature fluctuations, which is set up by inflation, k is the Fourier wave number, and l is a multipole in a spherical harmonic expansion. For a thorough derivation, see Dodelson [28].

2.5 The CMB

While the fluctuations in the CMB are an invaluable source of cosmological information, one must first understand what steps we need to take in order to extract this

information. In this section, I will outline the mathematical framework in which we will analyze the CMB radiation.

2.5.1 Spherical decomposition

The available data we have is the temperature of the CMB, measured on the sky. The temperature can be represented as a field

$$T(\theta, \phi) = T_{\mathrm{CMB}}(\theta, \phi) \left[1 + \Theta(\theta, \phi)\right], \tag{2.18}$$

where θ and ϕ represents angles in the sky, in galactic co-latitude and longitude, and Θ represents the anisotropy, I.E. deviation from the background value. The monopole is usually subtracted from the measurements, leaving only a measure of anisotropy.

It is standard practice to decompose the temperature field on the sky to spherical harmonics, such that an orthogonal basis is given by

$$a_{lm} = \int_{S_2} d\theta d\phi sin\left(\theta\right) T(\theta, \phi) Y^*_{lm}\left(\theta, \phi\right), \tag{2.19}$$

where then Y^*_{lm} is the spherical harmonic functions, and a_{lm} is the harmonic coefficients. The multipoles l relate to the angular scale (in radians) on the sky via

$$\theta = \frac{\pi}{l}. \tag{2.20}$$

The two representations of the anisotropy, in the spatial domain or in spherical harmonic space, is equivalent, and thus the temperature field can be recovered by the inversion

$$T(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y^*_{lm}. \tag{2.21}$$

For a completely isotropic field, the a_{lm} coefficients are uncorrelated and obey the relation

$$\langle a_{lm} a^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l, \tag{2.22}$$

where δ is the Kroenecker delta, C_l is the angular power spectrum and the angle brackets denote the *ensemble average*.

2.5.2 The angular power spectrum

The angular power spectrum, C_l, is the harmonic analogue of the two-point correlation function in real space. If the distribution of the fluctuations in the temperature field is Gaussian, the distribution is statistically fully described by the angular power spectrum with higher order moments, or correlation functions, being zero. In this way, the power spectrum gives the variance of the harmonic coefficients a_{lm} in a zero-mean field or

equivalently is a measure of the strength of the fluctuations as a function of scale l. From the relation shown in equation 2.22, it naturally follows that the optimal estimator for the angular power spectrum C_l is

$$\hat{C}_l = |a_{lm}|^2 . \tag{2.23}$$

In theory, this satisfies the relation $\langle \hat{C}_l \rangle = C_l$, that is that the ensemble average over the estimator equals the true value, C_l. However, there is one problem: We can not obtain the ensemble average over the estimator, as we only have one universe to measure. To bypass this problem, the isotropic properties of the spherical harmonics come to the rescue. Instead of averaging over an ensemble of universes, we can average over azimuthal directions m, leaving us with the estimator

$$\hat{C}_l = \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}|^2 . \tag{2.24}$$

This, however, comes with the caveat of *Cosmic variance*. At low multipoles, we have fewer m-samples, thus the sample variance is high at these multipoles, an intrinsic effect of the estimator we can not minimize. The higher the multipole, the smaller the effect of cosmic variance.

2.6 Theoretical model

Several anomalies statistical anomalies have been reported in the CMB, which will be discussed in more detail later in this thesis in chapter 4, but briefly they are: the small-scale and large-scale hemispherical power asymmetry, the Cold Spot, the low power on large angular scales, the quadrupole and octopole alignment and the parity asymmetry on large angular scales. As a phenomenological model to reproduce these anomalies, Hansen et al. [1] proposed a scale-dependent $g_{\rm NL}$-type model,

$$\begin{aligned} T(\theta,\phi) &= T_G(\theta,\phi) + \gamma \left[T_G(\theta,\phi) T_F^2(\theta,\phi) \right]^{\text{Filtered}} \\ &= T_G(\theta,\phi) + \gamma \sum_{lm} g_l Y_{lm}(\theta,\phi) \int d\Omega' Y^*_{lm}(\theta',\phi') \cdot T_G(\theta',\phi') T_F^2(\theta',\phi'), \end{aligned} \tag{2.25}$$

where $T_G(\theta,\phi)$ is an isotropic Gaussian CMB temperature realization, and

$$T_F(\theta',\phi') = \sum_{lm} w_l Y_{lm}(\theta,\phi) \int d\Omega' Y^*_{lm}(\theta',\phi') T_G(\theta',\phi'). \tag{2.26}$$

The filters w_l and g_l, and the amplitude γ, act as parameters that can be tuned to reproduce the CMB anomalies. In particular, it's easy to see that the parameter γ serves as a measure of strength for the non-Gaussianity, and we use 0.43.

The w_ℓ filter can be seen in figure 2.2. The oscillations in the low multipoles reproduce the parity asymmetry for the largest scales. The increments up to $\ell = 21$

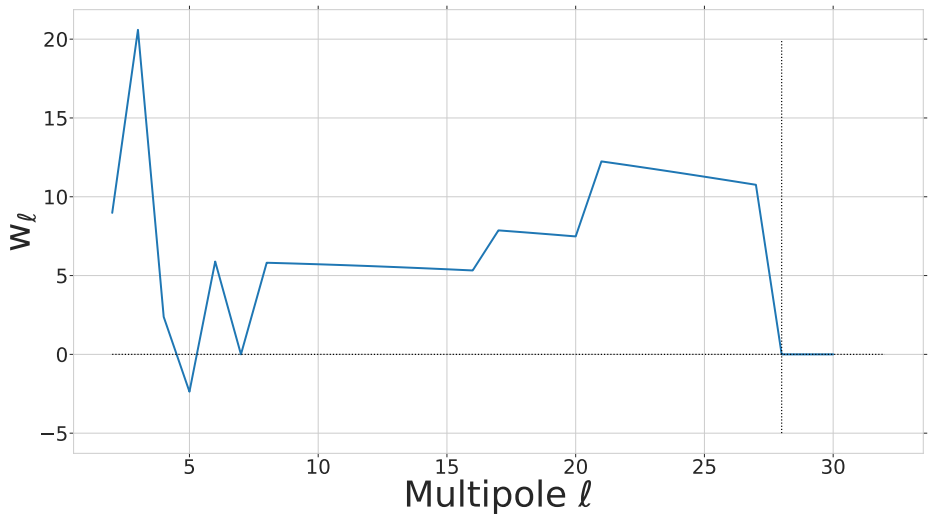


Figure 2.2: The w_ℓ filter, in the multipole range $l = [2, 30]$. The zero-line, as well as the cut-off at $l = 27$, is plotted for reference.

give rise to the large trough observed at this multipole in the observed power spectrum. There is a cut-off at $\ell = 27$, as the observed power spectrum is no longer abnormally low compared to the best-fit model. The g_ℓ filter is shown in figure 2.3. For the lowest multipoles, the filter is negative in order to suppress the power on large scales, and non-zero for multipoles up to $\ell \sim 1500$ to reproduce the small scale hemispherical asymmetry. The first value at $\ell = 2$, -8.8, is cropped for readability but will establish a small quadrupole in addition to creating correlations with the octopole, reproducing the quadru-octopole alignment.

The process of reproducing the anomalies in the CMB can be summed up in the following steps

1. Apply the first filter, w_l, to a Gaussian map.
2. Square the resulting filtered map.
3. Modulate the original Gaussian map by the square filtered map to obtain a non-Gaussian term.
4. Apply the second filter, g_l, to the non-Gaussian term.
5. Add the non-Gaussian term to the original Gaussian map to obtain a non-Gaussian map.

The process applied to a map from the Planck 2015 FFP8.1 MC simulations, illustrating the steps, can be seen in figure 2.4. To further illustrate the effects, an exaggerated

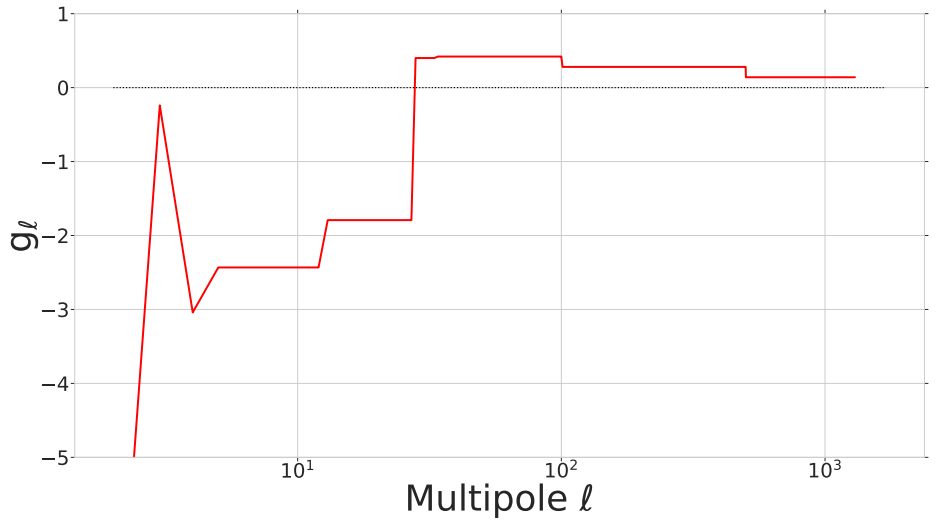
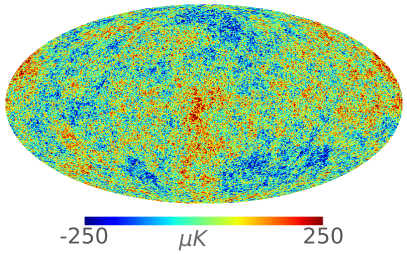
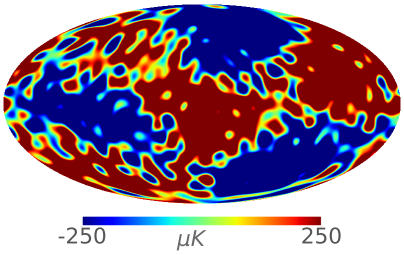


Figure 2.3: The g_ℓ filter. The filter is decreasing slowly towards zero at the higher multipoles. The zero-line is plotted for reference.

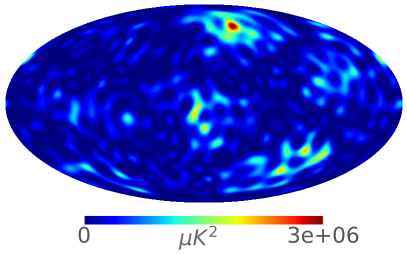
comparison is shown in figure 2.5. Especially visible in both figure 2.4 and 2.5 is the creation of strong hot and cold spots in the distribution. You can also see a strong indication of hemispherical asymmetry, as the central hemisphere contains much stronger fluctuations than the hemisphere at the edges of the map.



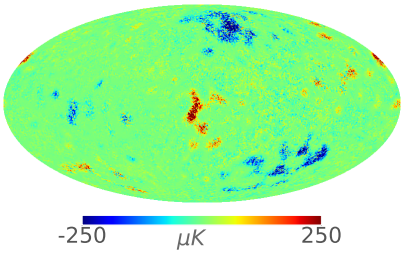
(a) Original Planck MC map



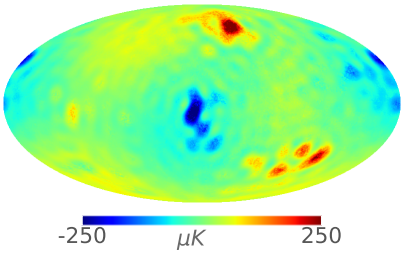
(b) w_l-filtered map



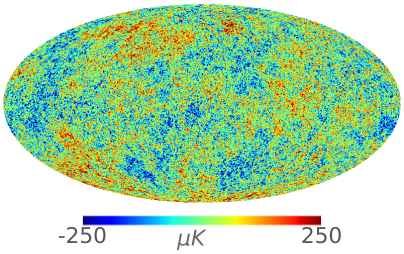
(c) Square filtered map



(d) NG-term

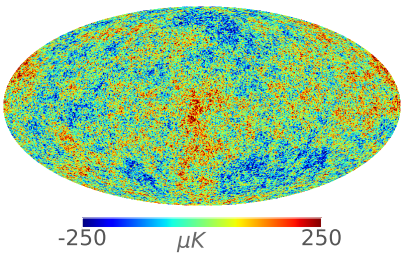


(e) g_l-filtered normalized NG-term

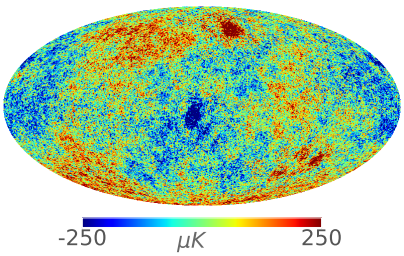


(f) Original + normalized NG-term

Figure 2.4: The entire proposed process of applying the filters to reproduce the 6 anomalies in a CMB map. Units are given in micro Kelvin.



(a) Original Planck MC map



(b) NG map with 2γ

Figure 2.5: Exaggerated Non-Gaussian map by doubling the strength parameter γ.

Chapter 3

Data analysis and statistics

3.1 Pipeline

Before one can draw conclusions from the observational raw data from the CMB, the data must first be processed through a relatively extensive data pipeline. The following section will give a rough outline of the steps taken to go from raw data to fit a model of the universe to the data.

3.1.1 Map making

The actual measurements from a satellite are raw data. This data is first pre-processed into Time Ordered Data (TOD). The TOD can be described as

$$d_t = s_t + n_t = A_{tp} s_p + n_t, \tag{3.1}$$

where d denotes the total data stream, that is the temperature measured by the detector at time t. The quantity s is the actual sky signal and n is a noise vector. The quantity A is the pointing matrix, which contains weights for each pixel, dependent on where the telescope has "been pointed" as a function of time. More weight is given to areas of the sky which has been observed the most. The pointing matrix thus converts between a time-dependent data stream and a spatially dependent data stream. The subscript p denotes a pixel number.

To compress the TOD into a much smaller data set, it can be converted into sky-maps, by the map-making equations

$$N_{pp'}^{-1} = A_{tp}^T N_{tt'}^{-1} A_{t'p'} \tag{3.2}$$

$$z_p = A_{tp}^T N_{tt'}^{-1} d_{t'} \tag{3.3}$$

$$d_p = N_{pp'} z_{p'}, \tag{3.4}$$

where N is a matrix describing the noise spectrum, an exponent of -1 denotes matrix inversion and t, t' p, p' are two different points in time and space respectively [36]. More often than not, however, a different approach to map-making is taken, as solving these equations is extremely computationally intensive.

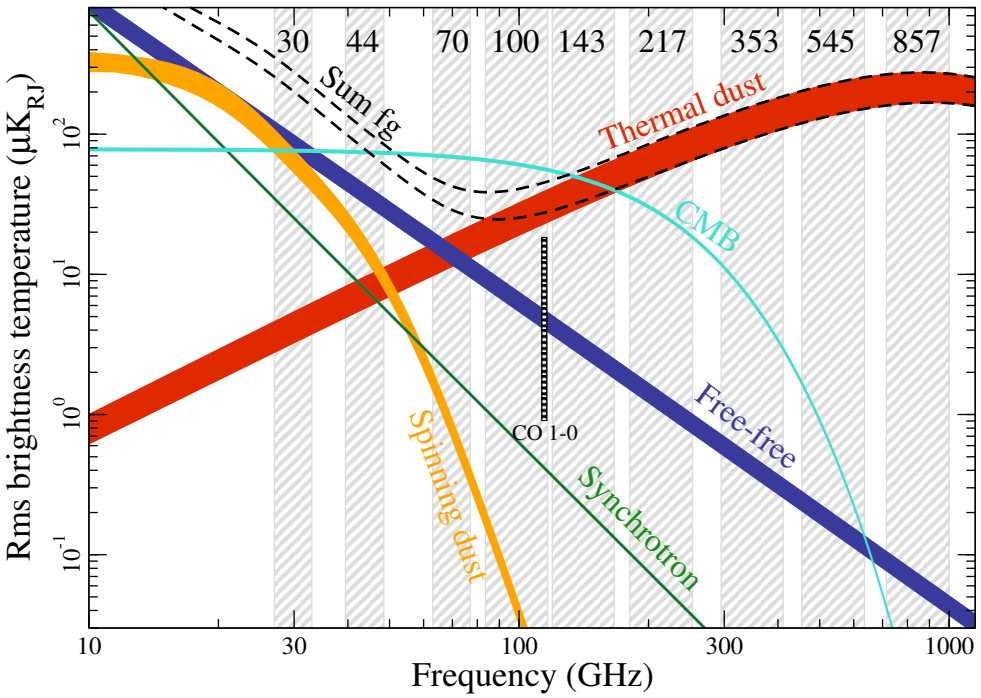


Figure 3.1: The foreground components in frequency. The grey shaded areas represent the detectors of Planck. Image credit: Planck Collaboration [37].

3.1.2 Galactic foregrounds

The light we receive on earth is (luckily) not all CMB radiation. To get an accurate measurement of the CMB radiation, one must first eliminate all sources of signal not originating from the surface of last scattering, so-called foregrounds. One of the principal sources of contamination in the CMB signal is the diffuse galactic emission. To effectively separate the CMB signal from the galactic emission, one must first have an understanding of which components make up the galactic foreground emission, and how they contribute. An image of how the different components contribute to the overall signal as a function of frequency is shown in figure 3.1.

Unless otherwise specified, the information in this section on galactic foregrounds and the following section on extra-galactic foregrounds is taken from the Planck 2013, 2015 and 2018 papers, "Diffuse Component Separation" [22, 37, 25].

Synchrotron

Synchrotron radiation is a dominant source of contamination at low frequencies. In the context of CMB observations, synchrotron radiation typically occurs when cosmic-ray electrons are spiraling in the Galactic magnetic field, releasing radiation due to the acceleration imposed on them by the field. Synchrotron radiation is empirically

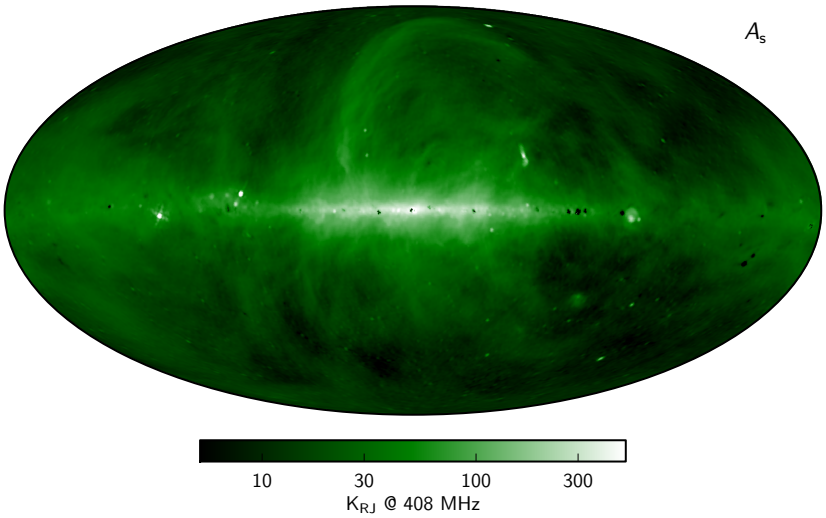


Figure 3.2: The planck 2015 synchrotron emission foreground map. Estimated from the maximum posterior in a Bayesian component separation approach, plotted with a non-linear high dynamic range colour scale. Image credit: Planck Collaboration [37].

modeled as a power law in frequency

$$T_{\mathrm{B,synch}}(\nu) \propto \nu^{\beta_{\mathrm{synch}}}, \tag{3.5}$$

where T_B is the brightness temperature, ν is the frequency, and the spectral index β_{synch} is a parameter to be tuned. For the Planck frequency range, $\nu \in [30, 857]$ GHz, β_{synch} lies in the range $[-3.2, -2.8]$. The planck synchrotron foreground template can be seen in figure 3.2.

Free-free

Free-free radiation, or Brehmsstrahlung, occurs in the CMB signal as a result of electron-electron or electron-ion scattering, where a free electron is accelerated and converts kinetic energy to radiation. It can also be, in the simple case, modeled as a power law in frequency

$$T_{\mathrm{B,ff}}(\nu) \propto \nu^{\beta_{\mathrm{ff}}}. \tag{3.6}$$

For Planck frequencies, the spectral index for free-free radiation has a slightly less steep slope than that of synchrotron radiation, lying in the range $\beta_{\mathrm{ff}} \in [-2.2, 2.1]$. The Planck free-free foreground template can be seen in figure 3.3. The Free-free radiation that contaminates the CMB typically occurs in hot clouds of gas in the galactic plane.

Anomalous Microwave emission (AME)

In the first analyses attempting component separation on the CMB signal, detection of residual emission was made in the frequency range below 40 GHz, not consistent with

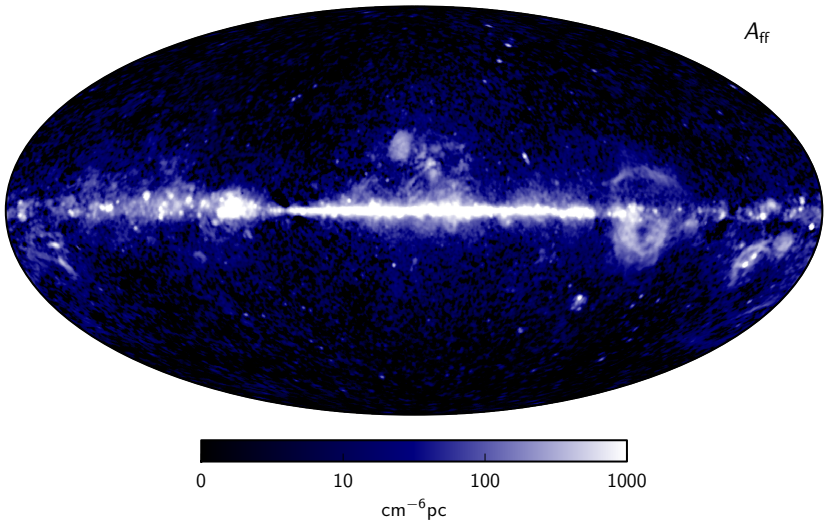


Figure 3.3: The planck 2015 free-free emission foreground map. Estimated from the maximum posterior in a Bayesian component separation approach, plotted with a non-linear high dynamic range colour scale. Image credit: Planck Collaboration [37].

the current foreground models for synchrotron and free-free radiation. This contaminant radiation was dubbed "Anomalous Microwave emission (AME)" in a paper by Leitch et al [38]. AME was later attributed to spinning dust grains, especially apparent below 70 GHz, in a paper by Draine, B. T. and Lazarian, A. [39]. This radiation has a sharp cutoff in its contribution, as there is a limit to how fast these dust grains can spin. The planck AME foreground template can be seen in figure 3.4.

Thermal dust emission

At frequencies above 100 GHz, the dominating contribution to the signal is emission from thermal dust grains. Thermal dust emission has been modeled empirically as a blackbody spectrum along with a power law component giving a two-parameter model in dust temperature T_d and spectral index for dust β_d,

$$T_{B,dust}\left(\nu, T_d\right) \propto \nu^{\beta_d+1} \frac{e^{\frac{h\nu_0}{k_B T_d}} - 1}{e^{\frac{h\nu}{k_B T_d}} - 1}, \tag{3.7}$$

where $T_{B,dust}$ is the brightness temperature for dust, T_D is the dust temperature and ν is the frequency [37]. The Planck thermal dust foreground template can be seen in figure 3.5.

CO rotational line emission

Discrete line emission from the $J = 1 \rightarrow 0$, $J = 2 \rightarrow 1$ and $J = 3 \rightarrow 2$ rotational transitions in the CO-molecule, at 115, 230 and 345 GHz respectively. The transition

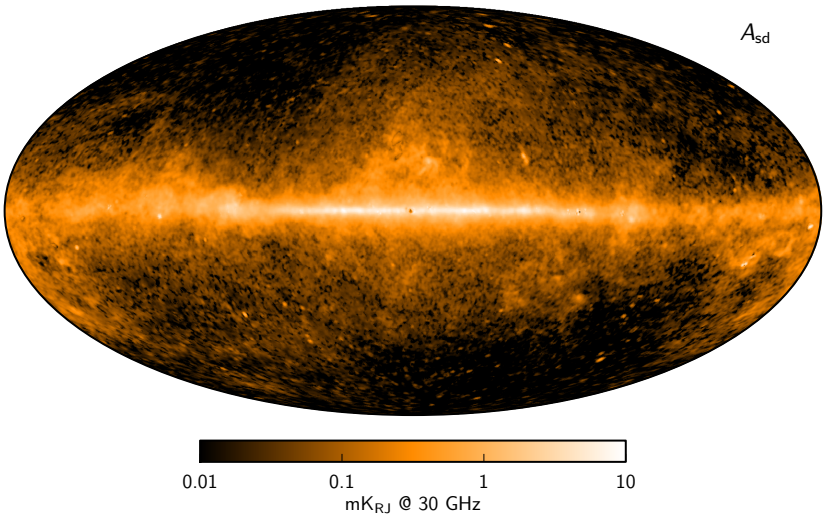


Figure 3.4: The planck 2015 spinning dust emission foreground map. Estimated from the maximum posterior in a Bayesian component separation approach, plotted with a non-linear high dynamic range colour scale. Image credit: Planck Collaboration [37].

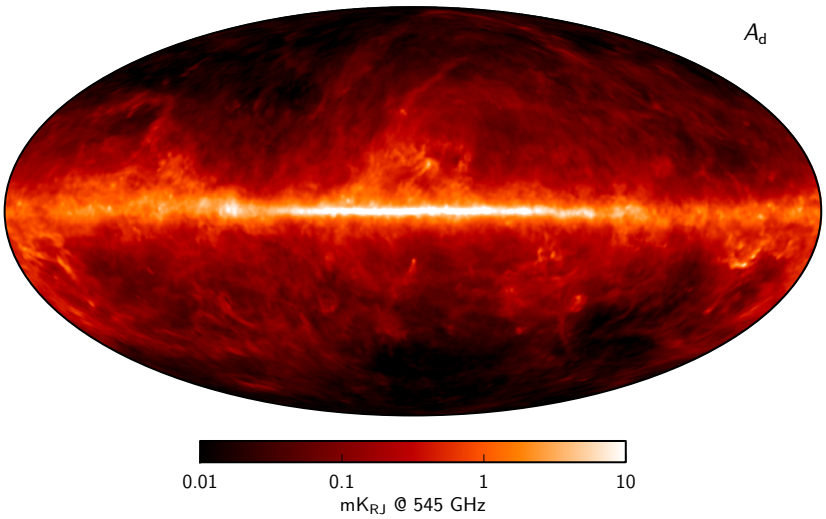


Figure 3.5: The planck 2015 thermal dust emission foreground map. Estimated from the maximum posterior in a Bayesian component separation approach, plotted with a non-linear high dynamic range colour scale. Image credit: Planck Collaboration [37].

is well understood, so the contribution from CO-line can be modeled accurately.

3.1.3 Extra-galactic foregrounds

There is also the contribution from sources outside our own galaxy to consider. A few of them are presented here.

Point sources

There is also a contribution from other compact, extragalactic radio sources. Many of these have been detected and masked individually in sky template maps. However, a residual contribution remains of unresolved sources. These are usually quasars, radio galaxies or other objects in the sky which emanate in the radio frequency domain [40]. The contribution from these extra-galactic foregrounds is generally degenerate with the emission of galactic origin and can be absorbed into the framework of galactic component separation.

(Thermal) Sunyaev-Zeldovich effect

The thermal Sunyaev-Zeldovich (tSZ) effect is produced by the inverse Compton scattering of CMB photons due to hot electrons along the line of sight, especially apparent in galaxy clusters. Named after its discoverers Sunyaev and Zeldovich in 1972 [41], it has been an important tool to study structure formation and galaxy clusters. The template used in Planck, as a function of multipole l, is shown in figure 3.6.

3.1.4 Foreground removal

Once the foregrounds have been modeled adequately, they are removed from the maps via a component separation technique (the modeling and creation of foreground-cleaned maps can be one algorithm). Popular methods include Commander, NILC, SEVEM, and SMICA, which are the ones used in the latest Planck release to extract the cleaned CMB signal from the full sky-maps [25].

Commander [45, 46, 37] is a Bayesian approach, using Gibbs sampling to fit a parametric model to the data.

NILC (Needlet Internal Linear Combination) [47, 48] is a method employing a needlet basis, which will be described later in section 3.2.2, calculating a linear combination of maps to minimize contamination from foregrounds and instrumental noise.

SEVEM [49, 50] is a template cleaning approach, typically creating difference maps from neighboring frequency channels such that the CMB vanishes, and you are left with a foreground template.

SMICA (Spectral Matching Independent Component Analysis) [51] is a method using spherical harmonic coefficients, combined with a weight vector describing the CMB emissivity, to freely match components through sky-maps of different frequencies.

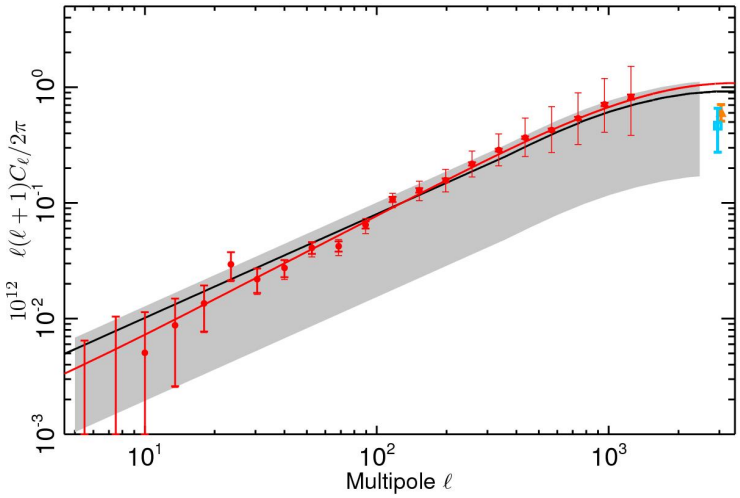


Figure 3.6: The power spectrum template of the SZ-effect used in Planck's cosmological parameter analysis. The red points are the NILC-MILCA F/L cross-power spectrum after foreground subtraction, with F/L referring to First and Last halves of the data. The black line shows the tSZ power spectrum used in the Planck cosmological analysis [42, 43], while the cyan and orange dot are the power spectrum estimates from the Atacama Cosmology Telescope and South Pole Telescope respectively. The grey region shows the $\pm 2\sigma$ confidence region. Image credit: [44].

3.1.5 Pixel window and beam

When we observe the CMB, we can not measure a continuous signal but a discrete one, limited by the resolution of our instruments. The effect of the finite size of the beam of a telescope, the fact that we can not restrict the pointing of our telescope to an infinitely small point, is encapsulated in the beam function. Furthermore, the map-making process introduces pixelation of the signal, when it is transformed from time steps to pixels. This process is described by the pixel window function. The beam function and pixel window function represents a convolution between the signal and a smoothing function in real space but are by the convolution theorem simple linear operators in harmonic space.

The pixel window function accounts for the pixelation of the continuous sky and can be represented in harmonic space as a product

$$a_{lm,\mathrm{pix}} = a_{lm,\mathrm{cont}} p_l, \tag{3.8}$$

where p_l is the pixel window function. Similarly, the beam window function, b_l, accounts for the fact that a point of observation is smoothed by surrounding contamination and can represented as the product

$$a_{lm,\mathrm{smoothed}} = a_{lm,\mathrm{cont}} b_l \tag{3.9}$$

Sometimes, these two effects are gathered into a single window function, accounting for the discrete representation of a continuous phenomenon.

In addition, the fact that we can not represent a sphere perfectly on a flat surface will introduce additional error, making the choice of pixelation scheme important for data analysis. As is common in CMB experiments, we will use the Hierarchical Equal Area isoLatitude Pixelation (HEALPix) [2, 52] to represent the sphere, shown in figure 3.7.

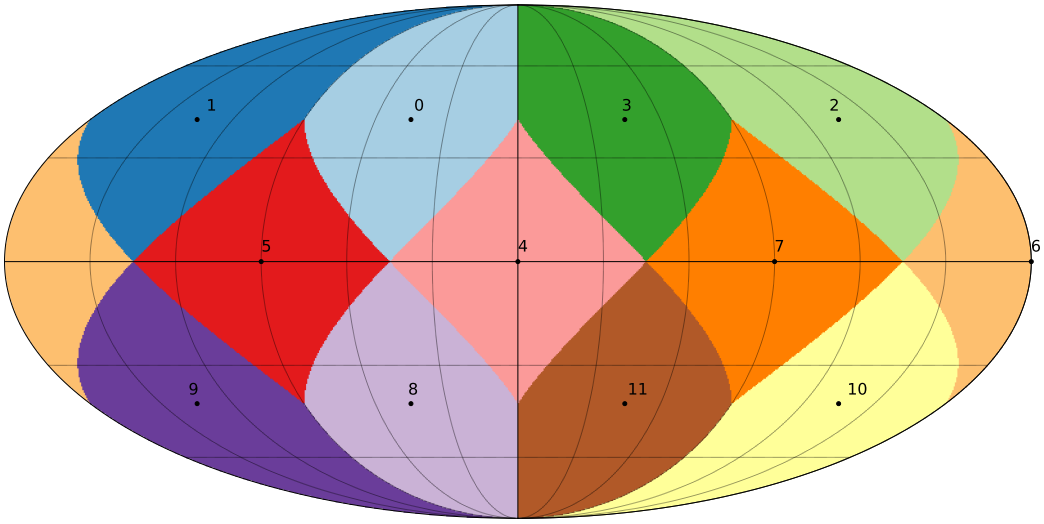


Figure 3.7: HEALPix pixellation scheme, RING ordering with NSIDE=1 showing the 12 base pixels.

In the HEALPix scheme, each pixel is of equal area, and the resolution is governed by the parameter N_{side}, governing how many "equal cuts" are applied to the 12 base pixels, such that each pixel will be divided into N_{side}^2 new ones, I.E. $N_{\text{side}} = 2$ will yield a pixellated sphere with 48 pixels as per the formula

$$N_{\text{pix}} = 12 \cdot N_{\text{side}}^2. \tag{3.10}$$

A visualization can be seen in figure 3.8.

3.1.6 Noise

We turn our attention to the noise term in equation 3.1. One possibility is to model the noise simply as "white-noise", meaning it is independent of one measurement to another. This, however, generally seems to be insufficient to properly model the noise properties of most modern CMB experiments. A more fitting approach is to split the noise into two components,

$$n_t = n_t^{\text{corr}} + n_t^{\text{wn}} \tag{3.11}$$

[2] `https://healpix.sourceforge.io/`

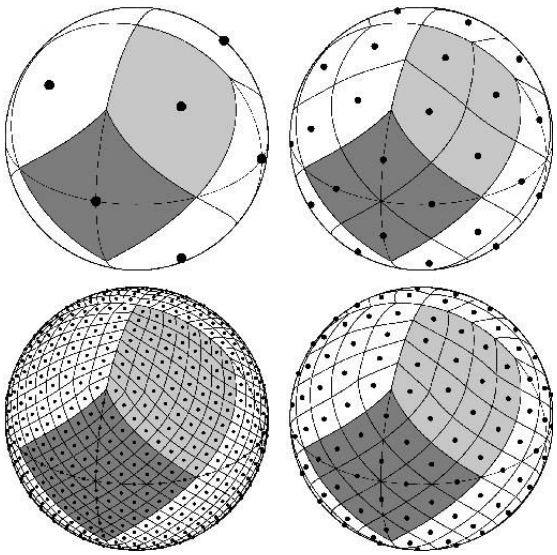


Figure 3.8: HEALPix pixellation scheme, with $N_{\text{side}} = 1, 2, 3, 4$. Original image from The Healpix Primer [52], direct access from HEALPix[2].

where n_t^{wn} is a white-noise component originating from thermal noise intrinsic to the detectors and amplifiers, and n_t^{corr} is a correlated noise term. This correlated term is often modeled as a "$1/f$-term", sometimes also known as "pink noise". It is called a $1/f$-term because of the inverse proportionality with the frequency, i.e. the noise power spectral density, $P_{\text{pink}}(f)$, goes as

$$P(f) \propto \frac{1}{f^{\gamma}}, \tag{3.12}$$

where γ is some constant. For the LFI instrument in the Planck mission, the noise was modeled as

$$P(f) = P_0^2 \left[1 + \left(\frac{f}{f_{\text{knee}}} \right)^{\beta} \right], \tag{3.13}$$

where P_0^2 is the white-noise spectrum estimated by averaging over frequency ranges where the level of pink noise is assumed to be low. β and f_{knee} are parameters governing the pink noise component. In the Planck 2018 results, the slope β was estimated to lie in the range $\beta \in [-1.30, -0.92]$, depending on the frequency range and radiometer [53]. The $1/f$-term is sourced by intrinsic instabilities in the detectors, amplifiers, and read-out electronics, environmental effects, and for ground-based experiments, atmospheric fluctuations. A recent analysis has shown that this approach works decently well for many frequency ranges. However, some deviations were found where the "$1/f$-model" was insufficient [54].

3.1.7 Estimating the power spectrum

What we ultimately want to do, is to connect the observations of the fluctuations in the CMB to a cosmological model, and estimate the cosmological parameters tied to this model. With isotropic and Gaussian fluctuations, the cosmological information in the full-sky map can be reduced to a few thousand numbers, in the form of the power spectrum [36]. Then, connecting the power spectrum to a cosmological model and estimating the cosmological parameters is a much simpler task. However, with real-world data, the spherical harmonics are no longer orthogonal because of incomplete sky coverage, and therefore we can not simply evaluate equation 2.24. So first, let us take a look at the necessary steps towards achieving the goal of parameter estimation.

The likelihood function

To actually use the power spectrum for parameter estimation, we need to know the complete likelihood function $P(d|C_l(\boldsymbol{\theta}))$, where d is the temperature data, C_l the angular power spectrum and $\boldsymbol{\theta}$ the vector containing all relevant cosmological parameters. For a full-sky map, it is not computationally feasible to compute the full likelihood function, therefore the estimation of the power spectrum and likelihood function is divided into high-l and low-l codes. I will a common method for power spectrum estimation here, the MASTER method.

High l: MASTER

In preparation for high-resolution data, it was clear that general maximum likelihood methods of extracting the power and noise spectra from CMB maps were too computationally prohibitive, with a scaling of $\sim N_{\mathrm{pix}}^3$. WMAP observed the sky with a maximum resolution of 3.15 million pixels [55], while Planck had a resolution of 1.16 billion pixels [26], creating a severe bottleneck if one were to solve with traditional full matrix inversions, and a faster method was needed. In 2001, Hivon et al. introduced the Monte Carlo Apodised Spherical Transform EstimatoR (MASTER) [56], as a method for fast and accurate estimation of the angular power spectrum, achieving a much faster computational scaling at $\sim \sqrt{N_{\mathrm{pix}}} N_{\mathrm{pix}}$. The following is an outline of the method presented in their paper.

When we measure the power spectrum, we measure the pseudo-spectrum $\tilde{C}_l$. The pseudo-power-spectrum is a collection of multiple components, and can be modeled as,

$$\langle \tilde{C}_l \rangle = \sum_{l'} M_{ll'} F_{l'} B_{l'}^2 \langle C_{l'} \rangle + \langle \tilde{N}_l \rangle, \tag{3.14}$$

where C_l is the true full-sky angular spectrum of which we want an estimator, $M_{ll'}$ is a mode-coupling matrix accounting for the sky-cut, $F_{l'}$ is a transfer function describing the filtering applied during the map-making process, $B_{l'}$ is the window function accounting for both the effect of the pixel size and beam as shown in equations 3.8 and 3.9 and $\langle \tilde{N}_l \rangle$ is the average noise power spectrum. The algorithm for computing these terms is as follows:

1) Evaluate the coupling kernel $M_{ll'}$, using the equation

$$M_{l_1 l_2} = \frac{2l_2+1}{4\pi} \sum_{l_3} (2l_3+1)\, \mathcal{W}_{l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2, \tag{3.15}$$

where the brackets represent the Wigner-3j symbols. For more on the Wigner-3j symbols, see appendix A.1. $\mathcal{W}_l$ is the power spectrum of a position-dependent window function (a mask)

$$\mathcal{W}_l = \frac{1}{2l+1} \sum_m |w_{lm}|^2, \tag{3.16}$$

in case the full sky is not observable. For example, if f_{sky} represents the fraction of sky that is non-zero, we have

$$f_{\text{sky}} w_i = \frac{1}{4\pi} \int_{4\pi} d\mathbf{u} W^i(\mathbf{u}), \tag{3.17}$$

where $W(\mathbf{u})$ is a position-dependent weighting scheme to reduce the effects of edges or noisy pixels.

2) The transfer function F_l is estimated through producing a series of noise-free Monte-Carlo simulations $N_{\text{MC}}^{(s)}$, projecting them onto the sky and analyzing them with a pseudo power spectrum estimator (HEALPix anafast). In our particular approach, we omit the transfer function F_l.

3) Step 2 is repeated only now with pure noise MC simulations to estimate $\langle \tilde{N} \rangle$, the power spectrum of the noise.

4) Make a (masked) map and calculate its pseudo power spectrum $\tilde{C}_l$.

5) Estimate a full-sky *binned* power spectrum, with the estimator $\hat{C}_b$ given as the inverse of equation 3.14,

$$\hat{C}_b = K^{-1}_{bb'} P_{b'l} \left(\tilde{C}_l - \langle \tilde{N}_l \rangle_{\text{MC}} \right), \tag{3.18}$$

where $P_{b'l}$ is the binning operator

$$P_{b'l} = \begin{cases} \frac{1}{2\pi} \frac{l(l+1)}{l^{b+1}_{\text{low}} - l^b_{\text{low}}} & 2 \leq l^b_{\text{low}} \leq l \leq l^{b+1}_{\text{low}} \\ 0 & \text{otherwise} \end{cases} \tag{3.19}$$

for a set of bins indexed by b with boundaries $l^b_{\text{low}} < l^b_{\text{high}} < l^{b+1}_{\text{low}}$. The reciprocal operator is then given as

$$Q_{lb} = \begin{cases} \frac{2\pi}{l(l+1)} & 2 \leq l^b_{\text{low}} \leq l \leq l^{b+1}_{\text{low}} \\ 0 & \text{otherwise} \end{cases}. \tag{3.20}$$

Then, $K_{bb'}$ is given as

$$K_{bb'} = P_{bl} M_{ll'} F_{l'} B^2_{l'} Q_{l'b'}. \tag{3.21}$$

6) Calculate the covariance matrix $\mathbf{C}_{bb'}$ with the equation

$$\mathbf{C}_{bb'} = \left\langle \left(\widehat{\mathcal{C}}_b - \langle \widehat{\mathcal{C}}_b \rangle_{\mathrm{MC}} \right) \left(\widehat{\mathcal{C}}_{b'} - \langle \widehat{\mathcal{C}}_{b'} \rangle_{\mathrm{MC}} \right) \right\rangle_{\mathrm{MC}} \tag{3.22}$$

from all the $N_{\mathrm{MC}}^{(s+n)}$ simulations, and the error bars on the binned spectrum are obtained through the diagonals of the covariance matrix

$$\Delta \widehat{\mathcal{C}}_b = \mathbf{C}_{bb}^{1/2}. \tag{3.23}$$

3.1.8 Parameter estimation

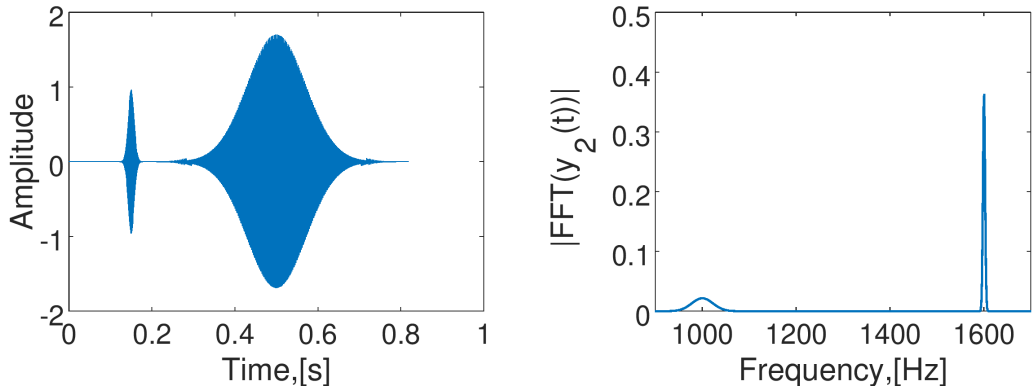
One of the most popular methods of estimating the cosmological parameters is through Markov Chain Monte Carlo (MCMC) simulations [36]. The MCMC simulations perform random walks through parameter space, calculating the likelihood at each proposed step, and adheres to an acceptance rule governing whether to take the next step or not. This process of drawing random samples, calculating the likelihood, and either moving forward a step or choosing a new point in parameter space is a class of algorithm known as the Metropolis-Hastings algorithm.

3.2 Wavelet- and needlet analysis

The use of wavelet- and needlet analysis has increased greatly in the field of cosmology over the last decade. Wavelet analysis enjoys the benefit of double-localization properties, making them excellent for analysis of the CMB due to foregrounds and masking preventing orthogonality of the coefficients in the harmonic domain. The double-localization property can be explained as follows:

Imagine you have a time-dependent signal, and a corresponding Fourier analysis, like shown in figure 3.9. Now the Fourier transform gives you excellent information about the frequency content in the signal, being very sharply peaked around frequencies 1000 and 1600 Hz. However, the Fourier transform gives you absolutely no information about the time component of the signal. If your data set adheres to the mathematical condition of "nice behavior", then of course the inverse Fourier transform is possible to retrieve the information about the time component of the signal. However, what you can not obtain is information in frequency and time simultaneously, i.e. which frequencies occur when (or where in the case of CMB analysis).

Consider now instead a wavelet analysis. This approach would keep some information about both time and frequency, with their respective resolutions parametrized. The localization properties of both time and frequency, however, come at the cost of one another, that is to say, the better the resolution in frequency space, the worse the resolution in time. To visualize this, two wavelet transforms (of the signal shown in figure 3.9), are shown as coherence plots with increasing frequency resolution in figures



(a) Signal (b) Fourier transform

Figure 3.9: A time dependent signal, consisting of two main frequencies, along with the corresponding Fourier transform of the signal.

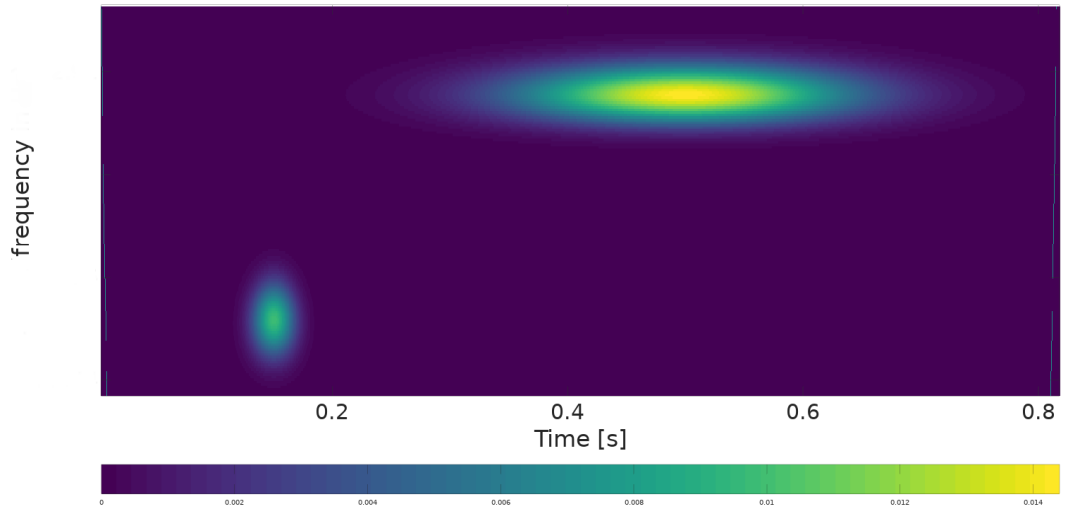


Figure 3.10: Wavelet coherence plot with time-frequency resolution. The wavelet shows a trade-off between resolution in time and frequency.

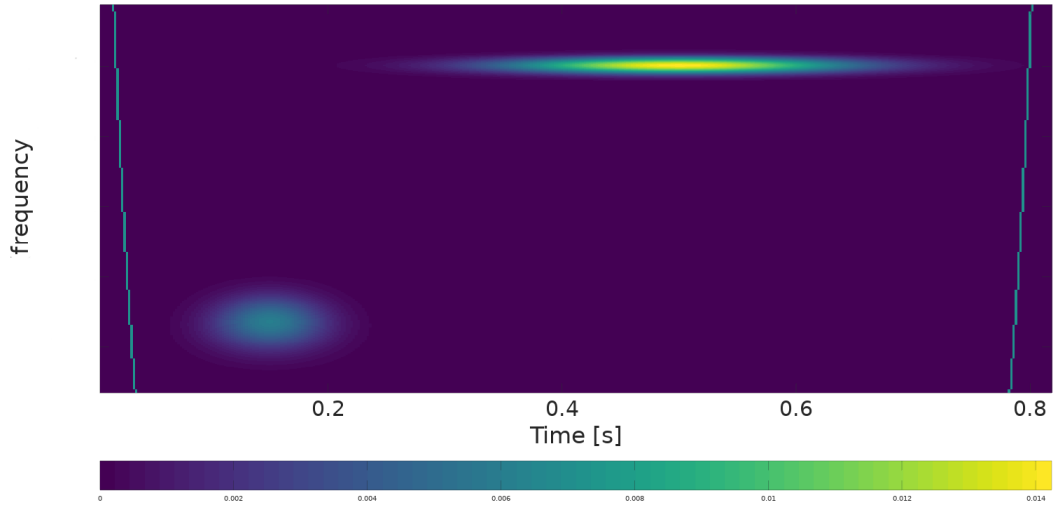


Figure 3.11: Wavelet coherence plot with time-frequency resolution. The localization in frequency space is increased here in comparison to figure 3.10, at the cost of smearing out the signal in time.

3.10 and 3.11. In the case of wavelet analysis on CMB, double-localization refers to resolution in real space (position) and resolution in harmonic space (scale, defined by multipole l). Two famous wavelets are shown in figure 3.12.

The idea behind wavelet analysis, in general, is that one constructs a "mother wavelet" of a certain shape, and a set of scale-dependent wavelets are built with corresponding weights from this mother wavelet. When the signal is then convoluted with the set of wavelets, we can quantify the response of the signal to the wavelet, giving a measure of power dependent on both position and scale. This process is called a wavelet transform, W, very similar to a Fourier or Gabor transform. A 2D continuous wavelet transform is given by,

$$w(R,\vec{b}) = \int_{\mathfrak{D}} d\vec{x} f(\vec{x}) \Psi(R,\vec{b};\vec{x}), \tag{3.24}$$

with

$$\Psi(R,\vec{b};\vec{x}) = \frac{1}{R}\psi\left(\frac{|\vec{x}-\vec{b}|}{R}\right), \tag{3.25}$$

where ψ is the mother wavelet, $f(\vec{x})$ is your function in position, $w(R,\vec{b})$ is the wavelet coefficient of scale R and coordinates $\vec{b}$ and $\mathfrak{D}$ is your domain on which the wavelets are defined [57]. We can see schematically that $\Psi(R,\vec{b};\vec{x})$ plays the role of the kernel in a general integral transform, just as a Fourier transform would have the kernel of $e^{-2\pi i t\nu}$.

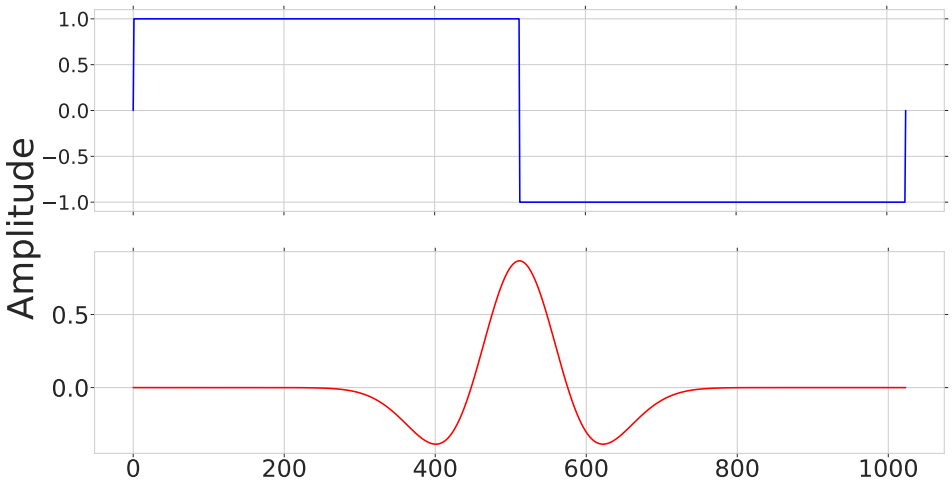


Figure 3.12: Two famous mother wavelet shapes: On top the Haar wavelet, and on the bottom the Mexican Hat wavelet. The dimension of the x-axis is arbitrary, but wavelet transforms are often used in time or space.

We can transform back via

$$f(\vec{x}) = \frac{1}{C_\psi} \int dR d\vec{b} \frac{1}{R^4} w(R, \vec{b}) \psi \left(\frac{|\vec{x} - \vec{b}|}{R} \right), \tag{3.26}$$

where $C_\psi = (2\pi)^2 \int dk k^{-1} \psi^2(k)$, $\psi(k)$ is the Fourier transform of $\psi(x)$ and k is the wave number.

Wavelets take advantage of the fact that low frequency components do not need high resolution in time, as they change slowly over short time intervals. A schematic overview of the idea is shown in figure 3.13, where a wavelet transform is applied over a large interval, increasing the resolution with smaller and smaller wavelets. If one is investigating, for example, a signal such as the one shown in figure 3.9, the upper panels represent excellent time resolution, but missing frequency information, while the lower panel represents excellent frequency resolution but missing time information.

3.2.1 Spherical Mexican Hat Wavelets

The traditional 2D Mexican Hat Wavelets has been widely used in a variety of different disciplines. The mother wavelet ψ is given as

$$\psi(x) = \frac{1}{(2\pi)^{1/2}} \left(2 - x^2\right) e^{-x^2/2}, x = |\vec{x}| \tag{3.27}$$

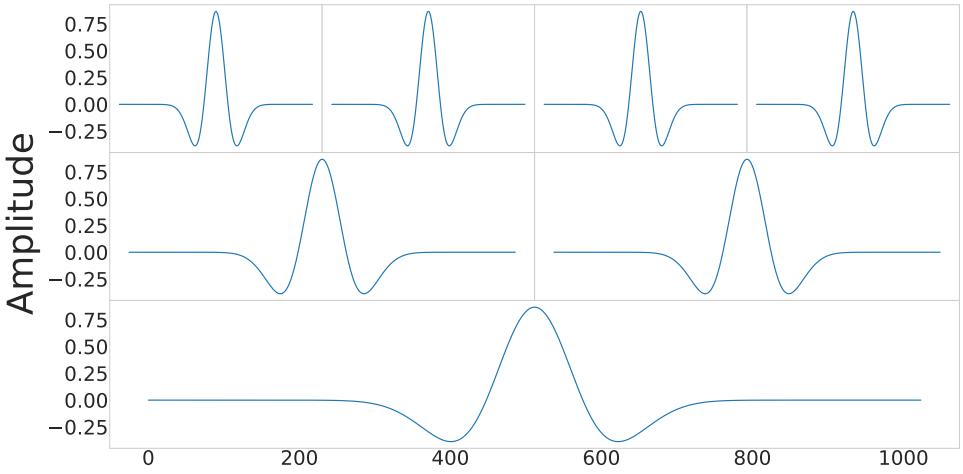


Figure 3.13: Applying the Mexican Hat Wavelet shown in figure 3.12, increasing time resolution with smaller and smaller wavelets.

A derivation of the continuous wavelet transform on the sphere by Antoine & Vandergheynst [58] gave way to adapting the Mexican Hat Wavelets for the sphere, as Spherical Mexican Hat Wavelets (SMHW). The points x are projected on the sphere via polar coordinates, $\vec{x} \rightarrow (\theta, \phi)$, as

$$x_1 = 2tan\left(\frac{\theta}{2}\right)\cos(\phi) \quad , \quad x_2 = 2tan\left(\frac{\theta}{2}\right)\sin(\phi), \tag{3.28}$$

where (θ, ϕ) are polar coordinates on the sphere, and $\left(y \equiv 2tan\left(\frac{\theta}{2}\right), \phi\right)$ are polar coordinates on the tangent plane at the north pole [59]. The idea is shown in figure 3.14.

The wavelet coefficients for the projected function $f(\theta, \phi)$ is then given by

$$w(R) = \int_{S_2} d\Omega f(\theta, \phi)\Psi_{S_2}(\theta; R), \tag{3.29}$$

where the kernel on the sphere Ψ_{S_2} is given by

$$\Psi_{S_2}(\theta; R) = N\frac{4}{(1 + cos\theta)^2}\psi\left(\frac{x}{R}\right) = N\frac{4}{(1 + cos\theta)^2}\psi\left(\frac{2}{R}tan\left(\frac{x}{R}\right)\right). \tag{3.30}$$

The shape of the Spherical Mexican Hat wavelets can be seen in figure 3.15, and the SMHW transformation on a map with a single non-zero pixel of arbitrary value can be seen in figure 3.16. Comparing scale $R = 600'$ from figure 3.15 with figure 3.16, we

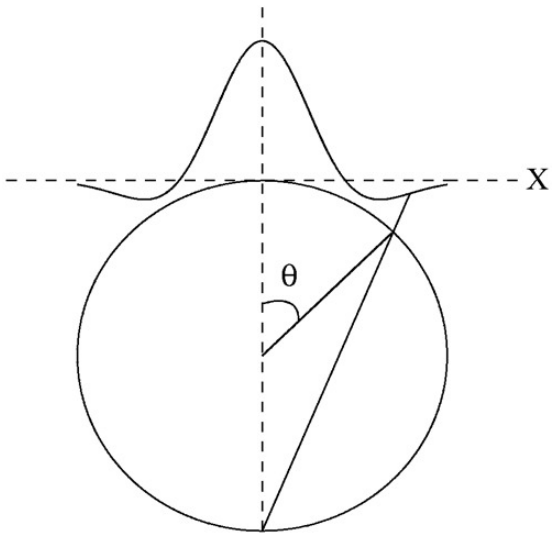


Figure 3.14: Projection of the 2D Mexican Hat Wavelet onto the sphere. Image credit: [59].

can see that the SMHW gives decent localization in real space, while still having the majority of its amplitude over 20 multipoles in harmonic space, providing information in both domains.

3.2.2 Needlets

Unless otherwise specified, this section relies heavily upon Scodeller et al. [60].

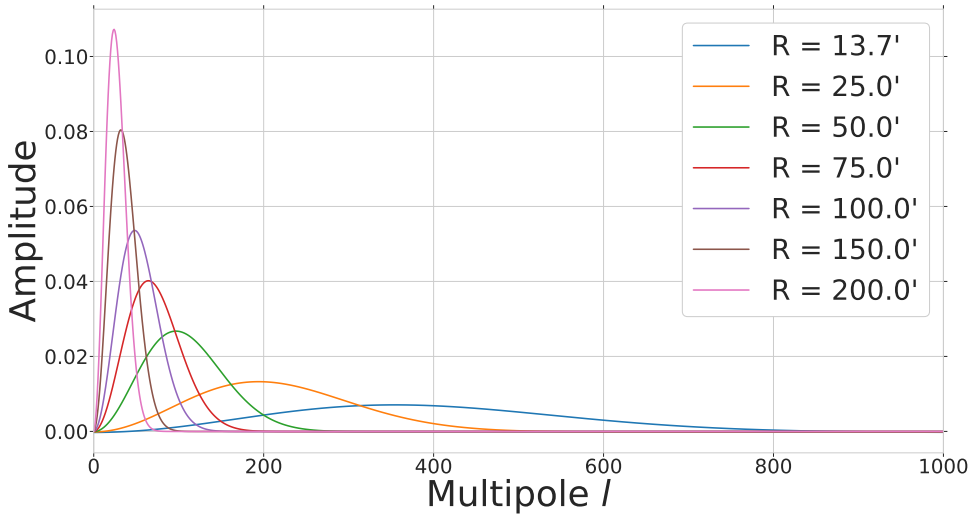
There is a class of wavelets, called needlets, first introduced in the literature by Narcowich et al. [61, 62], while Baldi et al. first discussed their statistical properties [63, 64].

The functional form of the mother needlet is given in real space as:

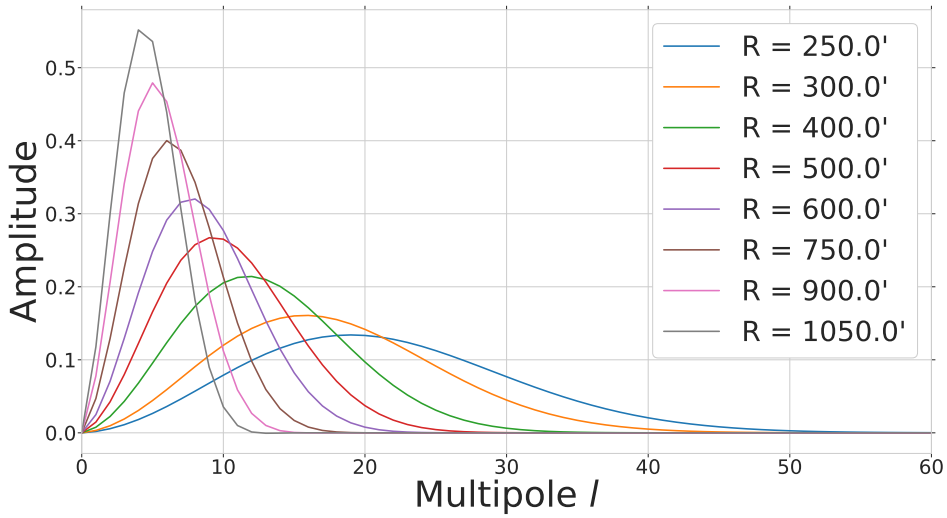
$$\psi_{jk}(x) \stackrel{\text{def}}{=} \sqrt{\lambda_{jk}} \sum_{l} b_l(B,j) \sum_{m=-l}^{l} Y_{lm}(\xi_{jk}) \overline{Y_{lm}}(x), \tag{3.31}$$

where $\{\lambda_{jk}\}$ is the set of *cubature weights* for scale(frequency) j and cubature point (pixel) ξ_{jk}, Y_{lm} are the spherical harmonics, x is the position and B is a parameter determining the properties of the needlet. The defining property determining the shape of the needlet is captured in the weight function b_l, and needlets are categorized in relation to how this weight function is defined. The weight function $b\left(\frac{l}{B^j}\right)$ is defined under the conditions that

1. $\sum_{j=0}^{\infty} b^2\left(\frac{l}{B^j}\right) = 1, \quad \forall l > B.$
2. $b\left(\frac{l}{B^j}\right)$ is non-zero in the domain $[B^{j-1}, B^{j+1}]$.
3. $b\left(\frac{l}{B^j}\right)$ is infinitely differentiable in $(0, \infty)$.



(a) Scales $R = \{13.7', 25.0', 50.0', 75.0', 100.0', 150.0', 200.0'\}$



(b) Scales $R = \{250.0', 300.0', 400.0', 500.0', 600.0', 750.0', 900.0', 1050.0'\}$

Figure 3.15: The shape of the SMHW, shown for different scales R, over the multipole range $l = [0, 1000]$, and a close up of the largest scales. The scales of the SMHW are given by R in units arcminutes.

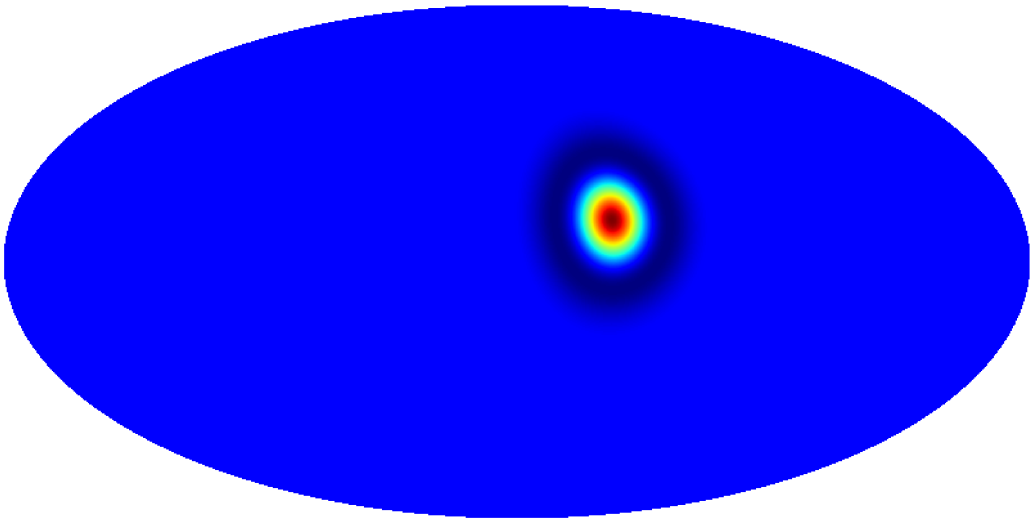


Figure 3.16: The SMHW transformation performed on a single non-zero pixel, with scale $R = 600'$ [arcmin].

Standard needlets

The following outlines how to construct the weight function for the standard needlets. For a function $\phi(\xi)$, we wish to obtain the weight functions $b\left(\frac{l}{B^j}\right)$. First, construct the function

$$f(t) = \begin{cases} \exp\left(-\frac{1}{1-t^2}\right) & , \quad -1 \leq t \leq 1 \\ 0 & , \quad otherwise \end{cases} \tag{3.32}$$

Then, construct

$$\phi(u) = \frac{\int_{-1}^{u} f(t)dt}{\int_{-1}^{1} f(t)dt}. \tag{3.33}$$

Now, construct

$$\varphi(t) = \begin{cases} 1 & , \quad 0 \leq t \leq \frac{1}{B} \\ \phi\left(1 - \frac{2B}{B-1}\left(t - \frac{1}{B}\right)\right) & , \quad \frac{1}{B} < t \leq 1 \\ 0 & , \quad t > 1 \end{cases} \tag{3.34}$$

And finally, we have

$$b^2(\xi) = \varphi\left(\frac{\xi}{B}\right) - \varphi(\xi), \tag{3.35}$$

and then $b(\xi)$ is the positive root of b^2. For more details on the properties of the weight function for the standard needlets, see Marinucci et al. [65].

The standard needlets enjoy several useful properties:

1. They are computationally simple and fit well with pixellation schemes such as HEALPix.

2. They do not require a tangent plane approximation as the SMHW does.

3. They depend only on a finite number of multipoles.

4. They are quasi-exponentially localized in real space. This means their tails decay faster than any polynomial.

5. They have have the property of uncorrelation; the coefficients are asymptotically uncorrelated as the frequency increases, for a fixed angular distance. Under Gaussianity, uncorrelation equals independence. Therefore, at very high frequencies, the needlets can be taken as nearly independent.

The shape and localization properties of the standard needlets can be seen in figure 3.17, while the needlet transformation applied to a map with a single non-zero pixel can be seen in figure 3.18. Looking at for example scale $j = 4$ from figure 3.17, we can see that the majority of the harmonic response spans only about 10 multipoles, while still giving information about real space localization, as seen for the same scale in figure 3.18.

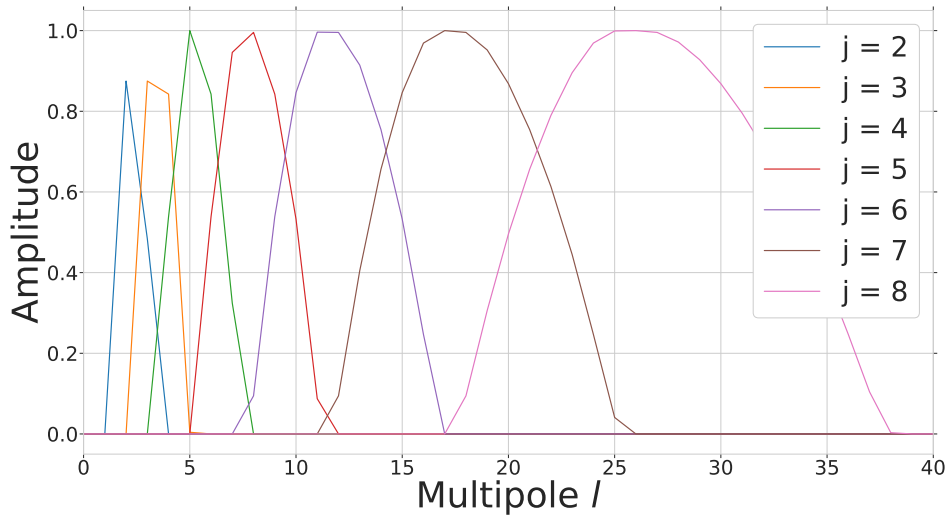
The effect of the parameter B on the shape of the needlet in harmonic space can be seen in figure 3.19. This parameter governs the localization properties in the harmonic domain versus the spatial domain. Increasing B will give you worse localization in harmonic space, smearing out the needlet over a larger range of multipoles ℓ, but will improve the localization in real space. The real space localization effect is shown in figure 3.20, where we have used the same scale $j = 4$ as in figure 3.18, only now width different values of B. We can see that the real space localization increases with higher values of B. There is no significant difference between $B = 1.5$ and $B = 2.0$ for this scale.

3.3 Polyspectra and statistics

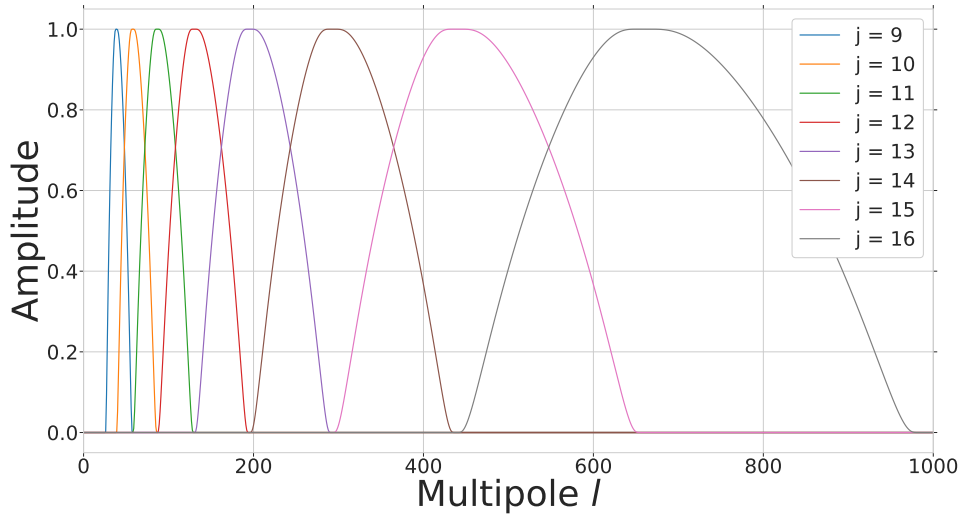
Polyspectra are the equivalent of moments of a distribution, within the framework of random spherical fields. They are very useful quantities when trying to determine the statistical properties of the distribution. In this section, I will outline some basic statistics relevant to the scope of this thesis.

3.3.1 One-dimensional statistics

For a single random variable, the moments of the distribution can provide us with valuable information.



(a) Scales $j = 2$ to $j = 8$



(b) Scales $j = 9$ to $j = 16$

Figure 3.17: The shape of the standard needlet, shown for different scales j, over the multipole range $l = [0, 1000]$, for $B = 1.5$.

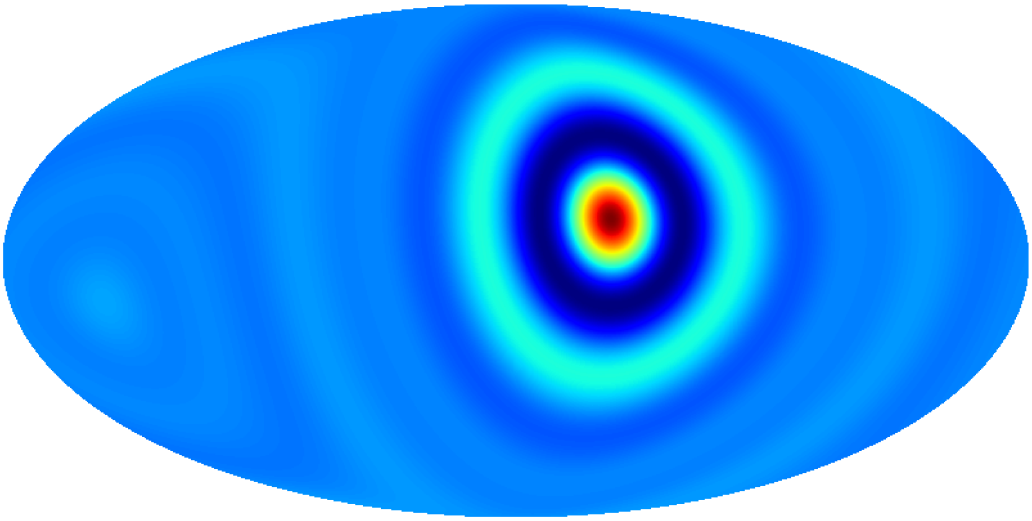


Figure 3.18: The standard needlet transformation performed on a single non-zero pixel, with $B = 1.5$ and scale $j = 4$.

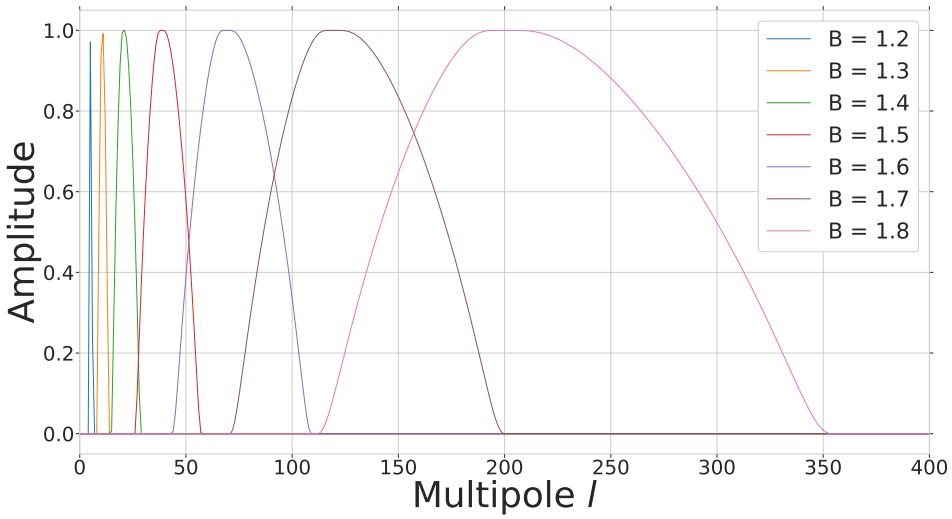
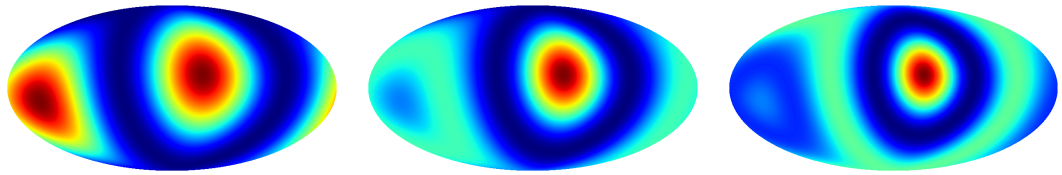
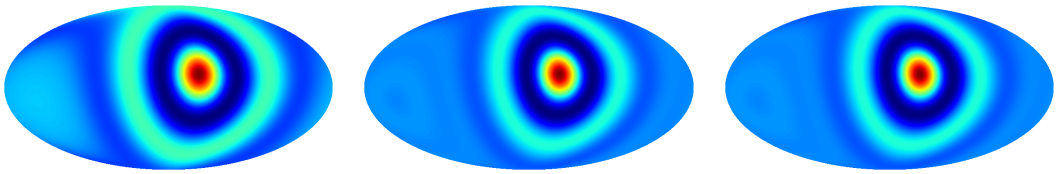


Figure 3.19: Different values of the localization parameter B, here for needlet scale $j = 9$.



(a) B = 1.15 (b) B = 1.25 (c) B = 1.35



(d) B = 1.45 (e) B = 1.50 (f) B = 2.0

Figure 3.20: Increasing the localization parameter B in real space, for needlet scale $j = 4$.

Mean and higher order central moments

For a random variable T, the mean is the expected value of T,

$$\mu = \langle T \rangle, \tag{3.36}$$

while the sample mean for a discrete data set, also known as the arithmetic mean, is given by

$$\langle T \rangle = \frac{1}{N} \sum_{i=1}^{N} T_i. \tag{3.37}$$

The higher order sample central moments are given by

$$\mu_n = \frac{1}{N-1} \sum_{i=1}^{N} \left(T_i - \langle T \rangle \right)^n . \tag{3.38}$$

For $n = 2$ we obtain the variance, the expected value of the squared deviation from the mean, while the standard deviation σ is given as the square root of the variance:

$$\sigma = \sqrt{\mathrm{Var}(T)} = \sqrt{\mu_2}. \tag{3.39}$$

Skewness and kurtosis

The skewness is defined as the normalized third order central moment,

$$\mathrm{Skew}[T] = S = \frac{\mu_3}{\sigma^3}. \tag{3.40}$$

The skewness is a measure of the asymmetry about the mean, for a given distribution. Examples of both positive and negative skewness is shown in figure 3.21, along with

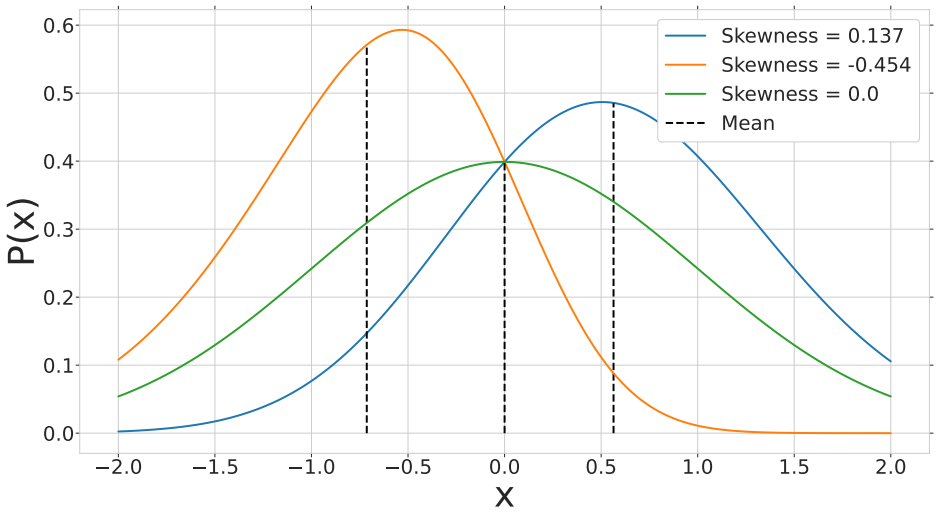


Figure 3.21: Three normal distributions with different skewness. A positive skewness, shown in red, will lead to a shift of the distribution to the left of the mean, while a negative skewness will lead to a shift of the distribution to the right of the mean.

a normal distribution with zero skewness for reference. We can see that a negative skewness will shift the distribution towards the right, such that the maximum of the distribution is to the right of the mean, while the opposite is true for a positive skewness.

The kurtosis K, is usually given as

$$\text{Kurt}\,(T) = K = \frac{\mu_4}{\sigma^4} - 3. \tag{3.41}$$

The subtraction of 3 ensures that the kurtosis of a Gaussian random variable is zero, as the fourth central moment of a Gaussian random variable always is $\mu_4 = 3\sigma^4$. The kurtosis measures the strength of the "tails" of the distribution. Examples of positive and negative kurtosis, again with the normal distribution of zero kurtosis for reference, are shown in figure 3.22. As per the definition, the "tailedness" of a normal distribution is the reference value of zero kurtosis. Figure 3.22 shows that for example a uniform distribution would result in a heavy negative kurtosis, as it has non-existant tails, while a Laplace distribution would yield a strongly positive kurtosis as the tails are much stronger than those of the normal distribution.

Transformed moments

The 1D moments for a harmonic or wavelet (needlet) decomposition are completely analogous, $T_i \to a_{lm}$ or $T_i \to \beta$, where β is the coefficients of the wavelet or needlet

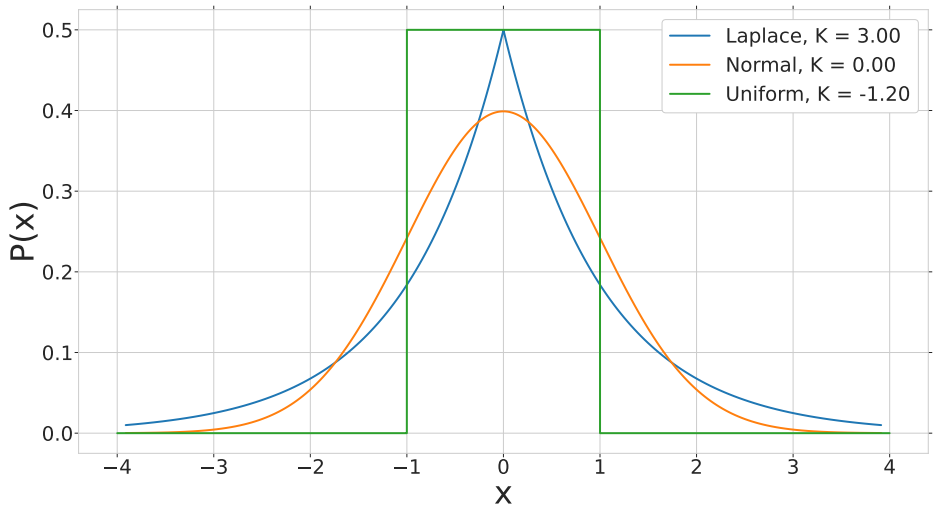


Figure 3.22: The laplace, normal and uniform (probability) distributions, along with their respective kurtosis values.

basis. The sample variance, skewness and kurtosis can be obtained by

$$S = \frac{1}{\sigma^3}\frac{1}{N}\sum_{i=1}^{N}(\beta - \langle\beta\rangle)^3, \tag{3.42}$$

$$K = \frac{1}{\sigma^4}\frac{1}{N}\sum_{i=1}^{N}(\beta - \langle\beta\rangle)^4 - 3, \tag{3.43}$$

$$\sigma = \left[\frac{1}{N}\sum_{i=1}^{N}\beta^2\right]^{1/2}. \tag{3.44}$$

Implementation-wise, we would have β dependent on a scale and position, $\beta = \beta(\theta_i, \phi_i, R)$ for wavelets, and $\beta = \beta(\theta_i, \phi_i, j)$ for needlets. The summation variable i would formally refer to cubature points, but in practice taken as pixel numbers in the HEALPix scheme, with the sum going up to $N = N_{\text{pix}}$.

3.3.2 The two-point correlation function or the power spectrum

In general, an n-point correlation function measures the correlation between n random variables, in our case the temperature fluctuations in the CMB as a function of distance x. The two-point correlation function is given as the ensemble average

$$C(|x_1 - x_2|) = \langle T(x_1)T(x_2)\rangle = \sum_{l_1,l_2}\sum_{m_1,m_2}\langle a_{l_1m_1}a^*_{l_2m_2}\rangle Y_{l_1m_1}(x)Y^*_{l_2m_2}(x). \tag{3.45}$$

The harmonic transform of the two-point correlation function is the angular power spectrum already discussed in section 2.5.2.

3.3.3 The three-point correlation function or the Bispectrum

If the primordial fluctuations are Gaussian distributed, the anisotropy is completely described by the two-point correlation function, or the "Power spectrum", C_l, in harmonic space. The reason for this is that a normal distribution is fully described by the mean and variance, and the power spectrum represents the variance of the a_{lm} coefficients with zero mean. However, if the fluctuations are non-Gaussian, there is additional information contained in the higher order moments. The first of these higher order moments is the three-point correlation function, or the Bispectrum B_l. The bispectrum measures the strength, or correlation, of trilateral configurations, i.e. triangles in scale-space. Analogous to how the power spectrum is the scale-dependent representation of the variance, the bispectrum is the scale-dependent representation of the skewness. The skewness as introduced earlier in this chapter is the diagonal of the bispectrum, "$B_{l_1 l_1 l_1}$".

The bispectrum relates back to inflation through the coupling parameter f_{NL} in equation 2.15. The simplest single-field inflation models predict a value $|f_{NL}| < 1$, while the expected value for f_{NL} varies in the more elaborate inflationary scenarios [66]. Thus quantifying f_{NL} is an excellent estimator for different types of inflation. We can model the non-Gaussian a_{lm}s to second order

$$a_{lm} = a_{lm}^{G} + f_{\mathrm{NL}} a_{lm}^{NG}, \tag{3.46}$$

where the non-Gaussian term scales as $a_{lm}^{NG} \sim \left(a_{lm}^{G}\right)^2$. The coupling parameter f_{NL} is a measure of the strength of non-Gaussianity in the distribution, and the bispectrum is the lowest order moment able to discern Gaussian from non-Gaussian models [66]. Planck's 2018 release provided the constraints $f_{NL}^{\mathrm{local}} = -0.9 \pm 5.1$, $f_{NL}^{\mathrm{equilateral}} = -26 \pm 47$, $f_{NL}^{\mathrm{orthogonal}} = -38 \pm 24$, with the names local, equilateral and orthogonal referring to different configurations of the bispectrum [24].

The angular, normalized wavelet bispectrum is given by

$$\begin{aligned} I_{j_1 j_2 j_3} &= \sum_k \frac{\beta_{j_1 k}\beta_{j_2 k}\beta_{j_3 k}}{\sigma_{j_1 k}\sigma_{j_2 k}\sigma_{j_3 k}} \\ &= \sum_k \sum_{l_1 l_2 l_3} \sum_{m_1 m_2 m_3} \frac{b_{l_1}}{\sigma_{j_1 k}} \frac{b_{l_2}}{\sigma_{j_2 k}} \frac{b_{l_3}}{\sigma_{j_3 k}} a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \\ &\cdot Y_{l_1 m_1}(\hat{\gamma}_k) Y_{l_2 m_2}(\hat{\gamma}_k) Y_{l_3 m_3}(\hat{\gamma}_k), \end{aligned} \tag{3.47}$$

where β_{jk} are the wavelet coefficients for scale j and cubature point k, b_l is the weight function, in our case either for SMHW or standard needlets, σ_{jk} is the standard deviation Y_{lm}, a_{lm} the spherical harmonics and their coefficients, and $\hat{\gamma}_k$ is a directional vector [67, 68].

3.3.4 The four-point correlation function or the Trispectrum

The next higher order moment is the 4-point correlation function, or the Trispectrum T_l. The trispectrum measures the strength, or correlation, of quadrilateral configurations, i.e. quadrilaterals in scale-space. Again, the trispectrum represents a scale-dependent measure of the kurtosis, with the diagonal $T_{l_1 l_1 l_1 l_1}$ giving the traditional 1D kurtosis.

In the same way as the Bispectrum, it relates to the third order perturbations in the primordial gravitational potential, through the coupling parameter g_{NL}. We can model the non-Gaussian a_{lm}s to third order as

$$a_{lm} = a_{lm}^{G} + f_{\mathrm{NL}} a_{lm}^{\mathrm{NG}} + g_{\mathrm{NL}} a_{lm}^{\mathrm{NG},2}, \tag{3.48}$$

with superscript 2 referring to second order perturbation, not power. The second order non-gaussianity term scales as $a_{lm}^{\mathrm{NG},2} \sim \left(a_{lm}^{G}\right)^3$. Similarly to the bispectrum, the needlet trispectrum is given by

$$\begin{aligned} T_{j_1 j_2 j_3 j_4} &= \sum_k \frac{\beta_{j_1 k}\beta_{j_2 k}\beta_{j_3 k}\beta_{j_4 k}}{\sigma_{j_1 k}\sigma_{j_2 k}\sigma_{j_3 k}\sigma_{j_4 k}} \\ &= \sum_k \sum_{l_1 l_2 l_3 l_4} \sum_{m_1 m_2 m_3 m_4} \frac{b_{l_1}}{\sigma_{j_1 k}} \frac{b_{l_2}}{\sigma_{j_2 k}} \frac{b_{l_3}}{\sigma_{j_3 k}} \frac{b_{l_4}}{\sigma_{j_4 k}} \\ &\cdot a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} a_{l_4 m_4} \cdot Y_{l_1 m_1}(\hat{\gamma}_k) Y_{l_2 m_2}(\hat{\gamma}_k) Y_{l_3 m_3}(\hat{\gamma}_k) Y_{l_4 m_4}(\hat{\gamma}_k), \end{aligned} \tag{3.49}$$

3.3.5 Scale-dependent coupling parameter $f_{NL}(k)$ and $g_{NL}(k)$

The quantities f_{NL} and g_{NL} are often introduced as a constant, relating to the amplitude of the bi- and trispectrum. The simplest model of inflation, a single-field slow-roll Inflaton field, with canonical kinetic energy and initial conditions, infer that the amplitude and scale-dependence of the higher order correlation functions are of the order of the slow-roll parameters, meaning they are too small to observe at the present time. However, several inflationary scenarios allow for a scale-dependent bi- and trispectrum, such that the non-linearity parameters are no longer constants, but dependent on a wave k in Fourier space. With this kind of scale-dependent non-Gaussianity, the relations given in equations 3.46 and 3.48 no longer hold, as the connection of f_{NL} and g_{NL} to the second and third order non-Gaussianity terms shown in equation 2.15 can no longer be transferred linearly to the harmonic coefficients a_{lm}.

3.3.6 Motivation for the trispectrum

In this thesis, we will focus on the trispectrum, and not the bispectrum. The motivation behind this is that, for our model as described in section 2.6, we expect the presence of strong scale-dependent non-Gaussianity with g_{NL}-like terms, and $f_{\mathrm{NL}} \sim 0$. Therefore, when considering the a_{lm} coefficients, we would have

$$a_{lm} \sim a_{lm}^{G} + g_{\mathrm{NL}} \left(a_{lm}^{G}\right)^3. \tag{3.50}$$

Now consider a schematic overview of the bispectrum under the assumption that $f_{\mathrm{NL}} = 0$,

$$\begin{aligned}\langle a_{lm}a_{lm}a_{lm}\rangle &\sim \left\langle a_{lm}^3 + 3a_{lm}^2 g_{\mathrm{NL}}\left(a_{lm}\right)^3 + 3a_{lm}(g_{\mathrm{NL}})^2\left(a_{lm}\right)^6 + \left(g_{\mathrm{NL}}\right)^3\left(a_{lm}\right)^9\right\rangle \\ &= \left\langle a_{lm}^3 + 3g_{\mathrm{NL}}\left(a_{lm}\right)^5 + 3(g_{\mathrm{NL}})^2\left(a_{lm}\right)^7 + \left(g_{\mathrm{NL}}\right)^3\left(a_{lm}\right)^9\right\rangle\end{aligned} \tag{3.51}$$

We can see that we only get odd moments for the a_{lm} coefficients, and for a Gaussian distribution, the odd moments are zero. Thus we expect a zero bispectrum from our model. However, consider instead the trispectrum

$$\begin{aligned}\langle a_{lm}a_{lm}a_{lm}a_{lm}\rangle \sim &\Big\langle \left(a_{lm}^G\right)^4 + 4\left(a_{lm}^G\right)^3\left(g_{\mathrm{NL}}\left(a_{lm}^G\right)^3\right) + 6\left(a_{lm}^G\right)^2\left(g_{\mathrm{NL}}\left(a_{lm}^G\right)^3\right)^2 \\ &+ 4\left(a_{lm}^G\right)\left(g_{\mathrm{NL}}\left(a_{lm}^G\right)^3\right)^3 + \left(g_{\mathrm{NL}}\left(a_{lm}^G\right)^3\right)^4\Big\rangle \\ = &\Big\langle \left(a_{lm}^G\right)^4 + 4g_{\mathrm{NL}}\left(a_{lm}^G\right)^6 + 6\left(g_{\mathrm{NL}}\right)^2\left(a_{lm}^G\right)^8 + 4\left(g_{\mathrm{NL}}\right)^3\left(a_{lm}^G\right)^{10} \\ &+ \left(g_{\mathrm{NL}}\right)^4\left(a_{lm}^G\right)^{12}\Big\rangle.\end{aligned} \tag{3.52}$$

We obtain non-zero even moments, which is why we are interested in investigating the trispectrum, and not the bispectrum.

Chapter 4

Anomalies in the CMB

Many statistical anomalies have been reported being present in the CMB, which calls into question whether the CMB really is perfectly Gaussian and isotropically distributed. It is therefore interesting to investigate whether other types of models can reproduce these anomalies, and fit the observed data better than the traditional Gaussian and isotropic hypothesis. To reiterate from section 2.6, the focus of this thesis is on the following 6 anomalies present in the CMB:

1. The large-scale hemispherical power asymmetry as a dipolar modulation of an isotropic sky.

2. The small-scale hemispherical power asymmetry.

3. The Cold Spot, giving rise to an excess kurtosis of wavelet coefficients for angular scales around 10°.

4. The low power on large angular scales.

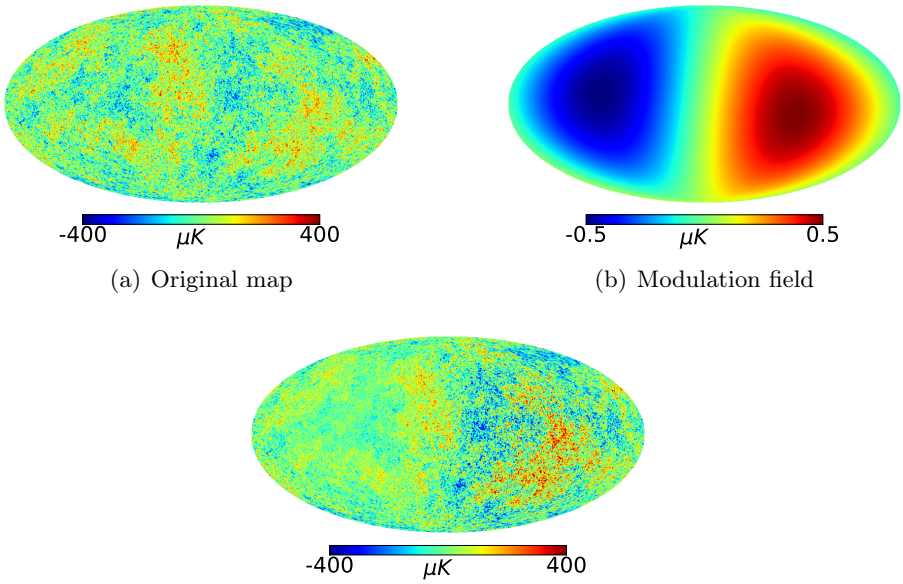
5. The quadrupole and octopole alignment.

6. Parity asymmetry on the large angular scales.

The following sections seek to provide a short description of each anomaly, along with evidence and methods for detection.

4.1 Large-scale hemispherical asymmetry

In the WMAP first-year data, local estimates of the angular power spectrum indicated an asymmetry in the distribution of power between two hemispheres in the sky. The anomaly was first reported by Eriksen et al. [69], and further confirmation was received from e.g. Hansen et. al [70]. A more recent study can also be seen in Akrami et al. [71]. A model consisting of an isotropic CMB sky modulated by a dipole was proposed



(a) Original map

(b) Modulation field

(c) Dipole modulated map

Figure 4.1: A dipole modulation of a map, resulting in large-scale hemispherical power asymmetry. Units are given in μK.

by Gordon [72], parametrized by a preferred direction $\hat{p} = (l, b)$ and a modulation amplitude A. The modulated signal is then given by

$$T(\hat{n}) = (1 + A\hat{p} \cdot \hat{n})B(\hat{n}), \tag{4.1}$$

where $\hat{n}$ is the direction in the sky, A is the modulation amplitude, $\hat{p}$ is the preferred direction, and B is an isotropic temperature field. A toy example showing the effect of dipole modulation can be seen in figure 4.1. From the figure, we can clearly see that fluctuations are enhanced in one hemisphere, that is the hot spots get hotter and the cold spots get colder. In the opposite hemisphere, fluctuations are suppressed smearing out fluctuations towards zero (i.e. the mean temperature).

The dipolar modulation model was found to fit the data by Eriksen et al. [73], reporting a dipole axis pointing towards galactic coordinates $(l, b) = (225°, -27°)$, with a best-fit modulation amplitude of 0.114. The analysis was extended up to multipoles $40 < l < 80$ by Eriksen et al. [74], and showed that the data prefers a dipolar modulation with amplitude $A = 0.072 \pm 0.022$ and preferred direction $(l, b) = (224°, -22°) \pm 24°$ detected at the 3.3σ level for $l \leq 64$. The latest Planck release [23] reports a dipole direction of $(l, b) = (221°, -20°)$, averaged over the 4 component separated maps Commander, NILC, SEVEM and SMICA with the combined dataset of temperature, polarization and the cross spectra (TT, EE, TE). The same averaging applied to the amplitude gives the value of $A = 0.0695$. For more information, see table 24 in Planck Collaboration VII [23].

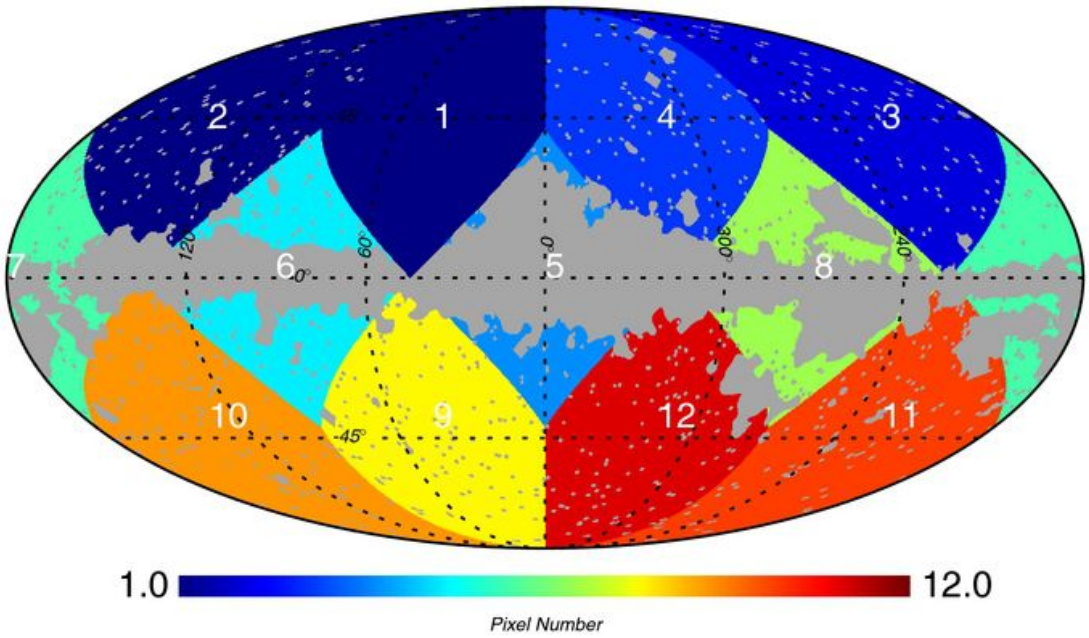


Figure 4.2: The 12 pixels used in analysis. Image credit: [76].

4.2 Small-scale hemispherical asymmetry

Similar to the hemispherical asymmetry on large scales, the asymmetry was found to extend to smaller scales, meaning higher multipoles l, first reported by Hansen et al. [75] in the WMAP 5 year-data, and later corroborated by Axelsson et al. [76] in the WMAP 9 year data, and also by the official Planck release [77, 78]. The asymmetry is not induced by a dipolar modulation as in the case for the largest scales; the modulation amplitude is within the expected range of Gaussian simulations. However, the alignment of the dipolar power distribution between multipoles persists, with a preferred direction close to what was found for the large-scale asymmetry.

Hansen et al. [75] performed power spectrum estimation on different sized disks of diameter $180°$(hemispheres)$, 90°, 45°, 22.5°$, using the MASTER algorithm for binned power spectra estimates. The analysis found a preferred direction $(l, b) = (226°, -17°)$ for multipole range $l = 2 - 600$, finding that the asymmetric models were favored over the isotropic models at 0.4% significance level.

The analysis was extended to parameter estimation by Axelsson et al. [76], finding a similar preferred direction for asymmetry, $(l, b) = (227°, -27°)$, at 3.4σ confidence level. The sky was divided into 12 patches, centered on the 12 base HEALPix pixels as shown in figure 4.2, where the value of each pixel was given by the binned power spectra for a given l-range. The dipole direction and amplitude was estimated for 6 separate multipole bins of size 100, in the range $l = [2 - 600]$. For an isotropic map,

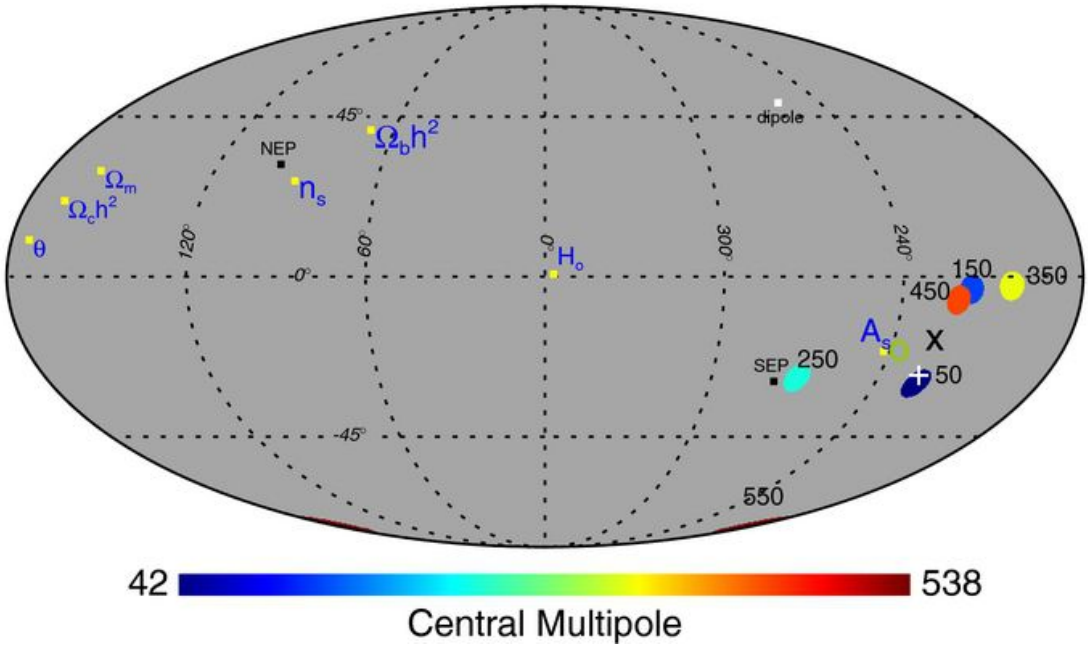


Figure 4.3: The filled circles shows the preferred dipole direction for a multipole bin, where the annotated number indicates the central multipole of this bin. The white cross, green circle and black cross represent the dipole direction estimates from the full range $l = 2 - 600$, the range $l = 2 - 40$ from WMAP 1 year data, and the full range $l = 2 - 600$ from the WMAP 5 year data respectively. NEP AND SEP represents the North and South ecliptic poles. The white square represents the cosmological CMB dipole. Dipole directions for the cosmological parameters are also shown as yellow squares, but these are not relevant for this section. Image credit: [76].

one would expect an uncorrelated power spectrum between the multipole bins, and thus randomly distributed dipole directions. However, their result shown in figure 4.3, shows a clear grouping of dipole directions between multipole bins.

4.3 The Cold Spot

During a blind Gaussianity test using SMHW statistics, performed by Vielva et al. on the WMAP first-year data, a significant positive excess of kurtosis was detected at scales around 10°. The excess kurtosis was attributed to a specific area in the southern hemisphere, a very large and cold spot in the CMB sky centered at $(b, l) = (-57°, 209°)$, aptly named "The Cold Spot". The kurtosis was so anomalous that for wavelet scales $R = 250'$ and $R = 300'$, it was outside the 1% acceptance interval, with a p-value of $p \approx 4 \cdot 10^{-3}$. The spot was also found to have a frequency dependence consistent with

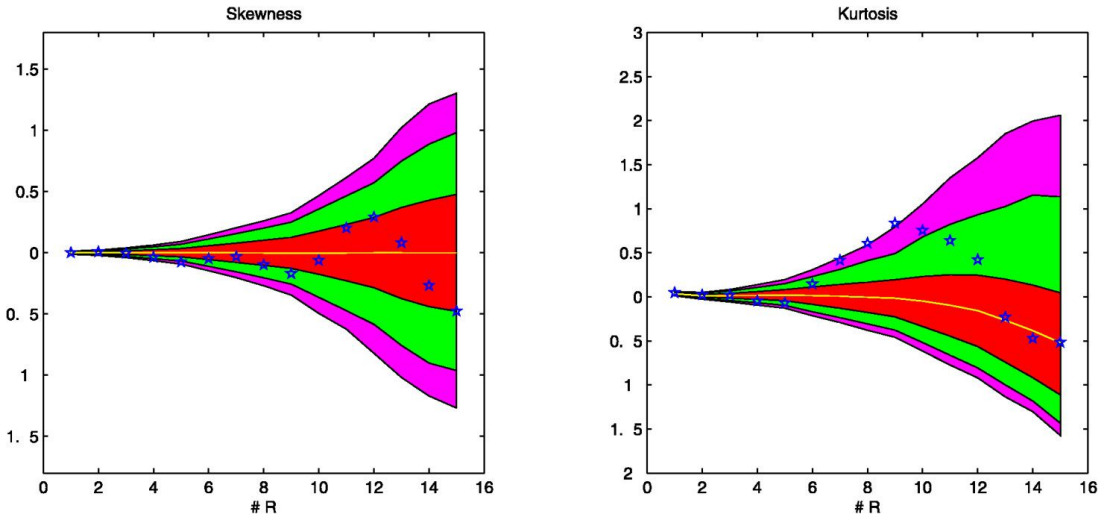
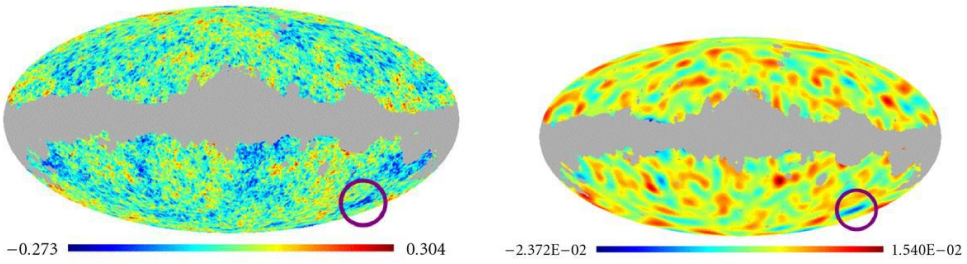


Figure 4.4: Positive deviation of the kurtosis at wavelet scales around $R_6 = 150'$ to $R_{13} = 750'$ in the WMAP data. The solid yellow shows the mean value obtained from 10000 Gaussian simulations, and the blue stars are the WMAP data points. The red, green and magenta regions show the 32%, 5% and 1% acceptance intervals, respectively. Image taken from [79].

the expected frequency behavior of the CMB [79]. Their detection of the positive excess kurtosis is shown in figure 4.4, showing that for scale number 8 and 9, corresponding to $R = 250'$ and $R = 300'$, the kurtosis is outside the magenta band representing 1% acceptance interval. This indicates a non-zero 4-point correlation function, inconsistent with Gaussianity. No significant deviation was found in the skewness. Later tests included measuring the amplitude with a MAX-statistic, calculating the area of the spot with Minkowski functionals, and a Higher Criticism (HC) test, all of which further confirmed its anomalous nature [80]. The Cold Spot consists of several smaller spots, none of which are very anomalous on their own in real space. The most notable single spot has a temperature of $\lesssim -350\mu K$, and a size of about 1° [79]. The Cold Spot is however very prominent at certain scales in wavelet space, shown in figure 4.5. We can see from this figure, even by eye, that the cold spot is a prominent feature in the wavelet map. Studies of the Spots morphology have shown it to be quite isotropic, almost being circular [81]. The Cold Spot is surrounded by a hot ring [78], making it very susceptible to being discovered through Spherical Mexican Hat Wavelets, as their shape matches very well (i.e. an inverted SMHW version of the Mexican Hat Wavelet shown in figure 3.12).



(a) Real Space (b) Wavelet space

Figure 4.5: The cold spot is circled, shown in real space and in wavelet space at scale $R = 250'$. Image credit [80].

4.4 Large-scale power deficit

One of the most debated anomalies in the CMB is the large-scale power deficit. The data shows consistently lower values than predicted from most mainstream cosmological models, at the largest angular scales. The quadrupole is especially low, originally reported as being anomalous at the 1-in-143 level [82] in WMAP. Further analysis by Tegmark et al. [83] argued that in a full-sky analysis, the significance dropped to a 1-in-20 level. This could be a statistical fluke, taking into account the large cosmic variance at low multipoles. There is also a possibility that the low quadrupole is connected with the parity asymmetry anomaly [84, 85, 86], and not a unique anomaly in its own right. The Planck team also reported the anomaly with a significance of $2.5-3\sigma$, manifesting as a general power deficit of $5-10\%$ at multipoles $l \leq 40$ compared to the Planck best-fit ΛCDM model [15]. The very low value of the quadrupole can be seen, even by eye, in the Planck TT-power spectrum presented in figure 1.1, while a statistic quantifying the general power deficit on the large scales can be seen in figure 4.6. In this statistic, the Planck and WMAP low-l likelihood was fitted to the Planck data, at different ranges of $2 \leq l \leq l_{\text{max}}$. If the low-l likelihood was a perfect fit, the expected amplitude would be 1. However, figure 4.6 shows the low-l spectrum trending significantly below the expected value of 1, indicating that the observed large scale power spectrum is anomalously low, by a factor of ~ 0.9 compared to the Planck best-fit model.

4.5 Quadrupole and octopole alignment

An apparent alignment between the quadrupole and octopole was found during a foreground analysis on the WMAP data performed by Tegmark et al. [83]. They reported a suppression of power along a common spatial axis, pointing towards $(b, l) \sim (60°, -110°)$. This alignment was (IS?) anomalous at the 1-in-60 level. Topological analyses excluded "plain bagel" small universe models as an explanation [87], as well as other twisted back-to-back models [88]. The alignment of the quadrupole and octopole can be seen in figure 4.7, from the original cleaned map by Tegmark et al. [83].

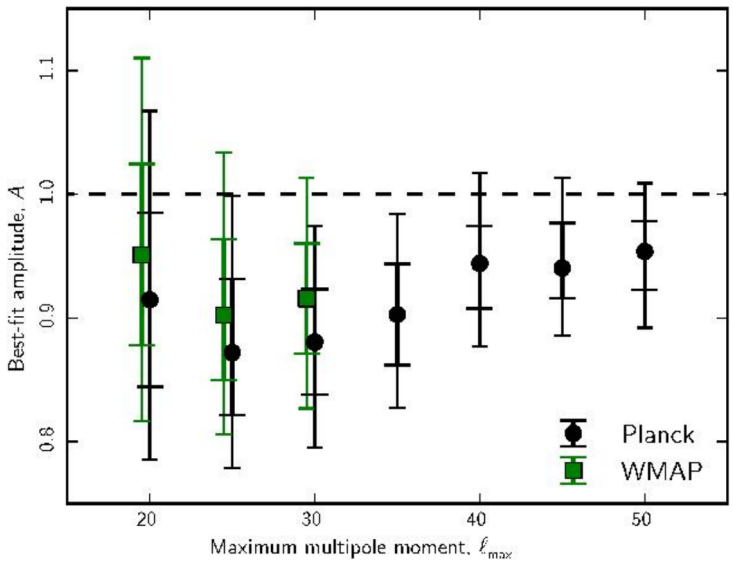


Figure 4.6: The relative amplitude of the low-l spectrum compared with the Planck best fit model. The error bars represent 68 and 95% confidence regions. Image credit: [15].

Figure 4.7 shows the alignment between the quadrupole and octopole (upper left and middle left panels), seemingly having their power supressed along a common axis.

4.6 Parity asymmetry on large angular scales

The CMB temperature anisotropy in the sky can be expanded in a sum of parity-symmetric $T^+(\hat{\mathbf{n}})$ and parity-antisymmetric $T^-(\hat{\mathbf{n}})$ functions, given by the relation

$$T^{\pm}(\hat{\mathbf{n}}) = \frac{1}{2}[T(\hat{\mathbf{n}}) \pm T(-\hat{\mathbf{n}})]. \tag{4.2}$$

The parity transformation in spherical coordinates corresponds to going from a point $\hat{\mathbf{n}}$ to the antipodal point $-\hat{\mathbf{n}}$:

$$r \to r \quad , \quad \theta \to \pi - \theta \quad , \quad \phi \to \phi + \pi. \tag{4.3}$$

In spherical harmonics, this then translates to

$$Y_{lm}(\theta, \phi) \to Y_{lm}(\pi - \theta, \phi + \pi). \tag{4.4}$$

From the definition of spherical harmonics,

$$Y_{lm}(\theta, \phi) = \underbrace{\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}}_{A} P_{lm}(\cos(\theta))e^{im\phi}, \tag{4.5}$$

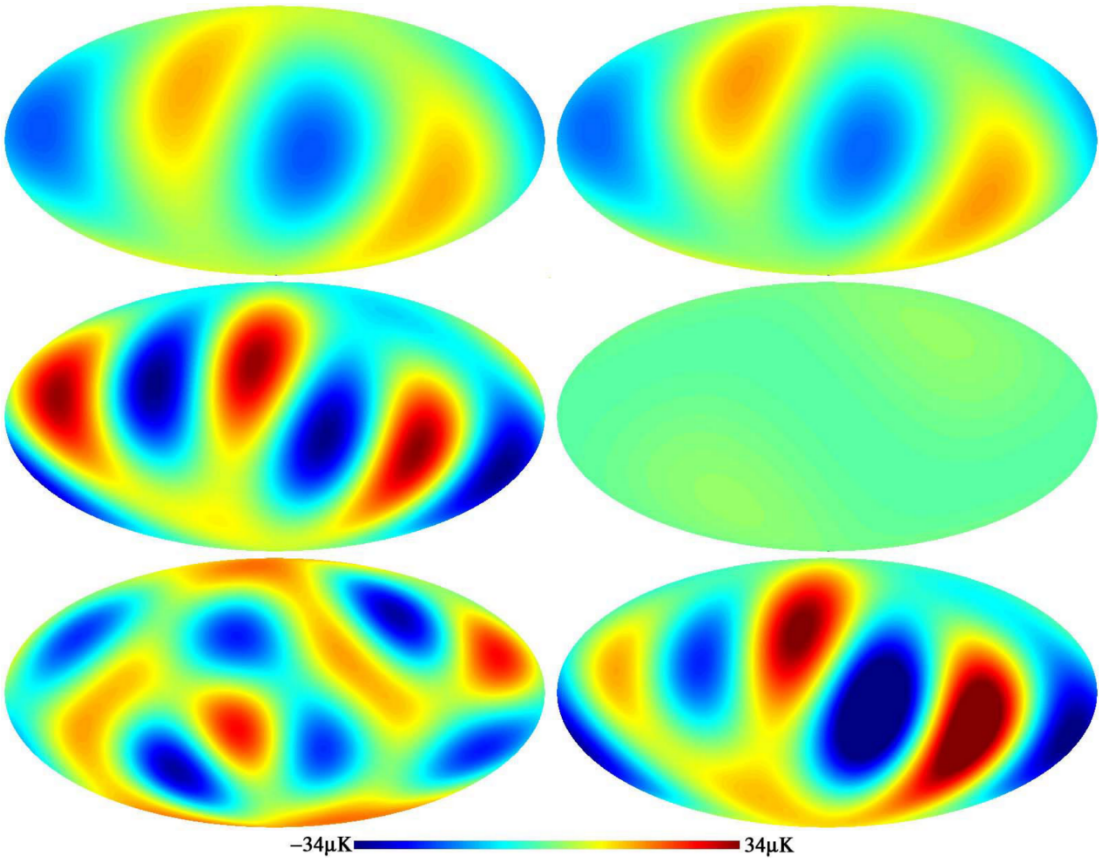


Figure 4.7: The quadrupole, octopole and hexadecapole is plotted from top to bottom in the left panel. The right panel show the cosmic quadrupole (top) after correcting for an estimate of the dynamic quadrupole (middle), resulting from our motion relative to the CMB. The lower right map shows the resulting map of adding the left panel quadrupole and octopole maps. Image credit: [83].

where P_{lm} is the associated Legendre polynomial, we then have

$$\begin{aligned} Y_{lm}(\pi-\theta,\phi+\pi) &= AP_{lm}(\cos(\pi-\theta))e^{im(\phi+\pi)} \\ &= AP_{lm}(-\cos(\theta))e^{im(\phi)}e^{i\pi m} \\ &\boxed{P_{lm}(-x) = (-1)^{l+m}P_{lm}(x)} \\ &= A(-1)^{l+m}P_{lm}(cos(\theta))e^{im\phi}(-1)^m \\ &= (-1)^l AP_{lm}(cos(\theta))e^{im\phi}(-1)^{2m} \\ &= (-1)^l Y_{lm}(\theta,\phi) \end{aligned}$$

or then equivalently

$$a_{lm}^{(P)} = (-1)^l a_{lm}, \tag{4.6}$$

where (P) represents the parity transformation. Thus, the symmetric function $T^+(\hat{\mathbf{n}})$ is made up of only spherical harmonics with even l-modes, while the anti-symmetric function $T^-(\hat{\mathbf{n}})$ conversely only with odd l-modes.

Because of isotropy, we expect a similar amplitude for the angular power spectra of the symmetric and anti-symmetric parts for $2 < l < 30$, on the largest angular scales [23]. However, a preference for the odd point-parity spectrum has been detected, at the $2-3\sigma$ significance level, in both WMAP data [89, 84, 90] and in Planck data [77, 78, 23]. Further studies [91] have also shown that this asymmetry can not be easily explained by the presence of residual Galactic foreground emission.

One statistic investigated in the Planck 2018 release [23], as well as the previous 2015 Planck data release [78], is the ratio

$$R^{\mathrm{TT}}(l_{\max}) = \frac{C_+^{\mathrm{TT}}(l_{\max})}{C_-^{\mathrm{TT}}(l_{\max})}, \tag{4.7}$$

where the temperature angular power spectra of the even and odd modes, C_+^{TT} and C_-^{TT} respectively, are given as

$$C_{+,-}^{\mathrm{TT}} = \frac{1}{l_{\mathrm{tot}}^{+,-}} \sum_{l=2,l_{\max}}^{+,-} \frac{l(l+1)}{2\pi} C_l^{TT}, \tag{4.8}$$

with $l_{\mathrm{tot}}^{+,-}$ as the total number of even and odd modes included in the sum to $l_{\max}$ [23]. The results of this statistic from the 2018 Planck data release can be seen in figure 4.8, where one can see that the statistic trends significantly below the expected value, showing a clear preference for the power spectra of the odd multipoles, C_-^{TT}.

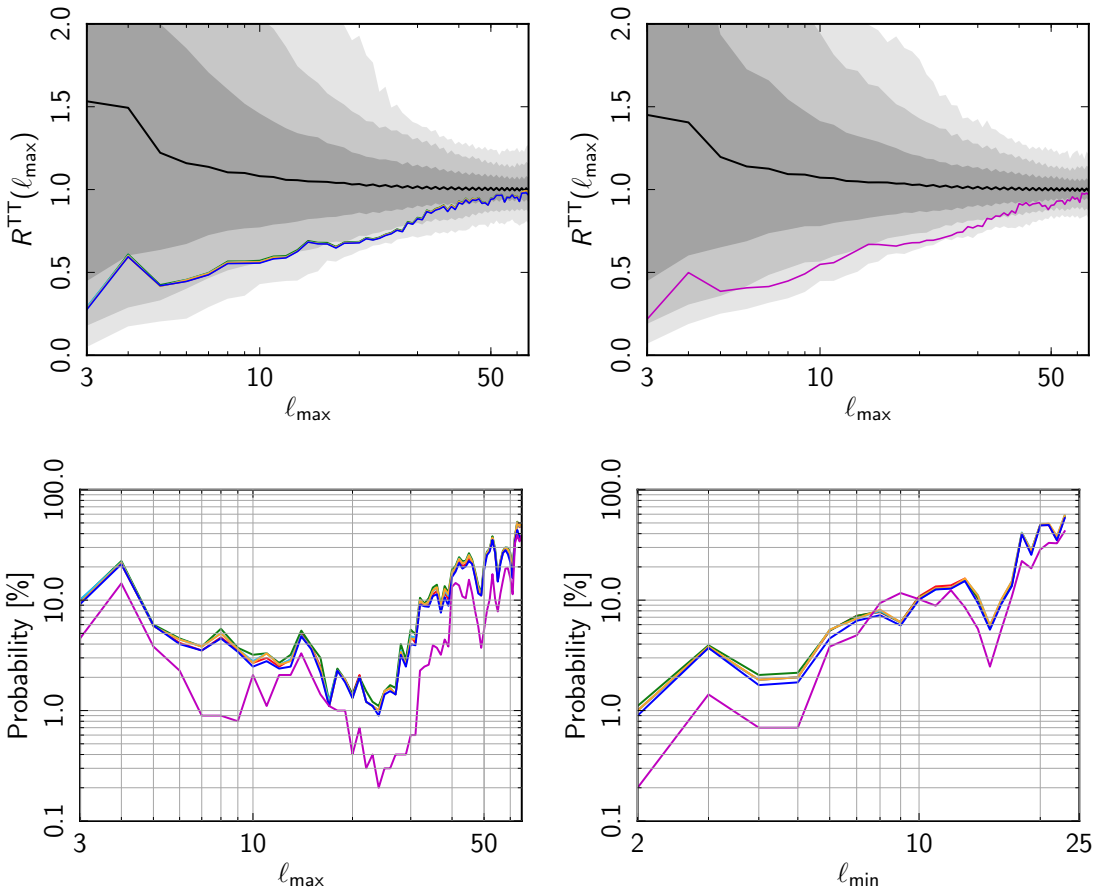


Figure 4.8: The statistic $R^{TT}(l_{max})$ applied to the Planck 2018 data. Upper left panel: the Black line along with the shaded areas are the expected distribution from SMICA MonteCarlo realizations, with the 1,2 and 3 σ bands. The red, orange, green and cyan lines are the ratios computed on maps with component separation Commander, NILC, SEVEM, and SMICA respectively, after application of the common mask. The upper right panel shows the same statistic, only with Lkl-Commander component-separated map (magenta line). The lower left and right panels show the lower-tail probability as a function of l_{max} and l_{min} respectively. Image credit: [23].

Chapter 5

Method

In this chapter, the methodology for investigating the non-Gaussian model presented in section 2.6 is outlined. First, we presents central estimators, then we discuss how noise is handled, before finally describing the technical algorithm for producing CMB realizations.

5.1 A simple needlet trispectrum estimator

Estimating the trispectrum is of interest, because of the connection to the presence of g_{NL}-like terms as discussed in section 3.3.6. We want to estimate the trispectrum to investigate whether it is a good estimator for detecting non-Gaussianity in our model, that is to check if it can distinguish Gaussian and non-Gaussian simulations. Referring back to equation 3.49, we wish to find an estimator for the (non-normalized) trispectrum,

$$T_{j_1 j_2 j_3 j_4} = \left\langle \sum_k \beta_{j_1,k}\beta_{j_2,k}\beta_{j_3,k}\beta_{j_4,k} \right\rangle \quad , \quad j_1 \leq j_2 \leq j_3 \leq j_4, \tag{5.1}$$

where β_j is the needlet coefficient for scale j, the angle brackets refer to ensemble average, and the sum goes over cubature points k. The HEALPix pixelation scheme is very handy in that we can take the pixel centers to be the cubature points, and therefore obtain the discrete estimator

$$\hat{T}_{j_1 j_2 j_3 j_4} = \frac{1}{N_{\text{pix}}} \sum_{k=1}^{N_{\text{pix}}} \beta_{j_1 k}\beta_{j_2 k}\beta_{j_3 k}\beta_{j_4 k}, \tag{5.2}$$

where k represents the pixels, to estimate the wavelet trispectrum as shown in equation 3.49. To ensure closed quadrilaterals in scale space (a quadrilateral must have 4 connected sides, or else it is only a line), we impose the quadrilateral condition

$$max(|j_1 - j_2 - j_3|, |j_2 - j_1 - j_3|, |j_3 - j_2 - j_1|) \leq j_4 \leq j_1 + j_2 + j_3. \tag{5.3}$$

Thus we can estimate the trispectrum for a single realization i, and obtain the estimated mean trispectrum

$$\bar{T}_{j_1 j_2 j_3 j_4} = \langle \hat{T}_{j_1 j_2 j_3 j_4} \rangle, \tag{5.4}$$

where the brackets denote the simulation average. Our initial choice of scales are

$$R_{\mathrm{SMHW}} = \{13.7, 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 600, 750, 900, 1050\}, \quad (5.5)$$

for the SMHW, which are given in arcminutes denoting how large the span of the wavelet is on the sky. These are the scales used by Vielva et al. in their original detection of the cold spot [79].

For the standard needlets, we use the scales

$$j = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16\}, \quad (5.6)$$

which spans the multipole range of $[B^{j_{\min}-1}, B^{j_{\max}+1}]$. In our case, with $B = 1.5$, this translates to $l \in [1.5, 985]$. Our choice of $B = 1.5$ is a trade-off where we feel we get decent localization in both the harmonic and real space. With this choice of B and values of j for the needlets, they will span about the same range as the SMHW. Additionally, this is the range in which we expect the most detection of anomalies from our non-Gaussian model, as explained in section 2.6.

5.1.1 The optimal needlet trispectrum estimator

Initial results were produced with the simple trispectrum estimator. However, we found that this estimator produced distributions with a relatively large standard deviation, which led to ambiguous results, as there was a significant overlap between the histograms for the Gaussian and non-Gaussian distributions. We decided to search for and implement a more robust estimator. An optimal needlet trispectrum estimator was presented by Troja et al. [92],

$$T_{j_1 j_2 j_3 j_4} = \sum_k \frac{\beta_{j_1,k}\beta_{j_2,k}\beta_{j_3,k}\beta_{j_4,k} + C_{j_1 j_2 j_3 j_4,k}}{\sigma_{j_1}\sigma_{j_2}\sigma_{j_3}\sigma_{j_4}} \quad , \quad j_1 \leq j_2 \leq j_3 \leq j_4, \quad (5.7)$$

where σ_j is the standard deviation for the needlet coefficients for scale j. The quantity $C_{j_1 j_2 j_3 j_4,k}$ is a quadratic term that serves to minimize the variance, defined as

$$\begin{aligned} C_{j_1 j_2 j_3 j_4,k} = &- \Gamma_{j_1 j_2}\beta_{j_3,k}\beta_{j_4,k} - \Gamma_{j_1 j_3}\beta_{j_2,k}\beta_{j_4,k} - \Gamma_{j_1 j_4}\beta_{j_2,k}\beta_{j_3,k} \\ &- \Gamma_{j_2 j_3}\beta_{j_1,k}\beta_{j_4,k} - \Gamma_{j_2 j_4}\beta_{j_1,k}\beta_{j_3,k} - \Gamma_{j_3 j_4}\beta_{j_1,k}\beta_{j_2,k} \\ &+ \Gamma_{j_1 j_2}\Gamma_{j_3 j_4} + \Gamma_{j_1 j_3}\Gamma_{j_2 j_4} + \Gamma_{j_1 j_4}\Gamma_{j_2 j_3}, \end{aligned} \quad (5.8)$$

with $\Gamma_{j_1 j_2} = \langle \sum_k \beta_{j_1,k}\beta_{j_2,k} \rangle$ is the simulation average from Gaussian simulations. Due to both computational restraints, and the fact that we only have so many Planck MC simulations available, we use $\Gamma_{j_1 j_2}$ from our previous simulations with the sub-optimal estimator. Although the sub-optimal estimator produced trispectra with a larger standard deviation, sanity checks discussed later in section 6.1.1 still showed good agreement, thus this choice should not lead to a bias.

5.2 The chi-squared test

As a blind test for how well the trispectra are able to distinguish between the Gaussian and non-Gaussian simulations, we employ a χ^2-statistic,

$$\chi^2_{\text{sim}} = \sum_i \left(\frac{T_i^{\text{sim}} - \left\langle T^G \right\rangle}{\sigma_i} \right)^2, \tag{5.9}$$

where i refers to a scale combination $j_1 j_2 j_3 j_4$, sim refers to a single simulation or realization of a CMB map, and obs signifies the simulated trispectra, either Gaussian, T_i^{G}, or non-Gaussian, T_i^{NG}. The standard deviation, σ_i, is the simulation average standard deviation of the trispectrum, σ_T, for a given combination i. The χ^2-test allows us to test how significant the trispectrum is as an estimator for detecting deviations from a Gaussian trispectrum in our simulated maps. A significant detection should yield different χ^2 distributions for the Gaussian and non-Gaussian cases.

5.3 The α-estimator

The model will produce a (simulation) averaged trispectrum, which can be considered as a Gaussian trispectrum with non-Gaussian contribution,

$$\bar{T}_i^{\text{MOD}}(\alpha) = \bar{T}_i^{\text{G}} + \alpha \bar{T}_i^{\text{NG}}, \tag{5.10}$$

where α serves as a parameter governing the strength of the non-Gaussian part, superscript MOD refers to the model spectra, i refers to a scale combination $j_1 j_2 j_3 j_4$, and the bar signifies simulation average for ease of notation. The quantity $\alpha \bar{T}_i^{\text{NG}}$ requires some explanation. Looking back at equation 2.25, taking care to notice that $T(\theta, \phi)$ refers to temperature and not the trispectrum, a model realization consists of a Gaussian term added with a non-Gaussian term

$$T(\theta,\phi) = T_G(\theta,\phi) + \underbrace{\gamma \left[T_G(\theta,\phi) T_F^2(\theta,\phi) \right]^{\text{Filtered}}}_{\text{NG part}}. \tag{5.11}$$

In this way, α is a measure of how large the contribution from the non-Gaussian term in equation 5.11 should be. Now consider a χ^2-statistic similar to the one in section 5.2, but testing the model hypothesis as a function of α,

$$\chi^2_{\text{sim}} = \sum_i \left(\frac{T_i^{\text{sim}} - \bar{T}_i^{\text{MOD}}(\alpha)}{\sigma_i} \right)^2. \tag{5.12}$$

We can obtain an estimator for this strength parameter, $\hat{\alpha}$, by considering the derivative of this χ^2-statistic;

$$\begin{aligned}\frac{\partial\chi^2}{\partial\alpha} &= \frac{\partial}{\partial\alpha}\sum_i\left(\frac{T^{\text{sim}} - \left(\bar{T}_i^{\text{G}} + \alpha\bar{T}_i^{\text{NG}}\right)}{\sigma_i}\right)^2 \\ &= 2\sum_i\left(\frac{T^{\text{sim}} - \left(\bar{T}_i^{\text{G}} + \alpha\bar{T}_i^{\text{NG}}\right)}{\sigma_i}\right)\left(-\frac{\bar{T}_i^{\text{NG}}}{\sigma_i}\right) \\ &= 2\sum_i\left(\frac{-T_i^{\text{sim}}\bar{T}_i^{\text{NG}} + \bar{T}_i^{\text{G}}\bar{T}_i^{\text{NG}} + \alpha\left(\bar{T}_i^{\text{NG}}\right)^2}{\sigma_i^2}\right).\end{aligned}$$

Setting $\frac{\partial\chi^2}{\partial\alpha} = 0$, we get

$$\begin{aligned}\frac{\partial\chi^2}{\partial\alpha} = 2\sum_i\left(\frac{-T_i^{\text{sim}}\bar{T}_i^{\text{NG}} + \bar{T}_i^{\text{G}}\bar{T}_i^{\text{NG}} + \hat{\alpha}\left(\bar{T}_i^{\text{NG}}\right)^2}{\sigma_i^2}\right) &= 0 \\ \hat{\alpha}\sum_i\frac{\left(\bar{T}_i^{\text{NG}}\right)^2}{\sigma_i^2} &= \sum_i\frac{T_i^{\text{sim}}\bar{T}_i^{\text{NG}} - \bar{T}_i^{\text{G}}\bar{T}_i^{\text{NG}}}{\sigma_i^2} \\ \hat{\alpha} &= \frac{\sum_i\frac{T_i^{\text{sim}}\bar{T}_i^{\text{NG}} - \bar{T}_i^{\text{G}}\bar{T}_i^{\text{NG}}}{\sigma_i^2}}{\sum_i\frac{\left(\bar{T}_i^{\text{NG}}\right)^2}{\sigma_i^2}} \\ \hat{\alpha} &= \frac{\sum_i\frac{\bar{T}_i^{\text{NG}}\left(T_i^{\text{sim}} - \bar{T}_i^{\text{G}}\right)}{\sigma_i^2}}{\sum_i\frac{\left(\bar{T}_i^{\text{NG}}\right)^2}{\sigma_i^2}}\end{aligned} \tag{5.13}$$

Equation 5.13 then quantifies the preference for the non-Gaussian model, that is, if $\hat{\alpha} = 1$, the data prefers exactly the strength parameter $\gamma = 0.43$ in our model, if $\hat{\alpha} = 2$ the data prefers $\gamma = 0.86$, and if $\hat{\alpha} = 0$ the data prefers a completely Gaussian model with no added non-Gaussian term.

5.4 Noise

In this section, we consider the effect of noise on our results, more specifically the effect of an erroneous noise model. When noise is present, as it always will be in real-life experiments, there will be a signal from the true CMB, and a noise signal

$$a_{lm} = a_{lm}^{\text{CMB}} + a_{lm}^{\text{noise}}. \tag{5.14}$$

The use of half-mission maps, that is two maps from the first and second half of the observation time, serves to minimize the noise as it allows for cross-spectra in which

noise can be assumed uncorrelated. For the power spectrum, if we have two half-mission maps with superscript 1 and 2,

$$\begin{aligned}\langle a^1_{lm} a^2_{lm}\rangle &= \left\langle \left(a^{\mathrm{CMB},1}_{lm} + a^{\mathrm{noise},1}_{lm}\right)\left(a^{\mathrm{CMB},2}_{lm} + a^{\mathrm{noise},2}_{lm}\right)\right\rangle \\ &= \left\langle a^{\mathrm{CMB},1}_{lm} a^{\mathrm{CMB},2}_{lm} + a^{\mathrm{CMB},1}_{lm} a^{\mathrm{noise},2}_{lm} + a^{\mathrm{CMB},2}_{lm} a^{\mathrm{noise},1}_{lm} + a^{\mathrm{noise},1}_{lm} a^{\mathrm{noise},2}_{lm}\right\rangle \\ &= \left\langle a^{\mathrm{CMB},1}_{lm} a^{\mathrm{CMB},2}_{lm}\right\rangle,\end{aligned}$$

as one can assume that neither the noise and the CMB, nor the noise from each half-mission map, is correlated. The case with needlet coefficients is completely equivalent,

$$\beta_{j_1} = \beta^{\mathrm{CMB}}_{j_1} + \beta^{\mathrm{noise}}_{j_1}, \tag{5.15}$$

and the same relation holds. However, when considering the (needlet) trispectrum, for two half-mission maps 1 and 2 and with superscript n indicating noise terms, we have

$$\begin{aligned}&T_{j_1 j_2 j_3 j_4} = \langle \beta^1_{j_1}\beta^1_{j_2}\beta^2_{j_3}\beta^2_{j_4}\rangle \\ &= \left\langle \left(\beta^{\mathrm{CMB},1}_{j_1} + \beta^{\mathrm{n},1}_{j_1}\right)\left(\beta^{\mathrm{CMB},1}_{j_2} + \beta^{\mathrm{n},1}_{j_2}\right)\left(\beta^{\mathrm{CMB},2}_{j_3} + \beta^{\mathrm{n},2}_{j_3}\right)\left(\beta^{\mathrm{CMB},2}_{j_4} + \beta^{\mathrm{n},2}_{j_4}\right)\right\rangle \\ &= \Bigg\langle \beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{CMB},2}_{j_3}\beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},1}_{j_1} + \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},2}_{j_3}\beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},1}_{j_2} + \beta^{\mathrm{CMB},2}_{j_3}\beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},1}_{j_2} \\ &+ \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},2}_{j_3} + \beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},2}_{j_3} + \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},1}_{j_2}\beta^{\mathrm{n},2}_{j_3} \\ &+ \beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},1}_{j_2}\beta^{\mathrm{n},2}_{j_3} + \beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{CMB},1}_{j_3}\beta^{\mathrm{CMB},2}_{j_4}\beta^{\mathrm{n},2}_{j_1} + \beta^{\mathrm{CMB},2}_{j_3}\beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},2}_{j_4} \\ &+ \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},2}_{j_3}\beta^{\mathrm{n},1}_{j_2}\beta^{\mathrm{n},2}_{j_4} + \beta^{\mathrm{CMB},2}_{j_3}\beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},1}_{j_2}\beta^{\mathrm{n},2}_{j_4} + \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4} \\ &+ \beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4} + \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{n},1}_{j_2}\beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4} + \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{CMB},2}_{j_3}\beta^{\mathrm{CMB},2}_{j_4} \\ &+ \beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},1}_{j_2}\beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4}\Bigg\rangle.\end{aligned} \tag{5.16}$$

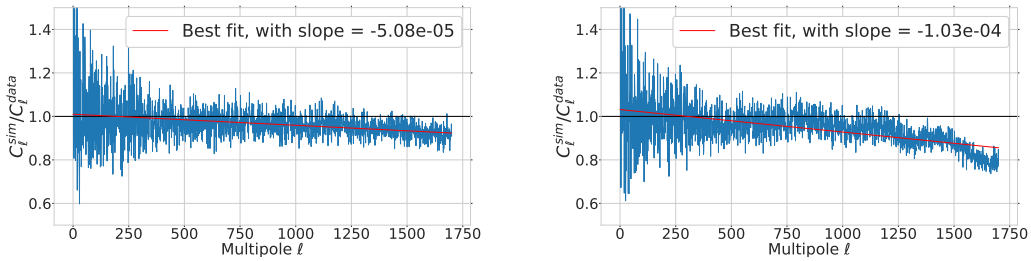
There are several correlated terms in the above expression which do not vanish. Not only will we obtain pure noise spectra of the type

$$\left\langle \beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},1}_{j_2}\beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4}\right\rangle = \left\langle \beta^{\mathrm{n},1}_{j_1}\beta^{\mathrm{n},1}_{j_2}\right\rangle \left\langle \beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4}\right\rangle, \tag{5.17}$$

but we also obtain noise terms amplified by the CMB signal, such as

$$\left\langle \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},1}_{j_2}\beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4}\right\rangle = \left\langle \beta^{\mathrm{CMB},1}_{j_1}\beta^{\mathrm{CMB},1}_{j_2}\right\rangle \left\langle \beta^{\mathrm{n},2}_{j_3}\beta^{\mathrm{n},2}_{j_4}\right\rangle. \tag{5.18}$$

These can be very hard to model if one does not fully know the noise properties of the detector. To get around this problem, we take a more empirical approach and try to estimate how well our noise model fits the data. To get an initial impression of how well the fit is, we extract a map with simulated CMB and noise, apply a mask, and compute the power spectrum C_l. The same procedure is done for the Planck data map. The



(a) SMHW (b) Standard needlets

Figure 5.1: The ratio of the power spectra, $\frac{C_l^{\text{sim}}}{C_l^{\text{data}}}$, computed from a simulated map with our noise model, and the Planck map, along with the best fit line and zero line for reference. The best fit as calculated without the mono- and dipole, as these are set to zero and the noise is not relevant.

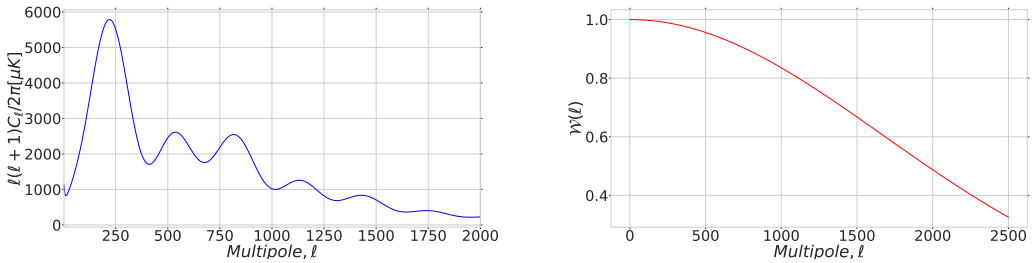
ratio of the two power spectra is plotted in figure 5.1, for both SMHW and standard needlets. Figure 5.1 shows that the disparity in the noise model is significant, especially at the highest multipoles. For the SMHW, the mean disparity between the simulated values and the data is 1.73% for the multipoles in the range $l \in [2, 1000]$, and 5.61% for the multipole range $l \in [1001, 1700]$. The corresponding numbers for the standard needlets are 1.93% and 10.97%. While this disparity is high for the large multipoles, $l = 1000$ and upwards is outside the range spanned by the wavelet coefficients discussed in section 5.1, and this should lessen the effect of the erroneous noise model. However, we still take precautions to minimize the effect of the noise, discussed in the following section.

5.4.1 Minimizing noise using α

We make an effort to reduce the effect of the noise model disparity, by comparing the effect on $\hat{\alpha}$. Since the highest scales are the most affected, a comparison was made by excluding the smallest scales, corresponding to the highest multipoles. We take a conservative approach and compare the noisy simulations to completely noise-free simulations, that is we assume a model with no noise and test $\hat{\alpha}$ on simulations with noise. This then represents the "worst-case scenario". Implementation-wise, we check the effect of noise on $\hat{\alpha}$ given in equation 5.15, by comparing the distributions from two cases:

1. T_i^{sim} from noise-free simulations,
2. T_i^{sim} from noisy simulations.

In both cases, the quantities $\bar{T}_i^{\text{NG}}$, $\bar{T}_i^{\text{G}}$ and σ_i are estimated from noise-free simulations. We perform this comparison while excluding more and more scales, that is disregard the trispectrum combinations containing these scales.



(a) Input C_l (b) SMICA beam

Figure 5.2: The input power spectrum and beam used to create CMB realizations.

5.5 Generation of CMB realizations

CMB maps are created from an input power spectrum, through HEALPix' *create_ alm* which generates Gaussian distributed a_{lm} coefficients of zero mean from a power spectrum. Furthermore, the HEALPix function *alm2map* creates a HEALPix temperature map with ring ordering pixel scheme, through a spherical harmonic transform. A window function, containing a Gaussian beam and pixel window function is added. Both the input power spectrum and the window function are shown in figure 5.2.

The process of estimating the trispectrum can be summed up as follows;

1. Create random CMB realizations from a given power spectrum C_l. The first $\frac{2}{3}$ simulations are Gaussian maps, used for comparison in statistical tests. The last $\frac{1}{3}$ simulations are used for non-Gaussian simulations.

2. Apply the process described in section 2.6 to the latter third of the simulations, for given values of the parameters γ and the filters w_ℓ and g_ℓ.

3. Perform a wavelet transformation on all the maps, to obtain the coefficients β.

4. Calculate the trispectrum $T_{j_1 j_2 j_3 j_4}$ as in section 5.1, along with $\Gamma_{j_1 j_2}$ as shown in section 6.1.1.

5. Step 2-4 can be repeated for different parameter values in order to test several models at once.

5.5.1 Masking the hot / cold spot

The process of creating the non-Gaussian maps will introduce a hot or cold spot into our maps [1], and from here on, it will simply be referred to as "the spot". We are interested in what effect the spot has on our results, and therefore the trispectrum is also calculated from the maps after masking the spot. This is achieved in the wavelet domain, where we search for the pixel with the largest SMHW coefficient in absolute value, and all pixels in a disc of radius $r = 5°$ centered on this pixel, are set to zero.

This radius is the actual radius of the cold spot as reported by Vielva [80] among others. For the standard needlets, this spot mask is then applied on a needlet map of the same realization.

5.6 The Planck 2015 FFP8.1 simulations

After the estimators have been validated through the randomly generated CMB realizations from an input power spectrum, the next step is to compare with simulations from the Planck experiment. We choose to use the Planck 2015 FFP8.1 simulation data, available from NERSC[1], to get a larger data set for better statistics as the 2018 release only had 1000 full-sky maps available [93]. Out of the total 4000 available Planck 2015 FFP8.1 maps, to simulate half-mission maps, we use 1920 of them with the corresponding 3840 noise maps, as our simulation code requires a simulation number divisible by 3.

5.7 Creation of a needlet mask

As wavelets and needlets are only semi-localized in space, they are susceptible to interference from masks, and the normal galactic and point source masks must be extended into scale-dependent masks. Here we explain the procedure for needlets. A similar procedure has been used to create the SMHW masks. We follow the approach presented by Scodeller et al. [60], using hyperbolic fits to deduce the critical angle θ_{crit} with which the mask must be extended. The extension angle θ_{crit} is given as

$$\theta_{\text{crit}} = \frac{\beta}{l^*}, \tag{5.19}$$

where l^* is the multipole for which the needlet has its peak, $l^* \approx B^j$, and β is a parameter to be fitted. For the standard needlets, β can be parametrized in terms of an acceptance criterion, τ, and the parameter B previously discussed in section 3.2.2. The acceptance criterion τ governs how much contamination from the mask is accepted. For example, for our choice of $\tau = 0.1$, 90% of the contamination from the mask is inside the critical angle θ_{crit}, and we thereby accept a 10% contamination from the mask in our data.

The hyperbolic fit depends on whether you are extending a galactic mask or a point source mask. For the galactic cut, we use

$$\beta_{\text{gal}} = (-0.337B + 4.98) \cdot \left(\frac{0.059 \cdot \tau^{-0.36}}{(B - 0.89)^2} + 0.44 \cdot \tau^{-0.32} \right), \tag{5.20}$$

while for the point source mask, or more specifically a Planck hole mask, we use

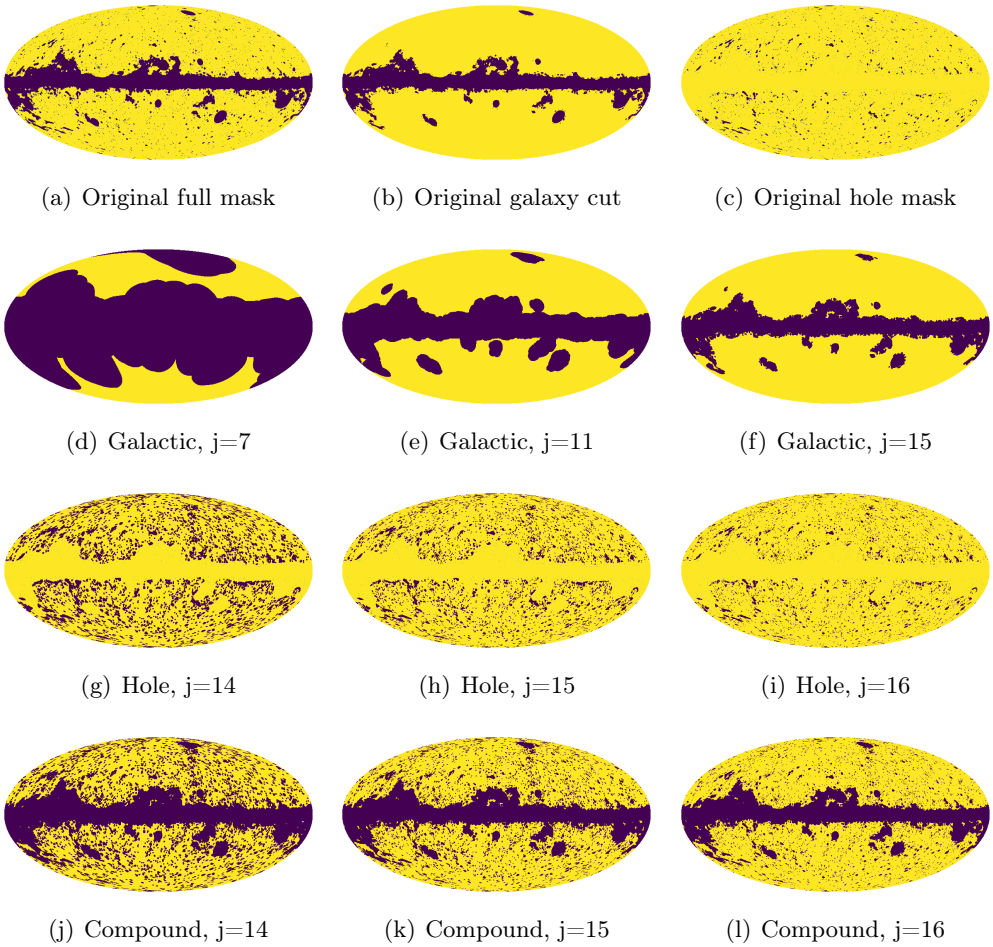
$$\beta_{\text{hole}} = (-0.337B + 4.98) \cdot \left(\frac{10^{f(\tau)}}{(B - 0.57)^2} + 0.18 \cdot \tau^{-0.30} \right), \tag{5.21}$$

[1] https://crd.lbl.gov/departments/computational-science/c3/c3-research/cosmic-microwave-background/cmb-data-at-nersc/

with

$$f(\tau) = \frac{-0.0065}{\sqrt{|\log_{10}(\tau)|}} \exp\left(\frac{1}{2}\frac{(\log_{10}(\tau) + 4.36)^2}{1.17^2}\right). \tag{5.22}$$

For the lowest scale needlets, corresponding to the largest scales on the sky, only the galactic mask is extended, as the point sources have very little impact on these scales. For the largest scale needlets, corresponding to small scales on the sky, both the galactic mask and the point source mask are extended, and then multiplied together to create a compound mask. The point source mask is obtained by $m_o + (1 - m_g)$, where m_o is the original full mask and m_g is the galaxy cut only. The "divide between the largest and smallest needlet scales" is found from table 1 and table 4 in Scodeller et al. [60]; The scales which have $l^* < 204$ are unaffacted by holes, and we therefore only extend the galactic mask, while for the scales with $l^* < 6$, the extension of the galaxy cut covers the whole sphere, i.e. the contamination from the galactic mask affects the whole sphere for these scales. Thus, these scales are rendered unusable, and they will be disregarded in the analysis. The scales with $l^* < 6$ correspond to needlet scales $j \in \{2, 3, 4\}$. A selection of the needlet masks created, with $N_{\text{side}} = 512$, along with the original masks, can be seen in figure 5.3.



(a) Original full mask (b) Original galaxy cut (c) Original hole mask

(d) Galactic, j=7 (e) Galactic, j=11 (f) Galactic, j=15

(g) Hole, j=14 (h) Hole, j=15 (i) Hole, j=16

(j) Compound, j=14 (k) Compound, j=15 (l) Compound, j=16

Figure 5.3: A selection of the extended masks for standard needlets, for $N_{\text{side}} = 512$. For the low needlet scales, $j < 5$, the extension of the galaxy cut covers the whole sphere, meaning these scales are unusable. The scales $j \in \{14, 15, 16\}$ have $l^* > 204$, and therefore the hole mask is also extended for these scales.

Chapter 6

Results

In this chapter, the results from the tests outlined in chapter 5 are presented. First, validation of the estimators are presented, before presenting the main results from the SMHW and needlet maps, first for the simple trispectrum, and then for the optimal trispectrum estimator.

6.1 Validation of estimators

6.1.1 Validation of the simple trispectrum estimator

It can be shown (see appendix A.2) that, for an isotropic field under isotropic noise, the analytic expression for the correlation of needlet coefficients is given as

$$\Gamma_{j_1 j_2} = \langle \beta_{j_1,k} \beta_{j_2,k} \rangle = \sum_{\text{shared l's}} b\left(\frac{l}{B^{j_1}}\right) b\left(\frac{l}{B^{j_2}}\right) \frac{2l+1}{4\pi} C_l \tag{6.1}$$

When comparing with actual simulations or data, we must also take the beam and pixel window into consideration. The convolution of these two functions are represented by the window function w_l, so the correlation of the needlet coefficients becomes

$$\Gamma_{j_1 j_2} = \langle \beta_{j_1,k} \beta_{j_2,k} \rangle = \sum_{\text{shared l's}} b\left(\frac{l}{B^{j_1}}\right) b\left(\frac{l}{B^{j_2}}\right) \frac{2l+1}{4\pi} C_l w_l. \tag{6.2}$$

The sum over "shared l's" refers to the fact that each needlet of scale j spans the range $l \in [B^{j-1}, B^{j+1}]$, and is zero outside. Therefore, only the overlapping needlets are correlated, when for two needlets with scale j_1, j_2, we have

$$\ell \in [B^{j_1-1}, B^{j_1+1}] \cup [B^{j_2-1}, B^{j_2+1}]. \tag{6.3}$$

A similar result can be obtained for wavelets, simply exchanging the shape of the weights b_l, and allowing the sum to go over the whole l-range, not simply the shared l's between j_1 and j_2. With the expression in equation 6.2, we can compare the simulated and analytic trispectrum mean values. For centered (of zero mean), Gaussian

random variables we should have, by the Wick Theorem [94] or the Isserlis Theorem [95] depending on the field, that

$$\langle \beta_{j_1}\beta_{j_2}\beta_{j_3}\beta_{j_4}\rangle = \langle \beta_{j_1}\beta_{j_2}\rangle \langle \beta_{j_3}\beta_{j_4}\rangle + \langle \beta_{j_1}\beta_{j_3}\rangle \langle \beta_{j_2}\beta_{j_4}\rangle + \langle \beta_{j_1}\beta_{j_4}\rangle \langle \beta_{j_2}\beta_{j_3}\rangle . \tag{6.4}$$

With our analytic expression for the quantity $\langle \beta_{j_1}\beta_{j_2}\rangle$, we can then perform the comparison of

$$\frac{\langle \beta_{j_1}\beta_{j_2}\beta_{j_3}\beta_{j_4}\rangle_{\text{sim}} - \langle \beta_{j_1}\beta_{j_2}\beta_{j_3}\beta_{j_4}\rangle_{\text{analytic}}}{\sigma_{j_1 j_2 j_3 j_4}} \tag{6.5}$$

for a given combination of $j_1 j_2 j_3 j_4$, where the standard deviation of the trispectrum is given as the square root of the variance, and the variance for a given scale combination being

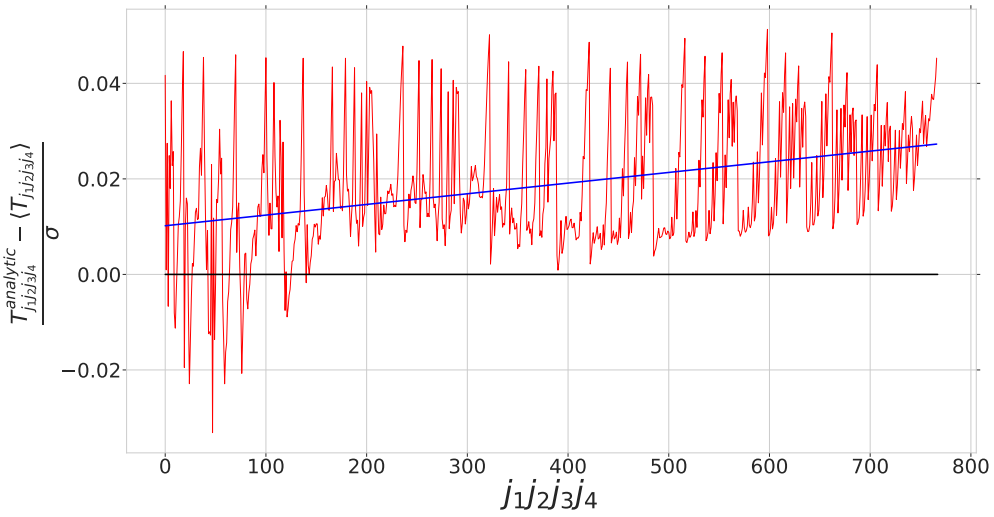
$$\sigma^2_{j_1 j_2 j_3 j_4} = \frac{1}{N}\sum_i \left(T^i_{j_1 j_2 j_3 j_4} - \langle T_{j_1 j_2 j_3 j_4}\rangle\right)^2, \tag{6.6}$$

where i goes over all Gaussian simulations, and N is the total number of Gaussian simulations. A plot of the test in equation 6.5 is shown in figure 6.1, for both SMHW and standard needlets. For the SMHW, there seems to be a consistent bias in the trispectrum, with the simulated values trending consistently below the analytic values. This could indicate that the simple estimator underestimates the trispectrum. The difference is however very small, at only about 0.02σ, which will be overshadowed by cosmic variance. For the standard needlets. there is no observable bias, but the spread around 0 indicates that the convergence is not perfect and could benefit from a larger number of simulations.

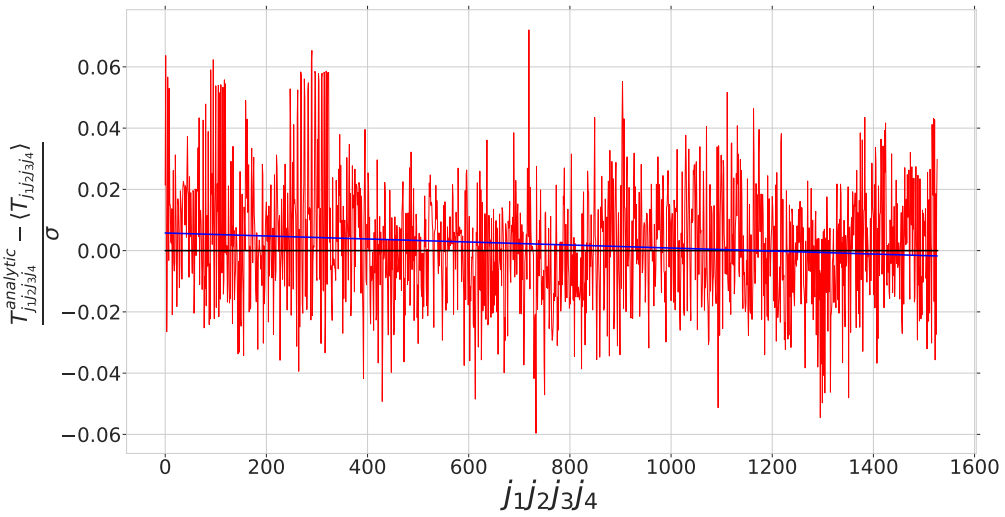
6.1.2 Validation of the χ^2 test

We validate the χ^2-estimator on 4800 simulated maps, split into two sets of 1600 Gaussian and 800 non-Gaussian simulations, one set for finding the average $\langle T^{\text{G}}\rangle$ and standard deviation σ_T, and one for comparing the simulated values T^{sim}. First and foremost, we can see from figure 6.2 that the χ^2 statistic is able to distinguish between the Gaussian and non-Gaussian simulations, as there is a distinct shift in the bulk of the distribution, as well as the mean. As the χ^2 distribution is highly asymmetric, we will use $68.27\%, 95.45\%$ and 99.73% limits instead of standard deviation to quantify the separation. For the SMHW, 88.38%, 57.63%, and 21.5% of the non-Gaussian distribution lie outside the $68.27\%, 95.45\%$ and 99.73% limits of the Gaussian distribution respectively. For the standard needlets, the equivalent percentages are 65.00%, 12.25%, and 3.62%, as the kurtosis is visibly smaller for the non-Gaussian standard needlet distribution, and the overlap is larger compared to the SMHW.

As a secondary indication, looking at the Gaussian simulations only, a χ^2 distribution where the degrees of freedom (D.O.F.), k, approaches infinity, $k \to \infty$, we should have that the mean approaches the D.O.F., $\langle \chi^2\rangle \to k$, by the central limit theorem. This is under the assumption that the distribution is made up of Gaussian random variables, that is to say, they will be Gaussian distributed. While we don't expect every trispectrum combination to be Gaussian distributed, many of them will, and thus the

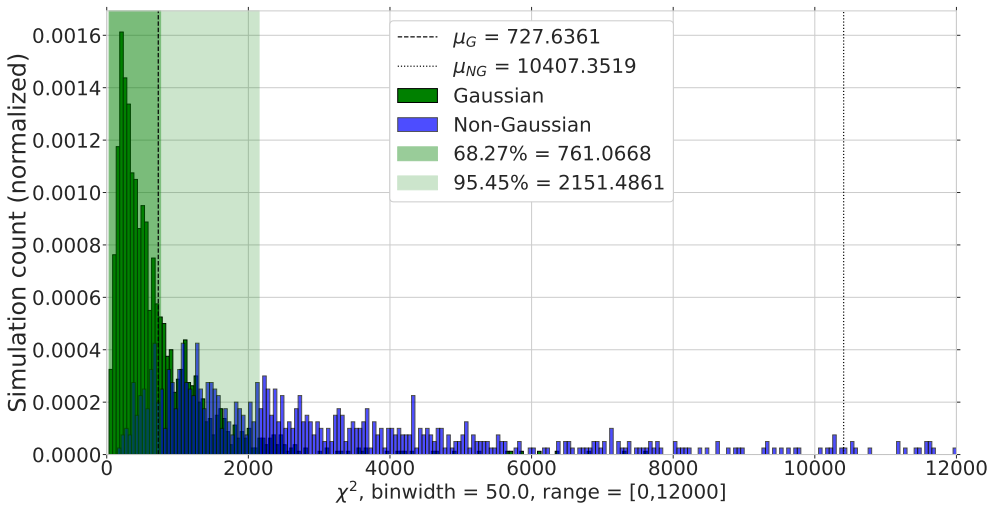


(a) SMHW

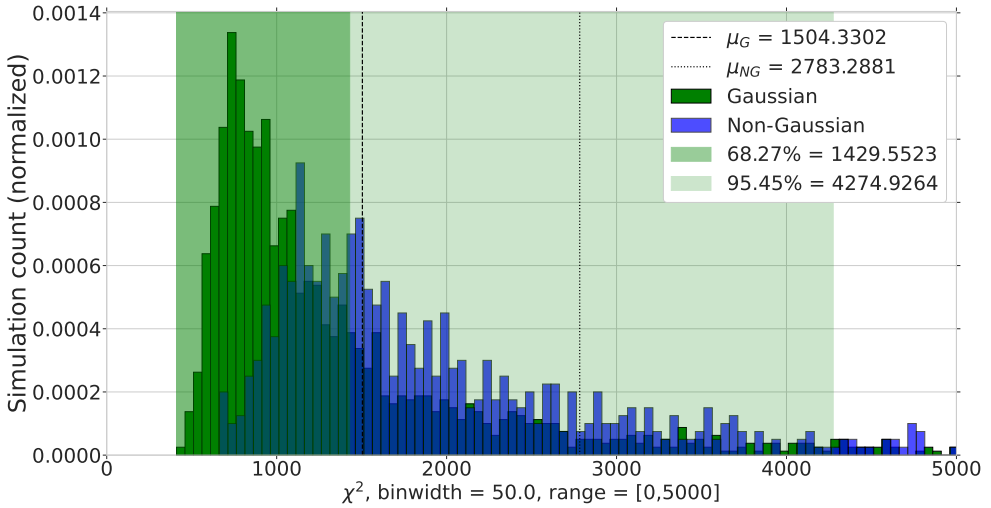


(b) Standard needlets

Figure 6.1: Difference between the mean of Gaussian simulations and analytic values, over 1σ, with the mean estimated from 3600 simulations without noise and masking. The best fit line is plotted in blue, while the zero line is plotted in black for reference.



(a) SMHW



(b) Standard needlets

Figure 6.2: Validation of the χ^2 estimator. The shaded green areas represent the intervals which contains 68.27% and 95.45% of the Gaussian distribution. The dashed line is the mean of the Gaussian distribution, while the dotted line is the mean of the non-Gaussian distribution. For the SMHW, there are outliers in the non-Gaussian distribution up to $\chi^2 = 635109$ not shown, while for the standard needlets, outliers up to $\chi^2 = 42643$ are omitted. Validated on 3600 Gaussian and 1200 non-Gaussian simulations, without noise and masking, for each of the wavelet shapes.

mean of the χ^2 distribution will be approximately equal to the D.O.F. The distribution of a random selection of trispectrum combinations, for 3600 simulated Gaussian maps without noise and mask, can be seen in figure 6.3. Typically, the larger needlet scales, corresponding to the smaller angular scales on the sky, are approximately Gaussian distributed, while the lower scales are not. This is confirmed by further testing not shown in the figure. This is expected by the Central Limit Theorem, as the larger needlet scales are localized at the high multipoles with a much larger sample size (the same principle as cosmic variance discussed in section 2.5.2). So while each trispectrum combination is not a Gaussian random variable, they will approach a Gaussian distribution given enough samples.

In our case, the degrees of freedom are the number of accepted trispectrum combinations minus one, that is $k_{\mathrm{SMHW}} = 766$ and $k_{\mathrm{Standard}} = 1526$. Our mean for the Gaussian distribution is indeed close to these values, but keep in mind this is not an exact test, and should only serve to complement other sanity checks.

6.1.3 Validation of $\hat{\alpha}$

A consistency check for our estimator $\hat{\alpha}$ is to check whether the estimator reproduces a distribution of mean 0 and 1 when T_i^{sim} is Gaussian and non-Gaussian simulated trispectra respectively. This, of course, because

$$\langle \hat{\alpha_{\mathrm{G}}} \rangle = \left\langle \frac{\sum_i \frac{\bar{T}_i^{\mathrm{NG}} \left(T_i^{\mathrm{G}} - \bar{T}_i^{\mathrm{G}} \right)}{\sigma_i^2}}{\sum_i \frac{\left(\bar{T}_i^{\mathrm{NG}} \right)^2}{\sigma_i^2}} \right\rangle \quad , \quad \langle \hat{\alpha_{\mathrm{NG}}} \rangle = \left\langle \frac{\sum_i \frac{\bar{T}_i^{\mathrm{NG}} \left(T_i^{\mathrm{NG}} - \bar{T}_i^{\mathrm{G}} \right)}{\sigma_i^2}}{\sum_i \frac{\left(\bar{T}_i^{\mathrm{NG}} \right)^2}{\sigma_i^2}} \right\rangle, \tag{6.7}$$

And as $\left\langle T_i^{\mathrm{G}} \right\rangle = \bar{T}_i^{\mathrm{G}}$ and $\left\langle T_i^{\mathrm{NG}} \right\rangle - \bar{T}_i^{\mathrm{G}} = \bar{T}_i^{\mathrm{NG}}$, it's easy to see that we should have $\langle \hat{\alpha_{\mathrm{G}}} \rangle = 0$ and $\langle \hat{\alpha_{\mathrm{NG}}} \rangle = 1$. Such a validation was done on 4800 simulations split into two sets of 1600 Gaussian and 800 non-Gaussian simulations, one set for finding the simulation averages $\bar{T}_i^{\mathrm{G}}$ and $\bar{T}_i^{\mathrm{NG}}$, and one set for T_i^{sim}, both for SMHW and standard needlets. The results can be seen in figure 6.4 for the SMHW, and 6.5 for the standard needlets, showing good agreement with the expected mean of 0 and 1.

6.1.4 The effect of noise on $\hat{\alpha}$

We return to results of the noise test presented in section 5.4.1. For the SMHW, we find that the exclusion of the three smallest wavelet scales, $R = \{13.7, 25.0, 50.0\}$ corresponding to the smallest scales on the sky, provides an acceptable fit, that is the histograms of the noise-free and noisy simulations are almost completely overlapping. Excluding more scales does not improve the fit meaningfully, while simultaneously giving worse statistics. The best-fit plot with the exclusion of the first three scales is shown in figure 6.6. There is no sharp cut-off in the multipole range spanned by these scales, as the SMHW are not sharply localized, but looking back at figure 3.15, this will leave very little amplitude above $\sim l = 200$.

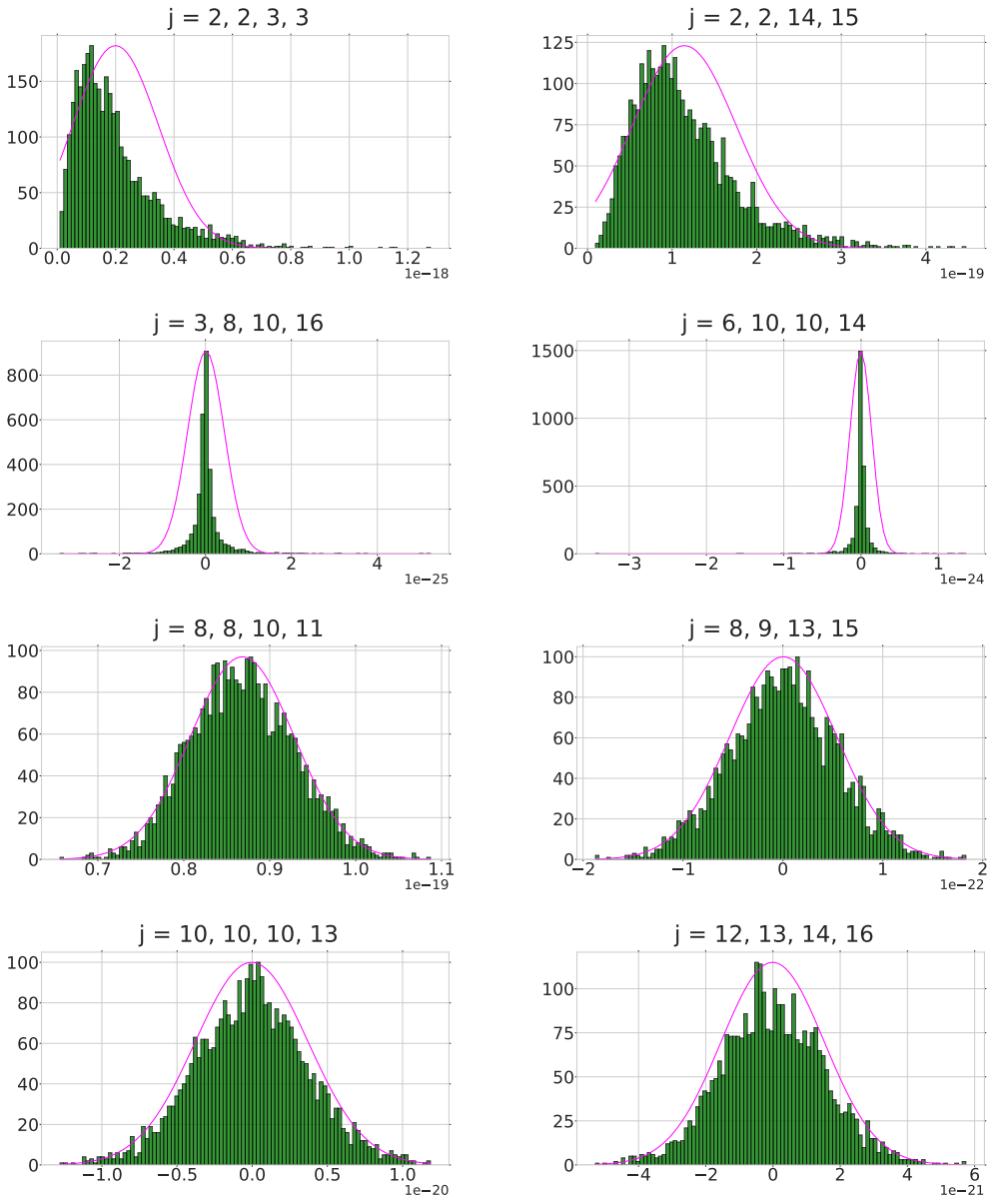
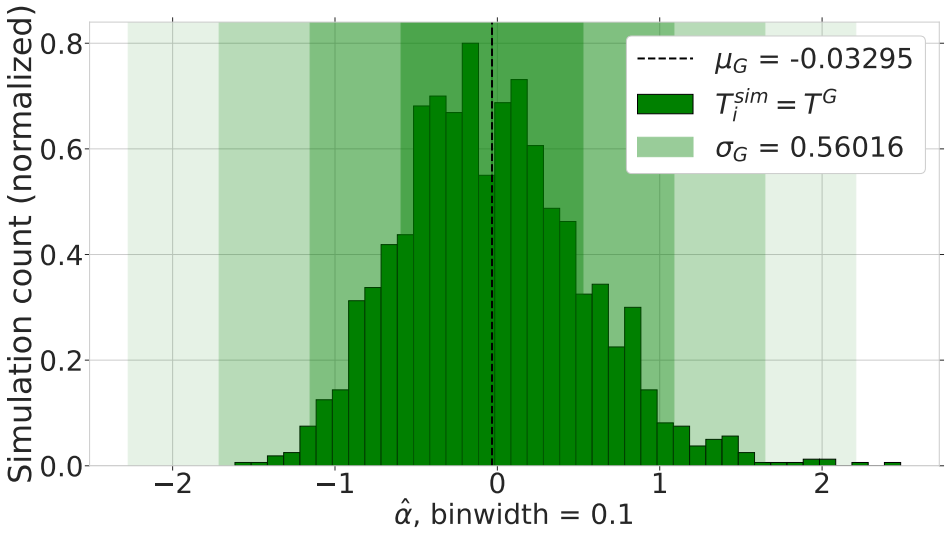
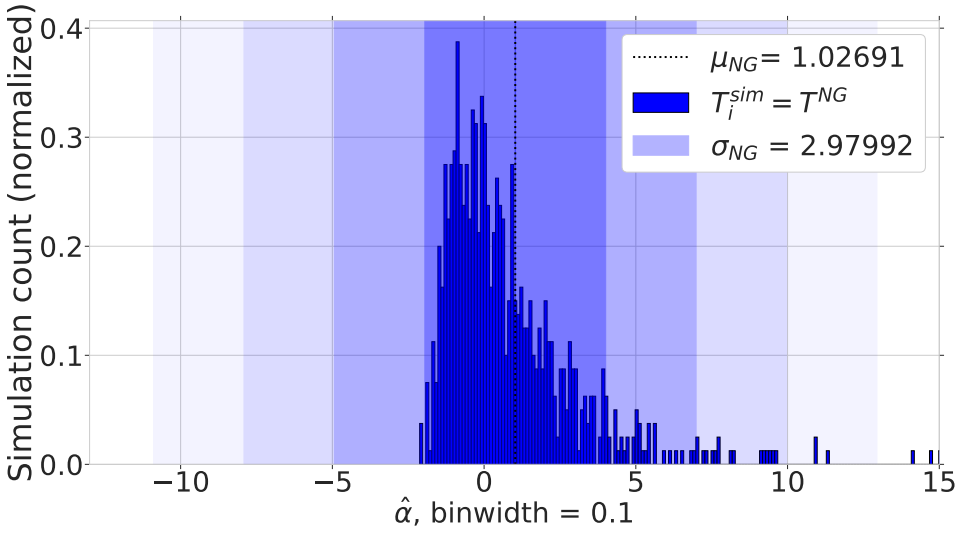


Figure 6.3: Histogram of a given trispectrum combinations for for 3600 Gaussian simulations. Fitted with a Gaussian curve (magenta line) of the same mean and standard deviation. Shown here for the standard needlets, but similar results are obtained for the SMHW.

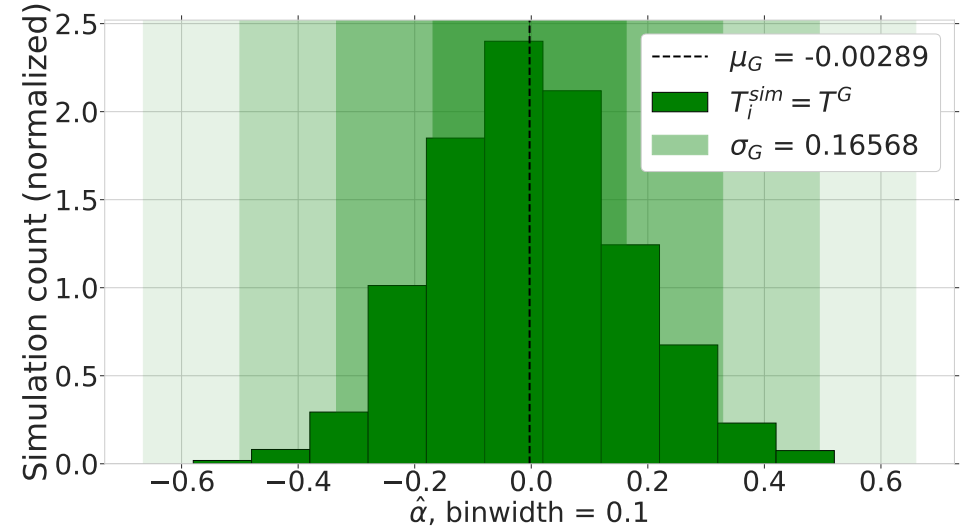


(a) Gaussian (SMHW)

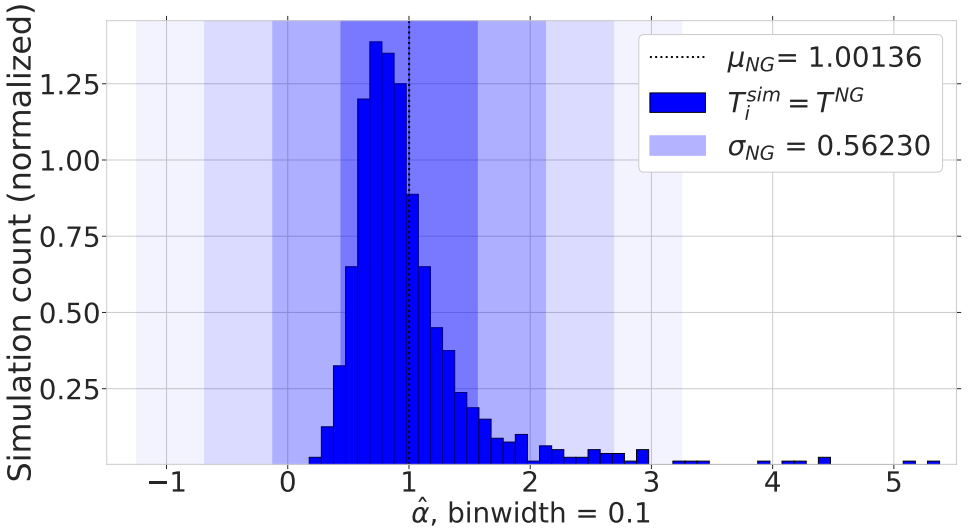


(b) Non-Gaussian (SMHW)

Figure 6.4: The normalized Gaussian and non-Gaussian distributions for $\hat{\alpha}$, with the 1,2,3,4 σ bands and mean plotted in for the SMHW. In the non-Gaussian distribution, outliers up to $\hat{\alpha} = 26.32$ have been omitted for readability. Validated on 3600 Gaussian and 1200 non-Gaussian simulations, without noise and masking.

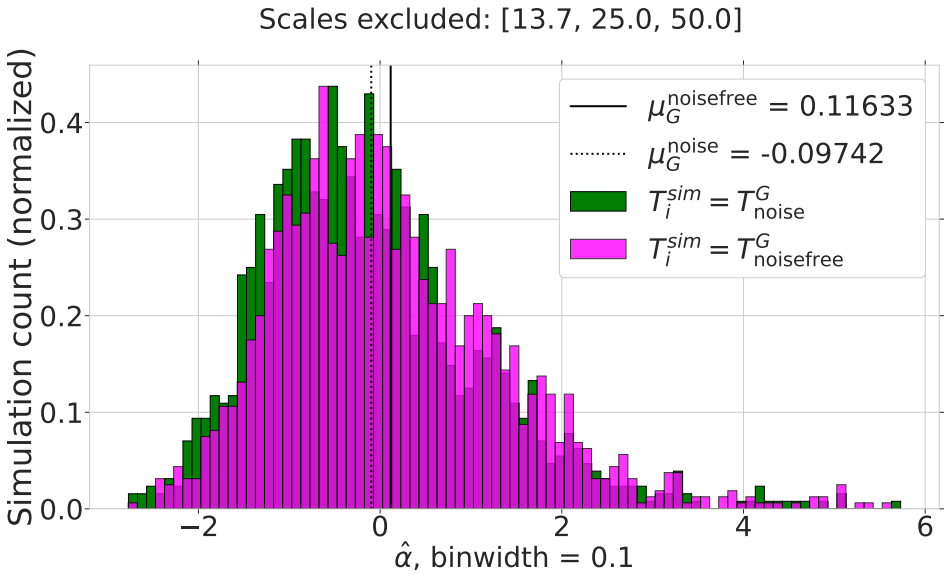


(a) Gaussian (Standard needlets)

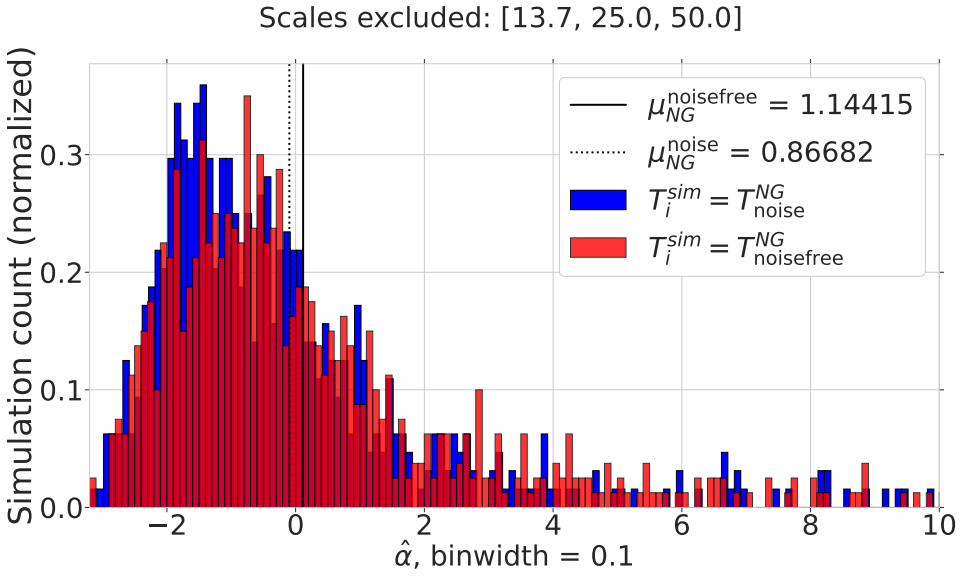


(b) Non-Gaussian (Standard needlets)

Figure 6.5: The normalized Gaussian and non-Gaussian distributions for $\hat{\alpha}$, with the 1,2,3,4 σ bands and mean plotted in for the standard needlets. All bins are shown. Validated on 3600 Gaussian and 1200 non-Gaussian simulations, without noise and masking.



(a) Gaussian



(b) Non-Gaussian

Figure 6.6: The $\hat{\alpha}$-estimator for SMHW, shown for both noisy and noise-free simulations, with the wavelet scales $[13.7, 25.0, 50.0]$ excluded. In the non-Gaussian distribution, outliers of up to $\hat{\alpha} = 118$ have been cropped for visibility.

For the standard needlets, we deem that the exclusion of the four largest j-values, $j = \{13, 14, 15, 16\}$ corresponding to the smallest scales on the sky, provides an acceptable fit. With these choices of excluded scales for the SMHW and needlets, the offset between the noise-free and noisy histograms are below 0.4σ for their respective validated $\hat{\alpha}$ distributions, and keep in mind that this is a worst-case scenario. Our noise model, while probably not perfect, is sure to yield a better fit to the noise in the Planck data than no noise at all. This exclusion limits our multipole range at the upper end to $l = 194$, still about the same range as the SMHW.

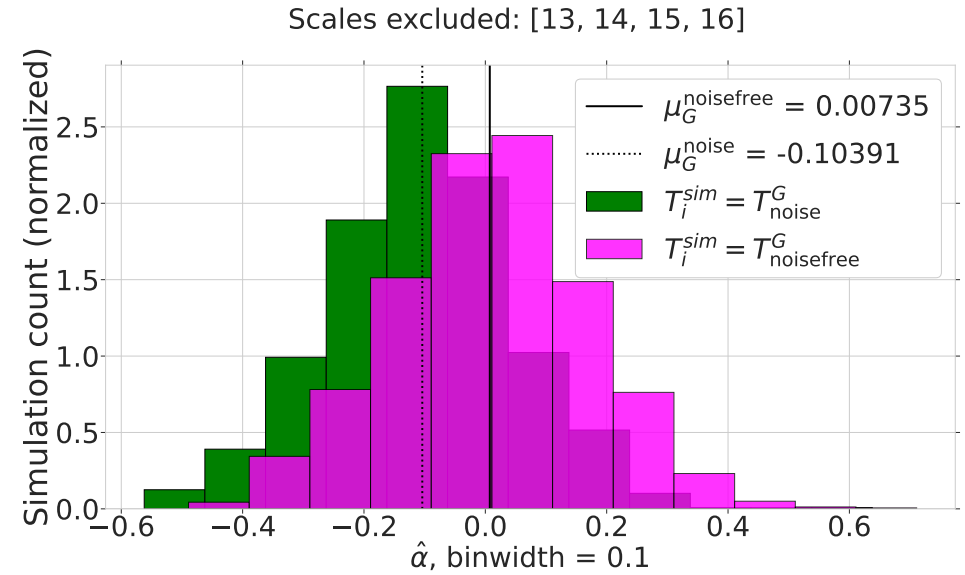
For comparison, the noise test with no scales excluded is shown in figure 6.8 for the SMHW and standard needlets. A broader selection of excluded scales tested, can be seen in appendix B.1 and B.2 for SMHW and standard needlets respectively.

6.1.5 Convergence with the Planck 2015 FFP8.1 simulations

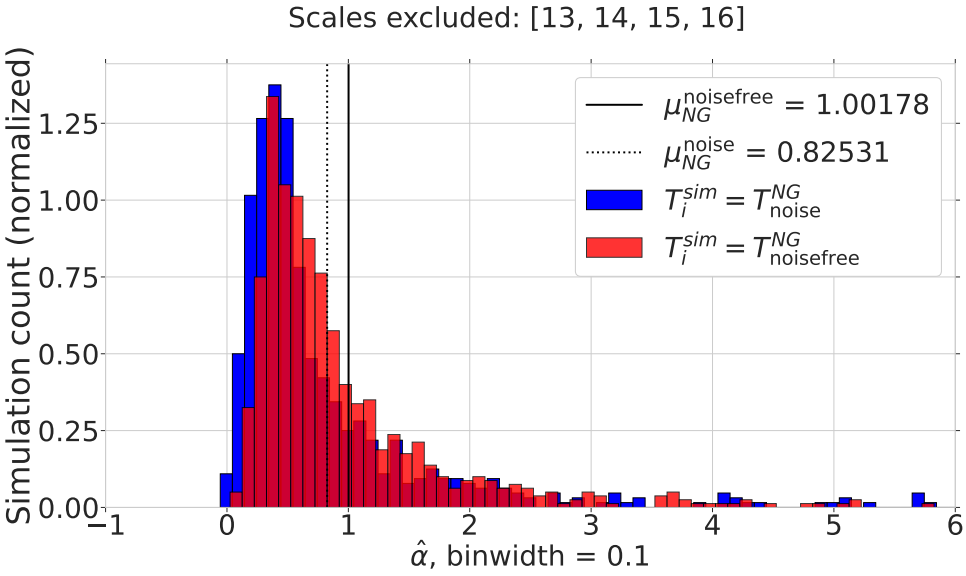
For our 1920 simulated maps from the 2015 Planck data release, with mask and noise, our initial choice is to proceed with our usual split of 1280 (2/3) Gaussian and 640 (1/3) non-Gaussian simulations. However, the initial result hinted at unsatisfactory convergence in the non-Gaussian $\hat{\alpha}$ distribution, as the mean was either too low or too high depending on which half of the data set, i.e. 1280/2 Gaussian and 640/2 non Gaussian simulations, was used for T^{sim} and which half was used to compute the means $\bar{T}^{\mathrm{G}}$ and $\bar{T}^{\mathrm{NG}}$. The two different cases are shown in figure 6.9 (only the non-Gaussian distribution is plotted), with the mean of the non-Gaussian distribution being 0.51 or 1.59 depending on which half was used for T^{sim}. In addition, the standard deviation was especially large for the high mean of 1.59, at $\sigma_{\mathrm{NG}} = 11.7$, rendering the results insignificant.

To further investigate the convergence, we test the sensitivity of the mean to the exclusion of the most extreme outliers, that is the values furthest away from the expected mean. This is motivated by the tendency of the non-Gaussian model to create a few non-Gaussian maps with very prominent spots, causing extreme outliers. Thus, we expect that these simulations will dominate the mean of the distribution if the number of simulations is too low. The effect of excluding outliers on the mean, for the data-split yielding a low mean of $\mu_{NG} \approx 0.515$, are shown in figure 6.10, indicating that the mean of the non-Gaussian distribution without the spot masked is very sensitive to a few values. This again implies that the non-Gaussian distribution has not converged properly, while the rest of the simulations seems to be relatively robust to the removal of statistical outliers.

Based on this, we make a decision to simulate 1920 non-Gaussian only maps reusing the noise maps, as we only have the 4000 noise maps provided by the Planck FFP8.1 available. This should not pose an issue, as there is no reason to suspect any noise correlation. All the results in the following section have been obtained from 1280 Gaussian and 1920 non-Gaussian maps from the Planck FFP8.1 MC simulation data, with noise and masking.

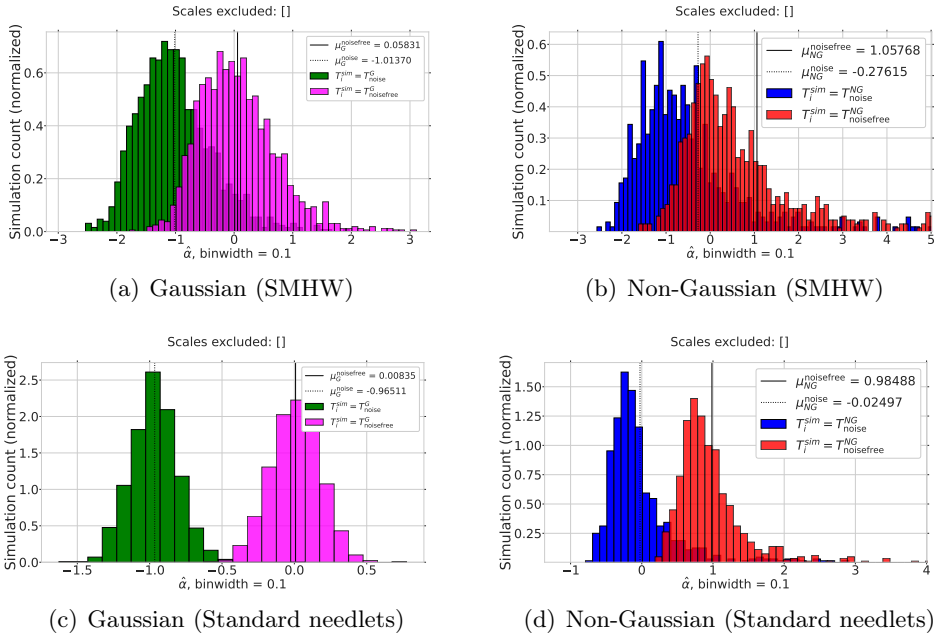


(a) Gaussian



(b) Non-Gaussian

Figure 6.7: The $\hat{\alpha}$-estimator for standard needlets, shown for both noisy and noise-free simulations, with the needlet scales $j \in \{13, 14, 15, 16\}$ excluded. In the non-Gaussian distributions, outliers up to $\hat{\alpha} = 20.54$ have been omitted from the plot.



(a) Gaussian (SMHW)

(b) Non-Gaussian (SMHW)

(c) Gaussian (Standard needlets)

(d) Non-Gaussian (Standard needlets)

Figure 6.8: The $\hat{\alpha}$-estimator for SMHW and standard needlets, shown for both noisy and noise-free simulations, with no scales excluded.

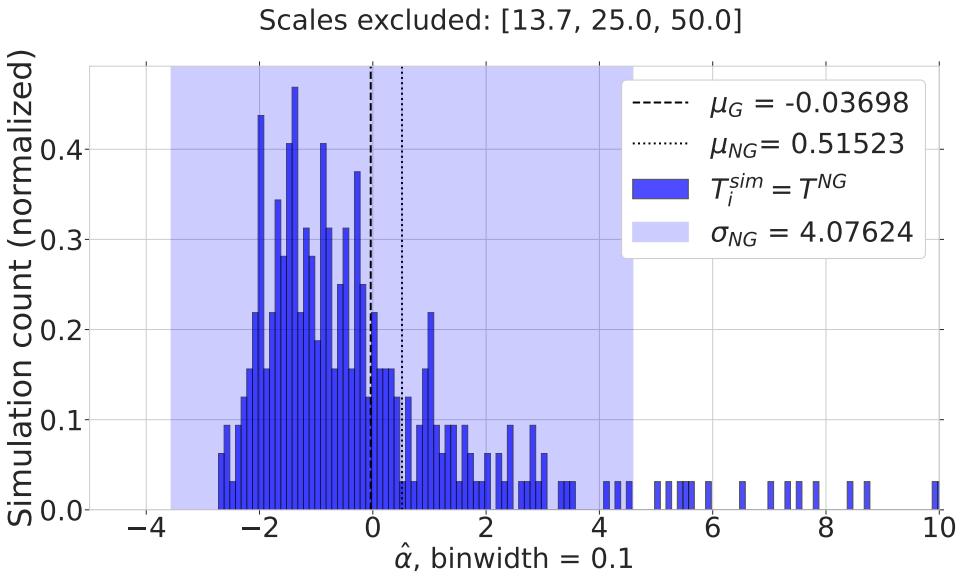
6.2 The simple trispectrum estimator

We present the results for the main two estimators, χ^2 and $\hat{\alpha}$, using the simple estimator for the trispectrum shown in section 5.1.

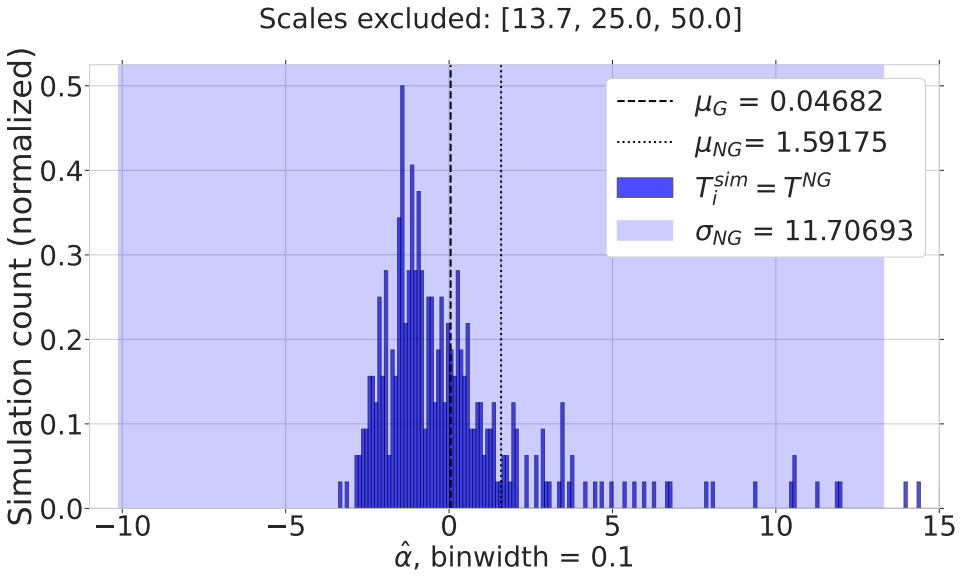
6.2.1 Results for χ^2

In figure 6.11, the χ^2 distribution from SMHW is shown. Regarding separation of the distributions, we have that $71.2\%, 17.1\%$ and 4.1% of the non-Gaussian distribution lies outside the $68.24\%, 95.45\%$ and 99.73% limits of the Gaussian distribution. The Planck data seems completely consistent with a Gaussian hypothesis, lying inside the 68.24% limit. More precisely, 32.08% of the Gaussian simulations achieved a higher χ^2 value than χ^2_{data}.

For the standard needlets, the χ^2 distribution is shown in figure 6.12. We have that $58.39\%, 17.52\%$ and 3.96% of the non-Gaussian distribution lies outside the $68.24\%, 95.45\%$ and 99.73% limits of the Gaussian distribution, not showing as good an ability as the SMHW to distinguish the Gaussian and non-Gaussian simulations. As the case for the SMHW, the needlets show a χ^2-value for the data fully consistent with the Gaussian null hypothesis.

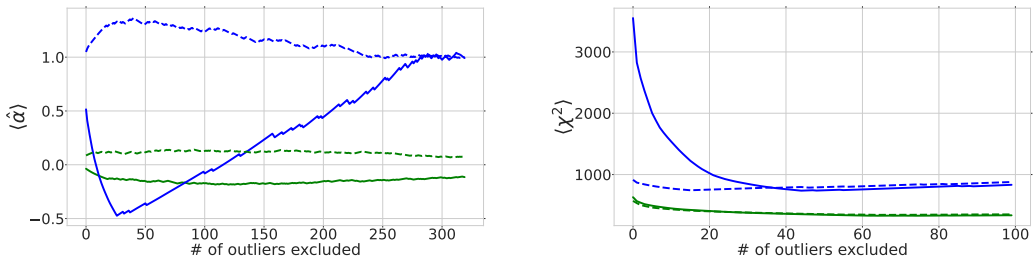


(a) First half for T^{sim}



(b) Latter half for T^{sim}

Figure 6.9: The $\hat{\alpha}$ estimator applied to the 640 non-Gaussian maps from the FFP8.1 Planck simulations with noise and mask. The mean, μ, is plotted in, with the subscripts G and NG referring to Gaussian and non-Gaussian simulations respectively, although only the non-Gaussian distribution is plotted. The blue band represents 1σ for the non-Gaussian distribution, that has an expected mean of 1. Outliers up to $\hat{\alpha} = 33.8$ and $\hat{\alpha} = 151.2$ have been cut in panel (a) and (b) respectively.



(a) Effect on $\hat{\alpha}$ (b) Effect on χ^2

Figure 6.10: The mean of χ^2 and $\hat{\alpha}$, plotted against number of outliers removed from a total of 640 Gaussian and 320 non-Gaussian simulations used for T^{sim}. The green color represents Gaussian simulations, while the blue represents the non-Gaussian. The dashed lines represent the Gaussian and non-Gaussian simulations with the spot masked, in green and blue respectively.

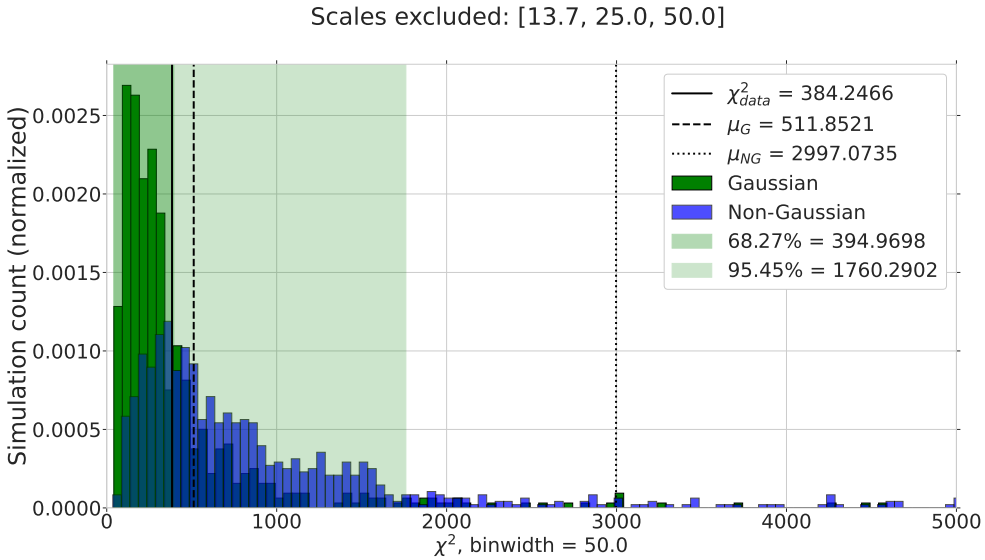


Figure 6.11: The SMHW χ^2 distribution for the simple trispectrum estimator. The histogram is plotted in the range $[0, 5000]$ for equispaced bins of width 50. The green shaded areas represent the interval containing 68.27% and 95.45% of the Gaussian distribution respectively. Outliers up to $\chi^2 = 161283.4$ are omitted for clarity. The distributions are obtained from 1280 Gaussian and 1920 non-Gaussian maps, with noise and masking.

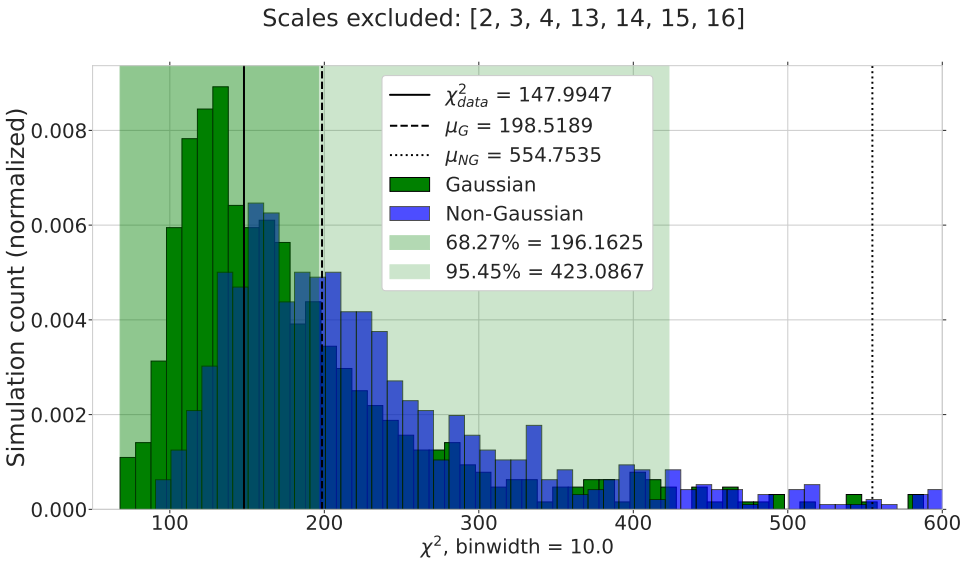


Figure 6.12: As in figure 6.11, only now with needlets. Plotted with equispaced bins of width 10, in the range $[50, 600]$. Outliers up to $\chi^2 = 58304.6$ have been omitted for visibility.

Masking the spot

The effect of masking the spot on the χ^2 distribution can be seen in figure 6.13 for the SMHW and figure 6.14 for the standard needlets. The percents lying outside the previously defined $68.24\%, 95.45\%$ and 99.73% limits are $64.2\%, 16.6\%$ and 4.9% for the SMHW and $58.81\%, 16.48\%$ and 3.96% for the standard needlets. For the SMHW, figure 6.13 shows that masking the spot worsens the ability to differentiate Gaussian and non-Gaussian simulations. A likely explanation is that the SMHW are excellent at detecting spots in the non-Gaussian maps, and thus a major part of "their detected non-Gaussian features" is removed. For the standard needlets, the effect of masking the spot is negligible, implying that the needlets are not very sensitive to features such as the spot. Masking the spot also has a large effect on the Planck data, χ^2_{data}, in the SMHW distribution. This can be understood as follows: if the data is already approximately Gaussian distributed, such a spot represents a tail outlier in the wavelet coefficient (temperature) distribution. Thus, removing a large part of the tail will skew the distribution, introducing a signal of non-Gaussian features.

6.2.2 Results for $\hat{\alpha}$

The initial results for $\hat{\alpha}$ for the SMHW is shown in figure 6.15, in the range of $\pm 5\sigma_G$ from the Gaussian mean. We can see that the data falls on the left side of the Gaussian mean, at $\hat{\alpha}_{\text{data}} = -0.77$, with 71.51% of the Gaussian simulations achieving a higher

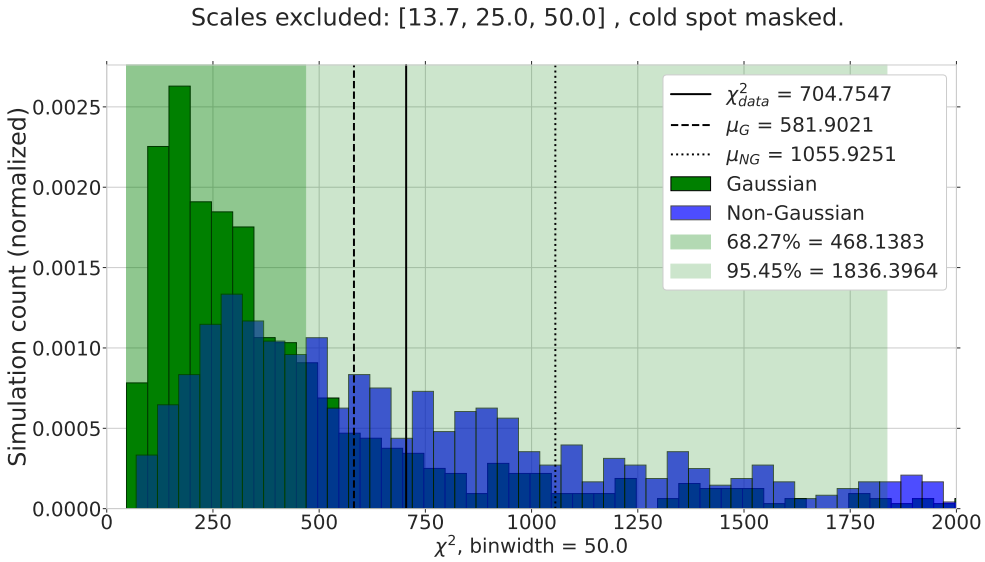


Figure 6.13: This figure shows the same as figure 6.11, only now with the spot masked. Plotted in the range $[0, 2000]$, with outliers up to $\chi^2 = 47374.6$ cropped.

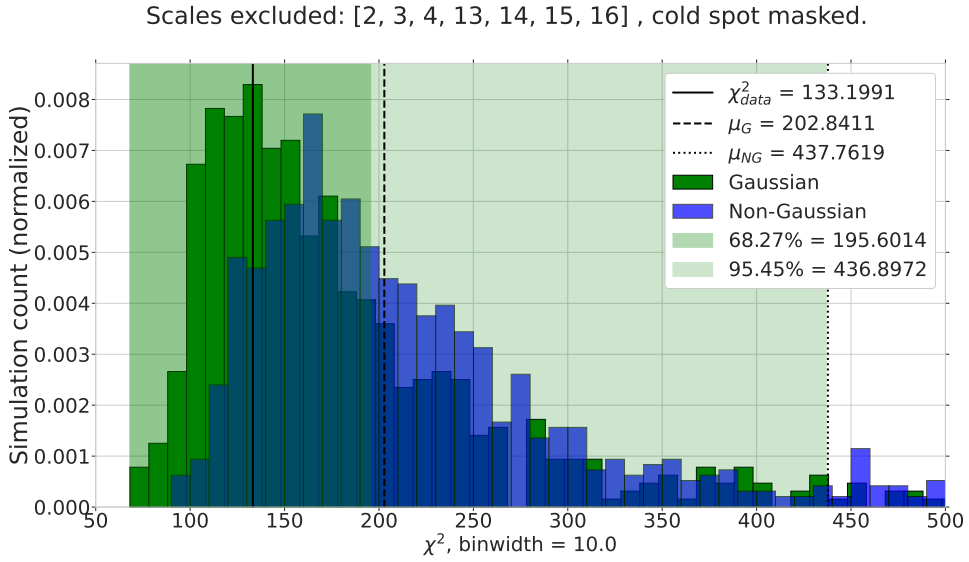


Figure 6.14: The same result as in figure 6.12, only now with the spot masked, in the range $[50, 500]$. Outliers up to $\chi^2 = 133233.5$ have been cropped.

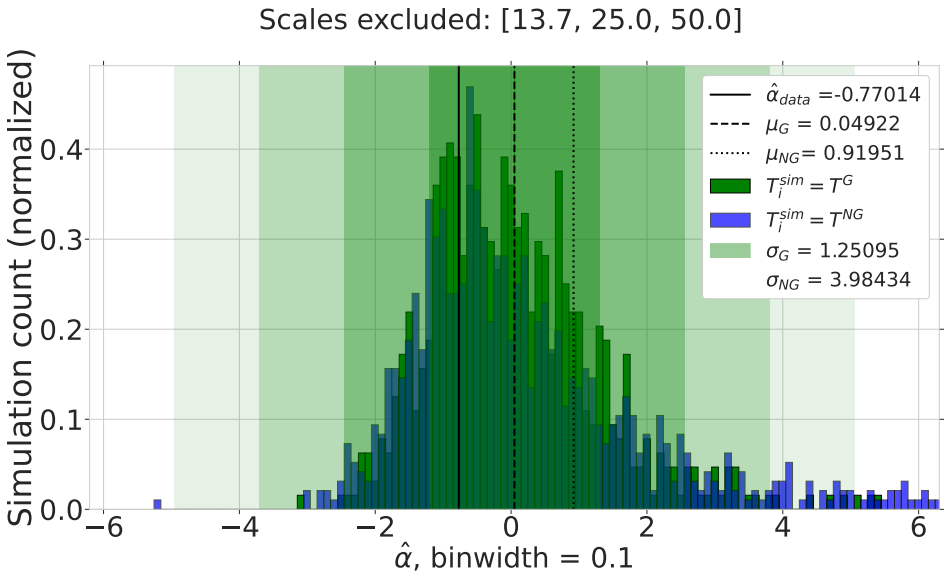


Figure 6.15: The initial results for $\hat{\alpha}$ for SMHW with the simple trispectrum estimator, for equispaced bins of width 0.1, in the range $\pm 5\sigma_G$. The shaded green areas represent 1, 2, 3 and 4σ respectively. The standard deviation for the non-Gaussian distribution is not plotted, but shown in the legend. Non-Gaussian outliers up to $\hat{\alpha} = 40.2$ are omitted. The distributions are obtained from 1280 Gaussian and 1920 non-Gaussian maps, with noise and masking.

value for $\hat{\alpha}$. However, we must be careful in our interpretation of the negative value for $\hat{\alpha}_{\text{data}}$. This does not necessarily tell us that the data prefers the Gaussian hypothesis, but that it prefers a negative value for the non-Gaussian strength parameter γ shown in section 2.6. Anyhow, the Gaussian and non-Gaussian distributions have large standard deviations at $\sigma_G = 1.25$ and $\sigma_{NG} = 3.98$ respectively, yielding results of low significance, as $\hat{\alpha}_{\text{data}}$ falls within 1σ of both distributions. This is what ultimately led us to search for a more robust estimator for the trispectrum.

For the standard needlets, the distribution of $\hat{\alpha}$ is shown in figure 6.16. This figure tells much of the same story as for the SMHW, with the data falling within 1σ of both distributions, albeit with much smaller standard deviation for the non-Gaussian distribution. This indicates that there are fewer dominating outliers detected by the needlets in the non-Gaussian distribution. 19.72% of the Gaussian simulations achieved a higher value of $\hat{\alpha}$ than the data. For both cases, SMHW and standard needlets, we still seem to have decent agreement with the expected means of the two distributions, namely $\langle \mu_G \rangle = 0$ and $\langle \mu_{NG} \rangle = 1$, even with noise and masking.

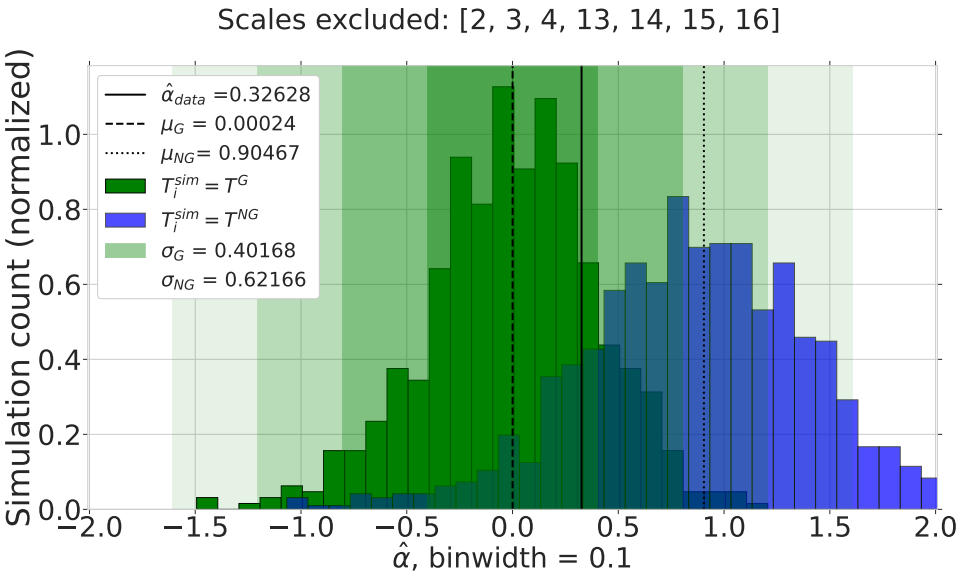


Figure 6.16: As in figure 6.15, only now with needlets. Outliers up to $\hat{\alpha} = 5.5$ are not shown.

Masking the spot

The effect of masking the spot on $\hat{\alpha}$ is shown in figure 6.17 for the SMHW maps. We can clearly see that masking a region in the sky corresponding to a tail end of a distribution has a strong effect on the data, probably due to skewing the distribution as discussed in the section on χ^2 results. The data is now even less aligned with the Gaussian mean than our non-Gaussian model distribution is. The same situation, only for needlets, are shown in figure 6.18. The effect is much smaller for the needlets than the SMHW, again most likely because they are worse at detecting features such as the spot.

6.3 The optimal trispectrum estimator

In this section, we present the χ^2 and $\hat{\alpha}$ distributions, evaluated on the Planck FFP8.1 MC simulations with the optimal trispectrum estimator presented in section 5.1.1. We start by presenting the SMHW results.

6.3.1 A possible detection in the SMHW maps

We report the results for the SMHW maps, with the optimal trispectrum estimator. The χ^2 test is plotted in figure 6.19, where a strong disfavoring of Gaussianity is seen in the data. Only 4 out of 639 total Gaussian simulations, $\sim 0.63\%$, achieves a similarly extreme result. The separation between the histograms is still relatively

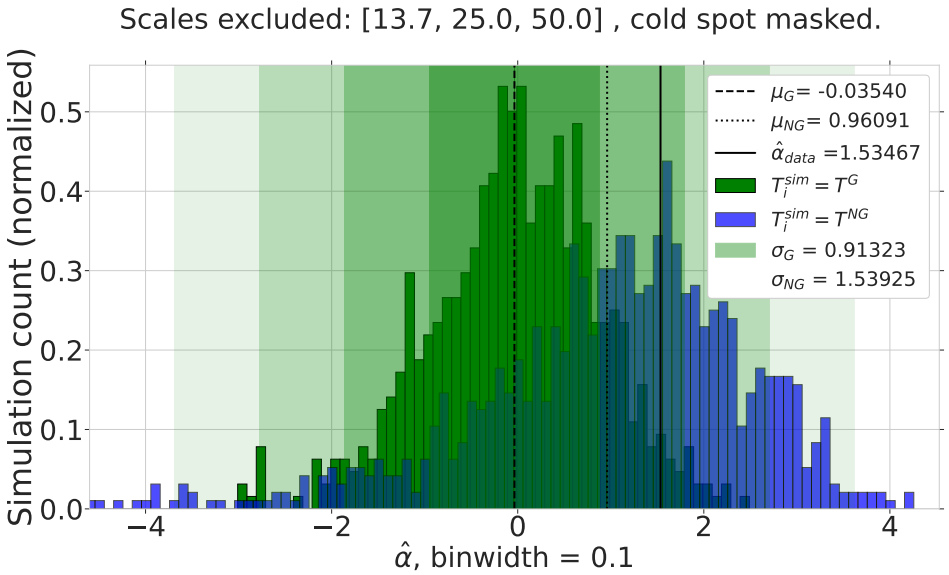


Figure 6.17: Shows $\hat{\alpha}$ as in figure 6.15, but with masking the spot. Non-Gaussian outliers are cropped on the lower end, down to $\hat{\alpha} = -9.0$

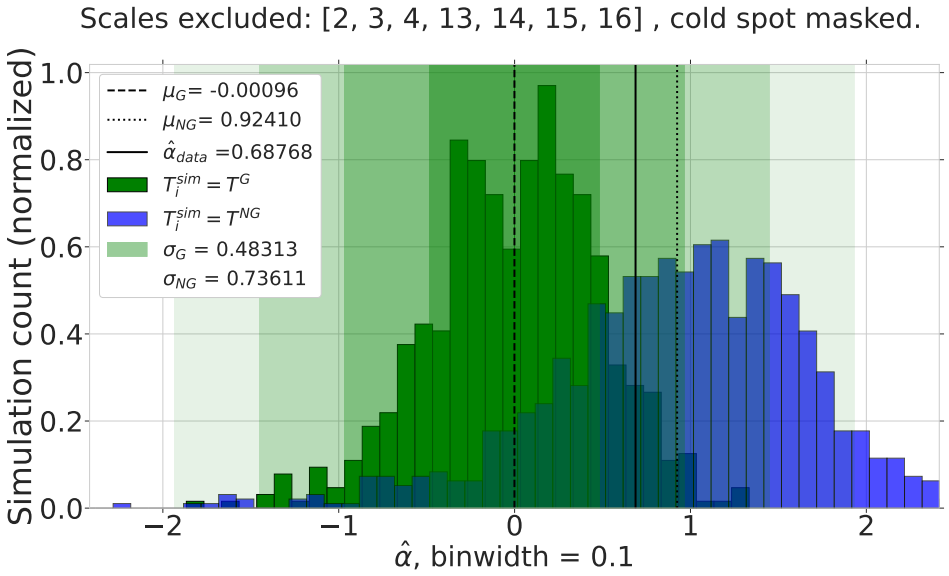


Figure 6.18: As in figure 6.16, but showing the effect of masking the spot. Outliers up to $\hat{\alpha} = 2.8$ are not shown.

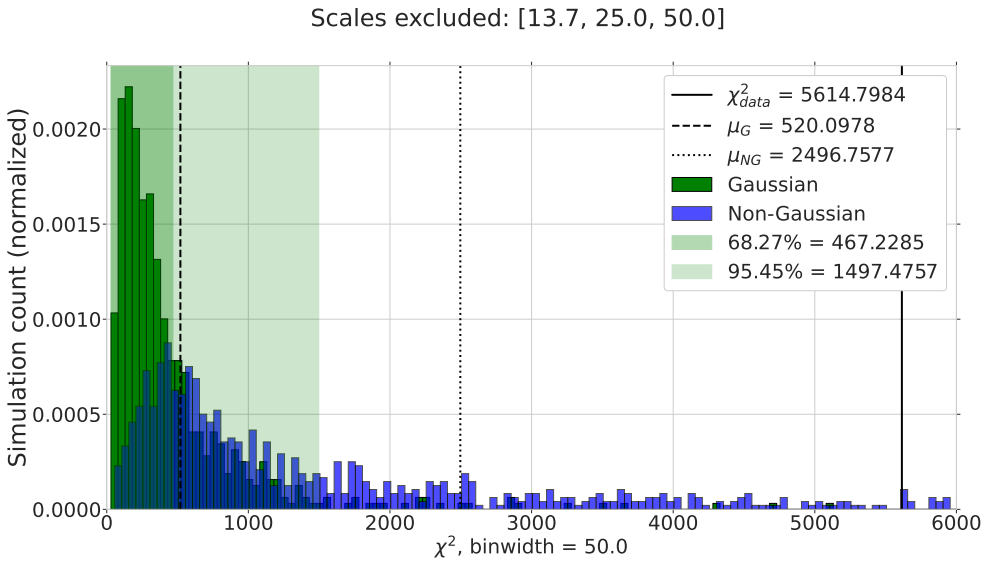


Figure 6.19: The χ^2 for SMHW, with the optimal trispectrum estimator. Plotted in the range $[0, 6000]$, with outliers up to $\chi^2 = 56487.7$ omitted. The distributions are obtained from 1280 Gaussian and 1920 non-Gaussian maps, with noise and masking.

good, with $76.96\%, 38.89\%$ and 7.30% of the non-Gaussian simulations lying outside the $68.24\%, 95.45\%$ and 99.73% limits of the Gaussian simulations.

The results for $\hat{\alpha}$, with the SMHW, are shown in figure 6.20. We can see that the data strongly disfavors the Gaussian hypothesis, seemingly at a $\sim 3.35\sigma$ level. However, after testing the tail of the distribution, we find that 4 out of 639 Gaussian simulations achieve a similarly extreme $\hat{\alpha}$, indicating that the $\hat{\alpha}$-distribution is not fully Gaussian, but slightly skewed. This lowers the significance of the detection to 0.63%. The data prefers a non-Gaussian strength parameter ~ 1.56 times as large as our current γ at 0.43, that is $\gamma = 0.67$. It is also worth noting that the optimal trispectrum estimator lowers the standard deviation of the SMHW $\hat{\alpha}$ distribution to $\sigma_G \approx 0.46$ and $\sigma_{NG} \approx 1.02$ for the Gaussian and non-Gaussian distributions respectively, making this result much more significant.

Both the results for χ^2, and the results for $\hat{\alpha}$ are very robust to masking the spot, shown in figure 6.21 and 6.22 respectively. The separation in the χ^2 distribution remains about the same, with $77.16\%, 36.29\%$ and 6.78% of the non-Gaussian simulations lying outside the $68.24\%, 95.45\%$ and 99.73% Gaussian limits. Exactly the same number of Gaussian simulations, 4 out of 639, achieves a similarly extreme result as χ^2_{data} with the cold spot masked. The results for $\hat{\alpha}$ tells much of the same story, with the data disfavoring the Gaussian hypothesis. The significance of the detection is marginally decreased as the standard deviation for the Gaussian $\hat{\alpha}$ is increased to $\sigma_G \approx 0.55$.

While we deemed the exclusion of the three smallest SMHW scales and the 4 largest

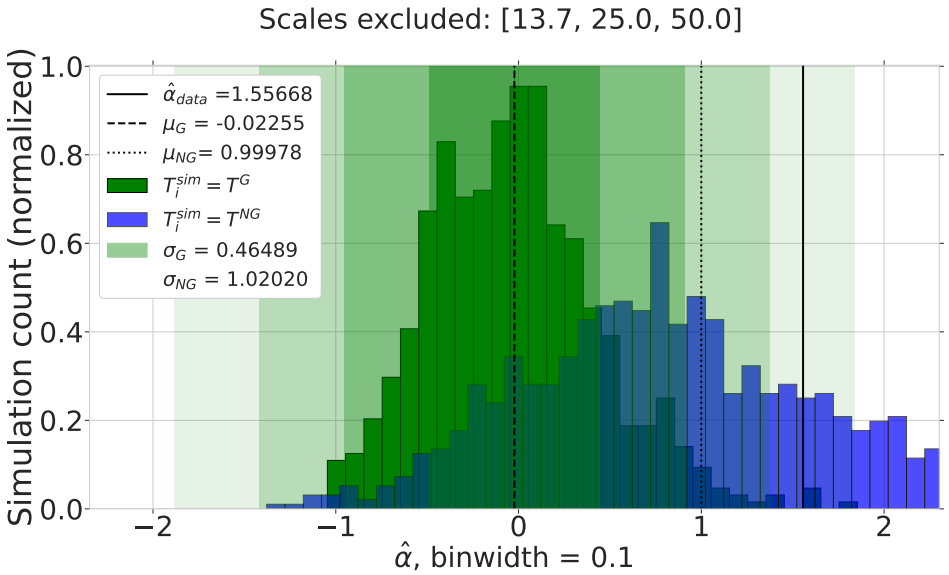


Figure 6.20: The distribution for $\hat{\alpha}$, for SMHW with the optimal trispectrum estimator. Outliers up to $\hat{\alpha} = 5.6$ are cut.

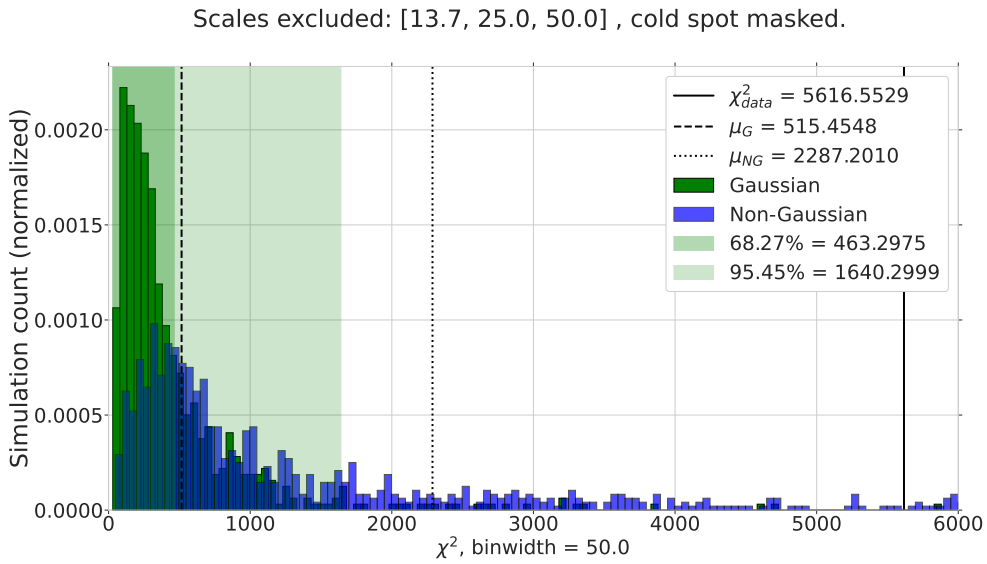


Figure 6.21: As in figure 6.19, but with the spot masked. Outliers up to $\chi^2 = 59908.5$ have been omitted.

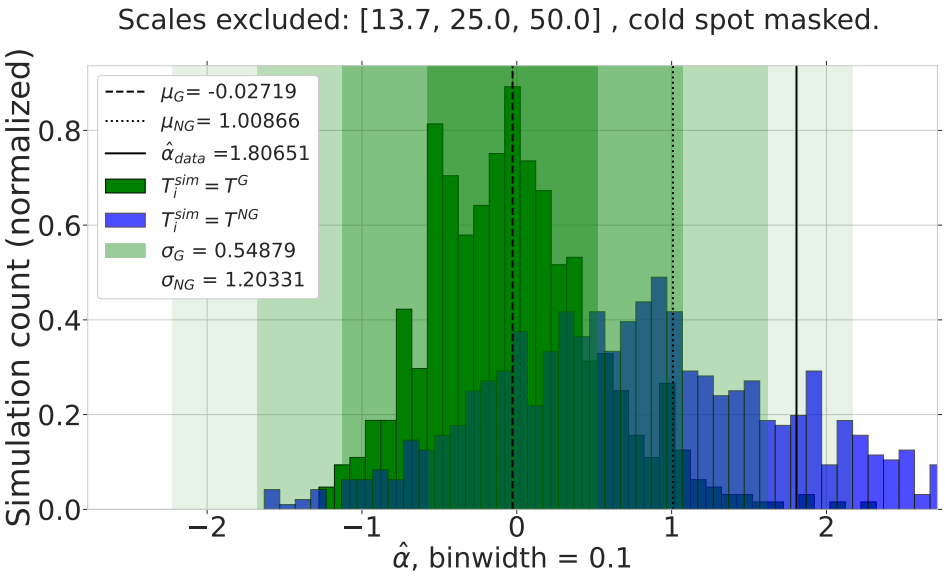


Figure 6.22: As figure 6.20, but with the spot masked. Outliers up to $\hat{\alpha} = 6.5$ are cut.

needlet scales the best fit in case of a large discrepancy in the noise modeling, we can re-investigate the effects on the actual data. The effect of including more scale combinations on $\hat{\alpha}$ is shown in figure 6.23. Including more scales does not drastically change the result, however the data shows a less extreme result the more scales you include, moving towards the Gaussian (and non-Gaussian) mean. One must take care when interpreting the results of including more scales, as this could introduce a noise bias in the result, which was the original reason for their exclusion. If there is a noise bias in these scales, the trend is an underestimation of the trispectrum, as the foremost effect is the lowering of $\hat{\alpha}_{\text{data}}$. A "more traditional" histogram plot is provided in appendix B.3 for the re-inclusion of scales.

6.3.2 The standard needlets

The χ^2 test is plotted in figure 6.24, where we can see that we achieve a good separation for the bulk of the distribution, with $87.59\%, 33.16\%$ and 1.25% of the non-Gaussian histogram lying outside the $68.24\%, 95.45\%$ and 99.73% limits of the Gaussian one. The data is more aligned with the bulk of the Gaussian distribution, lying inside the 68.24% limit, with 82.15% of the Gaussian simulations achieving a higher χ^2 value. The same consistency with Gaussianity can be seen from the distribution of $\hat{\alpha}$, shown in figure 6.25, with the data lying within 1σ of the Gaussian distribution, and barely outside the 1σ-limit for the non-Gaussian distribution.

The effect of masking the spot on χ^2 is shown in figure 6.26. The effect seems to be completely negligible, marginally changing the mean values. For the χ^2 distribution, we have we have very similar separation as with the spot included, with $87.70\%, 34.20\%$

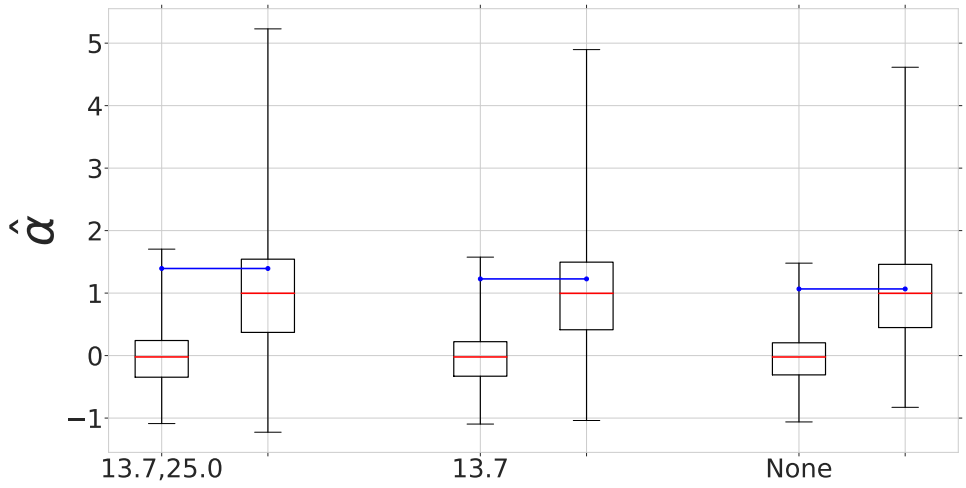


Figure 6.23: $\hat{\alpha}$ shown for inclusion of the previously omitted scales as a boxplot, to better illustrate the means of the distributions. The x-axis denotes which scales, R, are excluded. The left box in each group is the Gaussian distribution, and the right is the non-Gaussian. The box contains 50% of the distribution about the median. The red lines are the mean of the respective distribution, while the blue lines with dots are the Planck values $\hat{\alpha}_{\text{data}}$. The whiskers show the full range of the distributions.

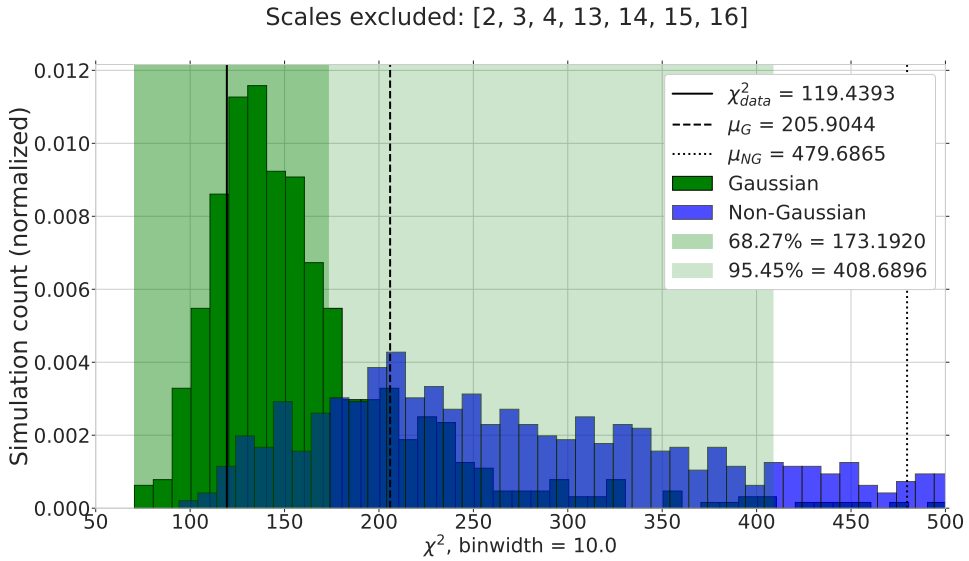


Figure 6.24: The χ^2 distribution for standard needlets, with the optimal trispectrum estimator. Plotted in the range $[50, 500]$, with outliers up to $\chi^2 = 10183.1$ cut. The results are obtained from 1280 Gaussian and 1920 non-Gaussian simulations with mask and noise.

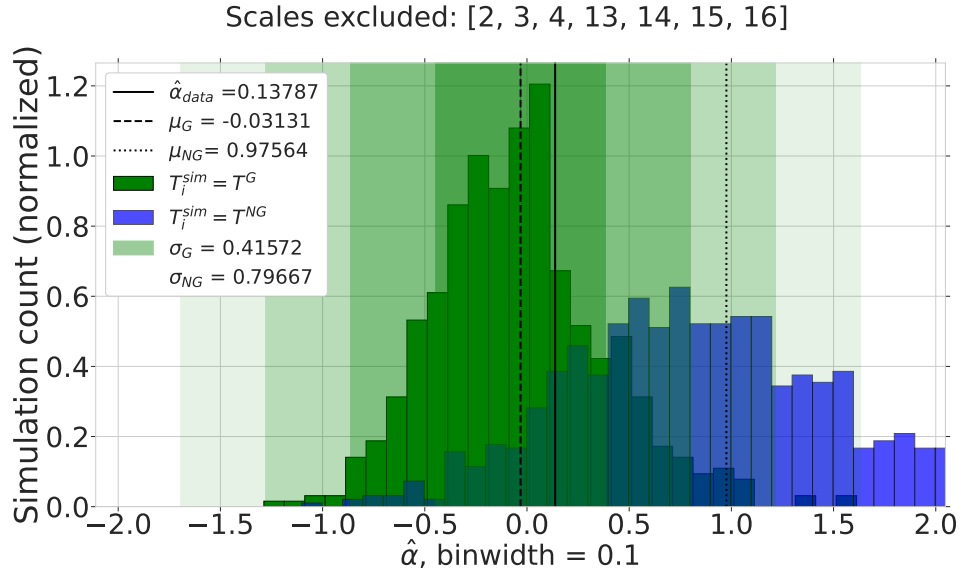


Figure 6.25: The $\hat{\alpha}$ distribution for standard needlets, with the optimal trispectrum estimator. Plotted in the range $\pm 5\sigma_G$. Outliers up to $\hat{\alpha} = 4.0$ are cut. The results are obtained from 1280 Gaussian and 1920 non-Gaussian simulations with mask and noise.

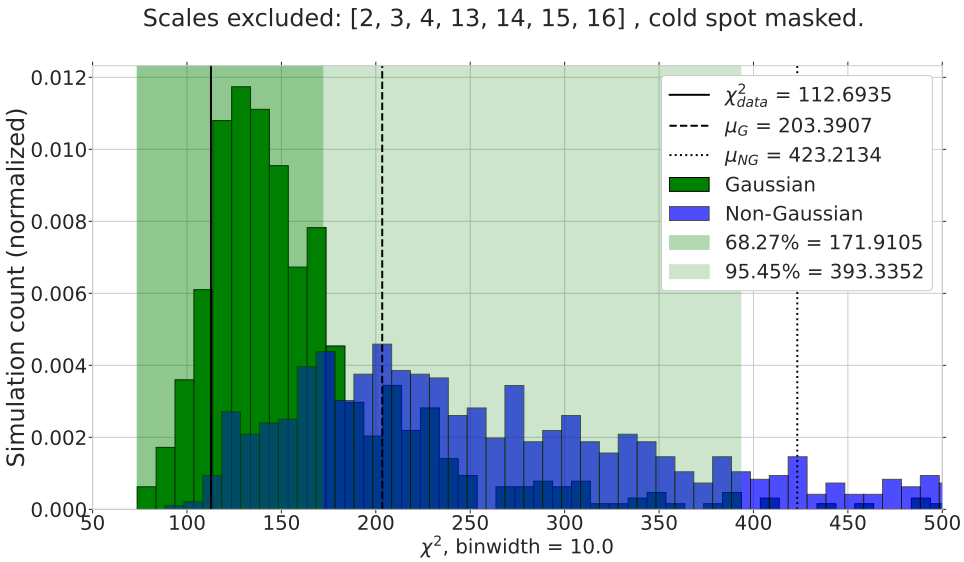


Figure 6.26: As in figure 6.24, only now with the spot masked. Plotted in the range $[50, 500]$. Outliers up to $\chi^2 = 9918.2$ are cut.

and 1.25% of the non-Gaussian histogram lying outside the 68.24%, 95.45% and 99.73% limits.

Especially interesting is the amount of non-Gaussian outliers detected outside the 99.73% limit, as a detection at this level is where the results start to be "interestingly significant". For the needlets, only 1.25% of the non-Gaussian simulations lie outside this limit, both with the spot included and masked. For the SMHW, the equivalent percentages are 7.30% with the spot included, and 6.78% with the spot masked. This indicates that the SMHW are much more apt at detecting the most extreme non-Gaussian outliers. This is a possible explanation for why a detection is seen in the SMHW maps, but not in the standard needlet maps.

The distribution of $\hat{\alpha}$ with the cold spot masked is shown in figure 6.27. This tells much of the same story as with the spot included, that the data is consistent with Gaussianity, lying within the 1σ band of the Gaussian distribution, preferring a contribution from the non Gaussian term αT^{NG} consistent with zero. The effect of including previously omitted scales is shown in figure 6.28 for the standard needlets, where no major effect is seen. The results seem relatively robust to the inclusion of the possibly "wrongly-noisy" scales, only marginally changing the means of the distributions, and slightly lowering $\hat{\alpha}_{\text{data}}$. Still, the data remains within 1σ of the Gaussian distribution for all cases. Again, a traditional histogram plot of the inclusion of more scales is avaiable in appendix B.3.

We conclude that the optimal standard needlet trispectrum estimator shows no preference for the non-Gaussian model, and is consistent with the Gaussian null hypothesis.

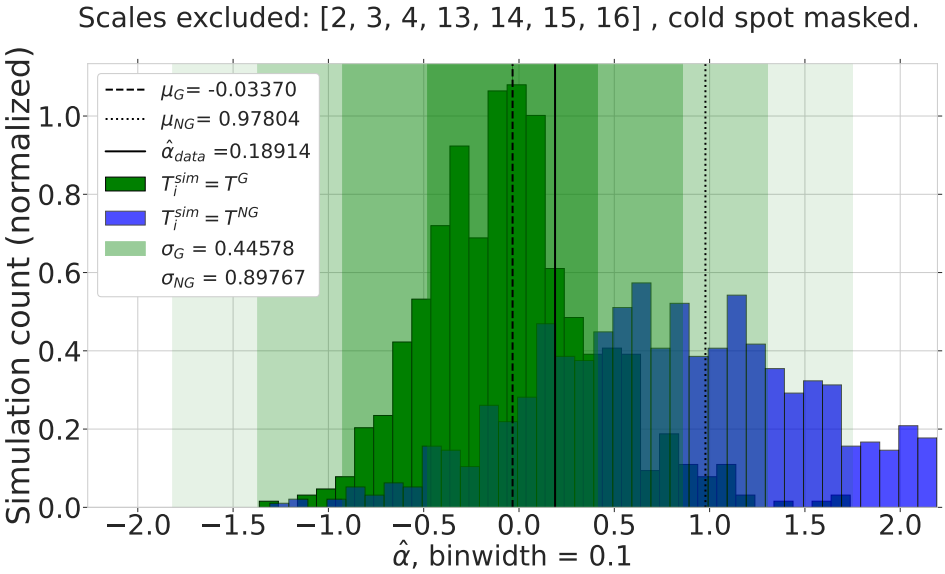


Figure 6.27: As in figure 6.25, only now with the spot masked, shown in the range $\pm 5\sigma_G$. Outliers up to $\hat{\alpha} = 4.7$ are cut.

6.3.3 Further investigations of the detection

To investigate the localization of the detection in harmonic space, we implement the test

$$\frac{T^{\text{Planck}}_{j_1 j_2 j_3 j_4} - \langle T^{\text{sim}}_{j_1 j_2 j_3 j_4} \rangle}{\sigma_{j_1 j_2 j_3 j_4}}, \tag{6.8}$$

where $T_{j_1 j_2 j_3 j_4}$ is the trispctrum and $\sigma_{j_1 j_2 j_3 j_4}$ is the standard deviation of the simulated trispectrum. In this way, we can test if the anomaly is localized to certain scale combinations. The result is shown in figure 6.29, telling us that this is systematic, not localized to a few scales. Unfortunately, this is the extent to which we were able to test the detection in the allotted time frame. Some steps that can be taken for further investigation will be presented in the concluding remarks.

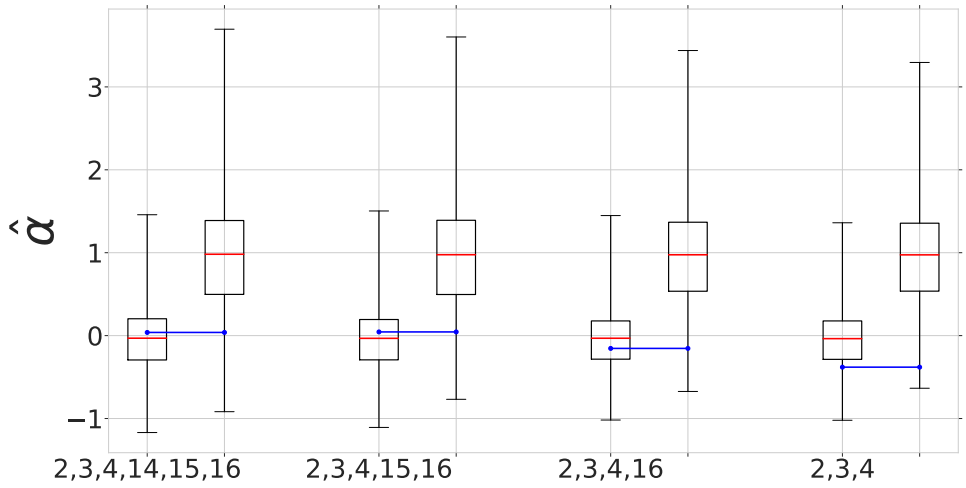
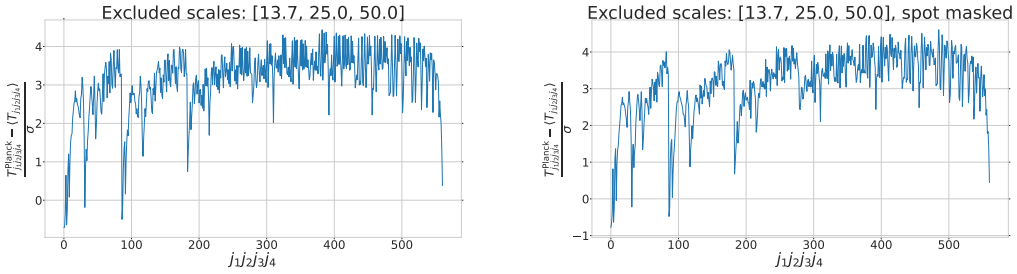


Figure 6.28: $\hat{\alpha}$ as in figure 6.23, only now for standard needlets. As before, the x-axis represents how many scales are excluded.



(a) With the spot

(b) Spot masked

Figure 6.29: Difference plot of the Planck trispectrum and the simulation mean trispectrum, normalized by the standard deviation. Shown with cold spot included in panel (a) and the cold spot masked in panel (b)

Chapter 7

Conclusion and future prospects

In this thesis, we have presented the theoretical framework and the data pipeline in which the CMB is analyzed, presenting concepts such as spherical decomposition of the temperature field, foreground removal, and noise handling. We have also discussed the use of wavelets and polyspectra in statistical tests, and why wavelets provide an excellent alternative to the traditional spherical harmonics in the presence of incomplete sky coverage.

We have focused on six anomalies reported in the CMB, namely the large- and small-scale hemispherical asymmetry, The Cold Spot, the large-scale power deficit, the alignment of the quadrupole and octopole, and the parity asymmetry on large angular scales. Evidence and methods for detection have been presented for all six anomalies.

Furthermore, we have presented the theory for and investigated the statistical properties of a model presented by Hansen et al. [1], reproducing the aforementioned anomalies. The model is a phenomenological model with connections to g_{NL}-like terms in the primordial gravitational potential, producing strong scale-dependent non-Gaussianity in CMB maps. The statistical properties of the model have been explored within the framework of Spherical Mexican Hat Wavelets (SMHW), and spherical standard needlets.

We have also implemented two different estimators for the wavelet 4-point correlation function, the trispectrum, the first of which was validated on 3200 Gaussian simulations. The choice to explore the trispectrum was motivated by the connection to the g_{NL}-terms arising from the model.

We proposed and investigated a blind non-Gaussianity test, using a χ^2 statistic on the estimated trispectrum, testing the Gaussian null hypothesis. The χ^2 distribution quantifies how well the data fits with the assumption of Gaussianity.

We also presented and implemented a model-specific test, an estimator $\hat{\alpha}$ which quantified the preferred contribution from the non-Gaussian term arising in the theoretical model, that is for a result of $\hat{\alpha} = 1$ the data prefers exactly our model, and a result of $\hat{\alpha} = 0$ prefers no contribution at all from the non-Gaussian term in our model. Both tests were validated on 3200 Gaussian simulations, pictured in figure 6.2, 6.4, and 6.5.

As a logical next step, a comparison with the Planck data was made using the 2015 Planck FFP8.1 Monte Carlo simulations. We report the results from the optimal trispectrum estimator. For the spherical standard needlets, the results were consistent with the Gaussian null hypothesis, presented in figure 6.24 and 6.25. However, a very interesting detection was made in the SMHW framework, disfavoring the Gaussian null hypothesis at 0.63% significance, shown in figure 6.19 and 6.20. We argued that the reason a detection was seen in the SMHW distributions and not for the standard needlets is that the SMHW proved themselves much better at detecting non-Gaussian statistical outliers. Unfortunately, due to time and computational constraints, we were not able to fully test these results, and the detection warrants further investigation.

With such an "atypical result", it is important to fully scrutinize the result through a host of tests. I will mention some of the most pressing tests here, which would have been the next step in investigating the result.

The first step is to fully validate the optimal estimator used. The choice to search for a more robust estimator was made late in the process of this thesis, and thus we had just enough time to apply it to the Planck data. Additionally, the simple problem of convergence is still present. Optimally, we would test all results with a simulation number on the order of 10^4 or higher, not 10^3 as is the case for this thesis. For the Planck data, this requires some tinkering, as there is no simulation data available of this magnitude.

A possible source of error is the method of foreground subtraction. In this thesis, we used the 4000 SMICA FFP maps. An immediate test should be to see if the result is robust to changing the foreground subtraction method to the other three methods NILC, SEVEM, and Commander. Unfortunately, there are only 1000 available maps for each of these methods, which is on the low end for proper statistics. Still, it would provide us with a reasonable sanity check.

When using half-mission maps, one must make a choice of what combination to use when estimating the trispectrum, i.e. for half-mission maps with superscript 1 and 2, $\beta^1_{j_1}\beta^1_{j_2}\beta^2_{j_4}\beta^2_{j_4}$ as opposed to for example $\beta^1_{j_1}\beta^2_{j_2}\beta^1_{j_4}\beta^2_{j_4}$, where we used the former in this thesis. By changing the ordering of the half-mission maps, we can introduce differences in the noise error, to check whether this significantly changes the result. However, as we took a conservative approach and excluded the noisiest scales, this effect is assumed to be small.

Another interesting test is to investigate whether the anomalous trispectrum is localized on the sky. One could, for example, estimate trispectra in different hemispheres, as a consistency check to see if the effect is a global one, or if the anomalies are localized as expected in the model.

Furthermore, as the χ^2 distribution is highly asymmetric, we can not use the standard deviation as a measure of significance. Therefore, to quantify just how extreme the χ^2 result for the data is, we would need to simulate a large number of Gaussian simulations as a ground for comparison. As we do not have a large number of noise maps available, we would first need to check if the χ^2 distributions for Gaussian simulations with a mask but without noise are a good fit to the simulations with both noise and mask. If they are, we can then run Gaussian simulations without noise which would be

comparable to the noisy simulations.

An additional test is to check whether errors are introduced in the masking process. One could extend the wavelet masks further, to check for anomalous boundary effects due to insufficient extension of the wavelet masks.

Finally, it would be reasonable to test the results against the newest 2018 Planck data. However, as there is a low number of available Monte Carlo simulations from the latest Planck data release, this would be on the low end of what we could consider proper statistics. Still, it would be a welcome sanity check.

Appendices

Appendix A

A.1 Wigner-3j and Gaunt integrals

The Wigner-3j symbols are a handy tool when working with spherical harmonics. They relate to the integral of the product of three spherical harmonics by

$$\begin{aligned}&\int Y_{l_1 m_1}(\theta,\phi) Y_{l_2 m_2}(\theta,\phi) Y_{l_3 m_3}(\theta,\phi) sin\theta d\theta d\phi \\ &= \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}.\end{aligned} \tag{A.1}$$

The Wigner-3j symbols are zero unless the following selection rules are ALL satisified:

$$m_i \in \{-l_i, l_i+1, ..., l_i\} \tag{A.2}$$

$$m_1 + m_2 + m_3 = 0 \tag{A.3}$$

$$|l_1 - l_2| \leq l_3 \leq l_1 + l_2 \tag{A.4}$$

$$(l_1 + l_2 + l_3) \text{ is an integer.} \tag{A.5}$$

Even permutations of the columns in the Wigner-3j leaves the symbol invariant,

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} l_2 & l_3 & l_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = \begin{pmatrix} l_3 & l_1 & l_2 \\ m_3 & m_1 & m_2 \end{pmatrix}, \tag{A.6}$$

while odd permutations adds a phase factor

$$\begin{aligned}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} &= (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_2 & l_1 & l_3 \\ m_2 & m_1 & m_3 \end{pmatrix} \\ &= (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_1 & l_3 & l_2 \\ m_1 & m_3 & m_2 \end{pmatrix} \\ &= (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_3 & l_2 & l_1 \\ m_3 & m_2 & m_1 \end{pmatrix}\end{aligned} \tag{A.7}$$

A similar phase factor occurs for a switch in signs in m:

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{l_1+l_2+l_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} \tag{A.8}$$

Another useful equation is the relation to Legendre polynomials, $P(x)$,

$$\frac{1}{2}\int_{-}^{1} P_{l_1}(x)P_{l_2}(x)P_l(x)dx = \begin{pmatrix} l & l_1 & l_2 \\ 0 & 0 & 0 \end{pmatrix}^2 . \tag{A.9}$$

A.2 Derivation of the correlation between needlet coefficients

$$\Gamma_{j_1 j_2} = \langle \beta_{j_1,k}\beta_{j_2,k} \rangle$$

$$= \left\langle \sum_{l=B^{j_1-1}}^{B^{j_1+1}} \sum_{m=-l}^{l} b\left(\frac{l}{B^{j_1}}\right) a_{lm} Y_{lm}(\xi_{j_1,k}) \sum_{l'=B^{j_2-1}}^{B^{j_2+1}} \sum_{m'=-l'}^{l'} b\left(\frac{l'}{B^{j_2}}\right) a_{l'm'} Y_{l'm'}(\xi_{j_2,k}) \right\rangle$$

$\boxed{a_{l-m} = (-1)^m a^*_{lm} \quad , \quad Y^*_{lm}(\theta,\phi) = (-1)^m Y_{l-m}(\theta,\phi)}$

$$= \left\langle \sum_{l=B^{j_1-1}}^{B^{j_1+1}} b\left(\frac{l}{B^{j_1}}\right) \sum_{l'=B^{j_2-1}}^{B^{j_2+1}} b\left(\frac{l'}{B^{j_2}}\right) \sum_{m=-l}^{l} \sum_{m'=-l'}^{l'} a_{lm}(-1)^{-m'} a^*_{l'-m'} Y_{lm}(\xi_{j_1,k})(-1)^{-m'} Y^*_{l'-m'}(\xi_{j_2,k}) \right\rangle$$

$\boxed{\text{Sum is symmetric over } m \in \{-l, -l+1, ..., 0, ..., l-1, l\}, \text{ i.e. } \sum_{m=-l}^{l} Y_{lm} = \sum_{m=-l}^{l} Y_{l-m}}$

$$= \sum_{l=B^{j_1-1}}^{B^{j_1+1}} b\left(\frac{l}{B^{j_1}}\right) \sum_{l'=B^{j_2-1}}^{B^{j_2+1}} b\left(\frac{l'}{B^{j_2}}\right) \sum_{m=-l}^{l} \sum_{m'=-l'}^{l'} \langle a_{lm} a^*_{l'm'} \rangle\, Y_{lm}(\xi_{j_1,k}) Y^*_{l'm'}(\xi_{j_2,k})$$

$$= \sum_{l=B^{j_1-1}}^{B^{j_1+1}} b\left(\frac{l}{B^{j_1}}\right) \sum_{l'=B^{j_2-1}}^{B^{j_2+1}} b\left(\frac{l'}{B^{j_2}}\right) C_l \delta^{l'}_{l} \delta^{m'}_{m} \sum_{m=-l}^{l} \sum_{m'=-l'}^{l'} Y_{lm}(\xi_{j_1,k}) Y^*_{l'm'}(\xi_{j_2,k})$$

$\boxed{\text{Because of the } \delta^{l'}_{l}\delta^{m'}_{m}, \text{ only } l = l', m = m' \text{ survives.}}$

$$= \sum_{\text{shared l's}} b\left(\frac{l}{B^{j_1}}\right) b\left(\frac{l}{B^{j_2}}\right) C_l \frac{2l+1}{4\pi} P_l(cos\Delta\theta)$$

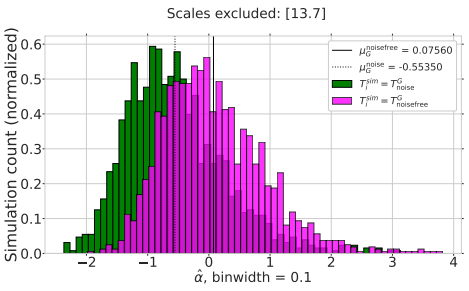
$\boxed{\Delta\theta = 0 \text{ for same k, giving } P_l(1) = 1}$

$$= \sum_{\text{shared l's}} b\left(\frac{l}{B^{j_1}}\right) b\left(\frac{l}{B^{j_2}}\right) \frac{2l+1}{4\pi} C_l$$

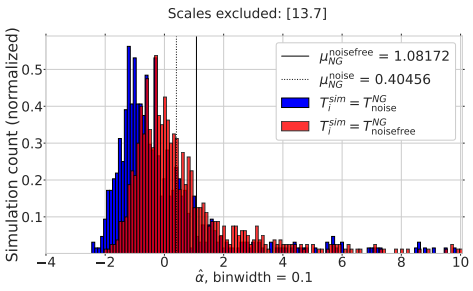
Appendix B

Noise comparisons for $\hat{\alpha}$

B.1 SMHW

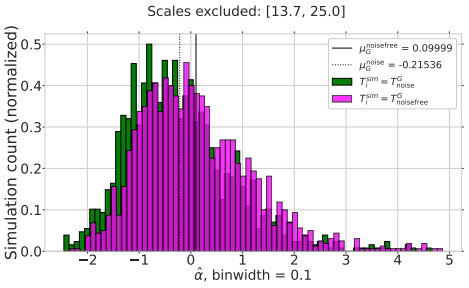


(a) Gaussian

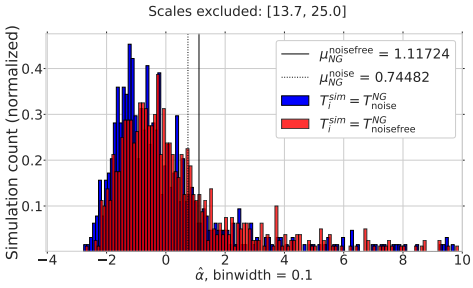


(b) Non-Gaussian

Figure B.1: Scales excluded: 13.7

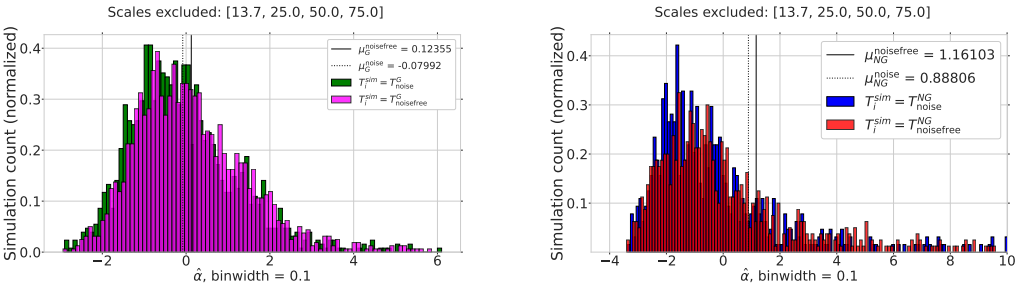


(a) Gaussian



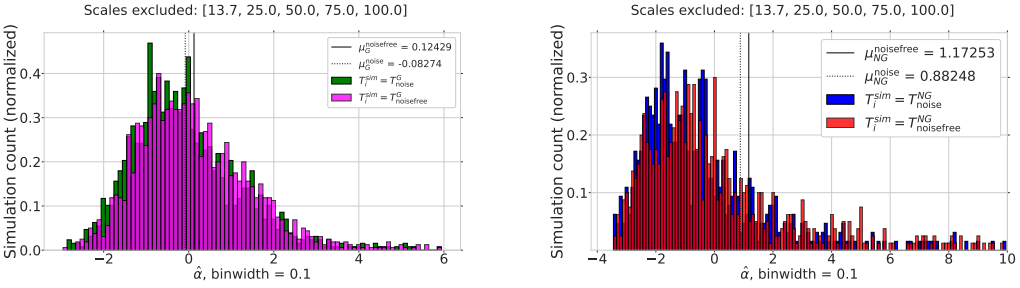
(b) Non-Gaussian

Figure B.2: Scales excluded: $[13.7, 25.0]$



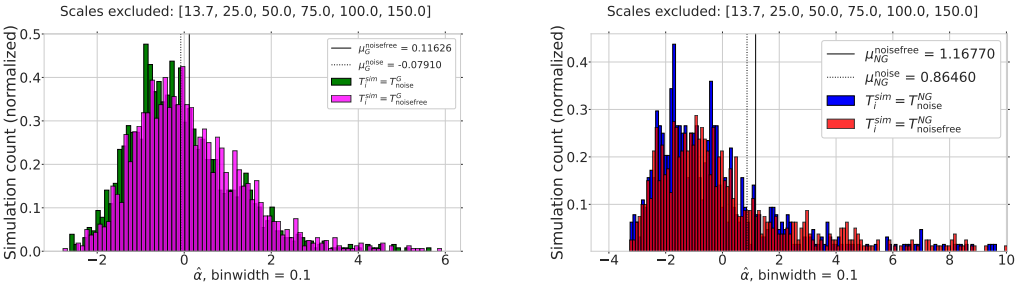
(a) Gaussian (b) Non-Gaussian

Figure B.3: Scales excluded: $[13.7, 25.0, 50.0, 75.0]$



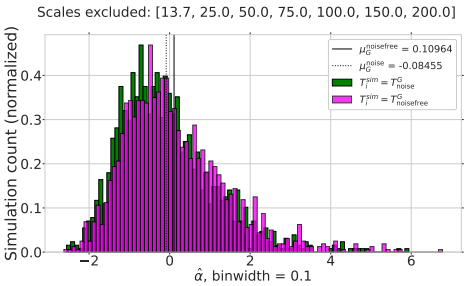
(a) Gaussian (b) Non-Gaussian

Figure B.4: Scales excluded: $[13.7, 25.0, 50.0, 75.0, 100.0]$

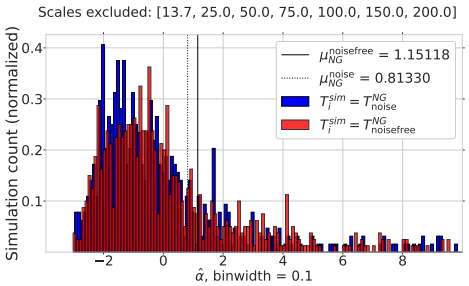


(a) Gaussian (b) Non-Gaussian

Figure B.5: Scales excluded: $[13.7, 25.0, 50.0, 75.0, 100.0, 150.0]$

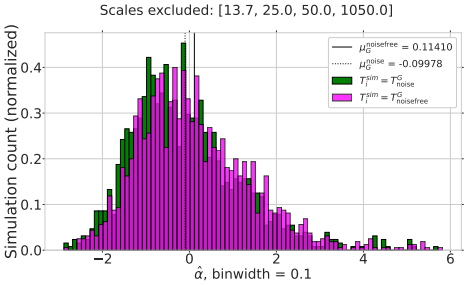


(a) Gaussian

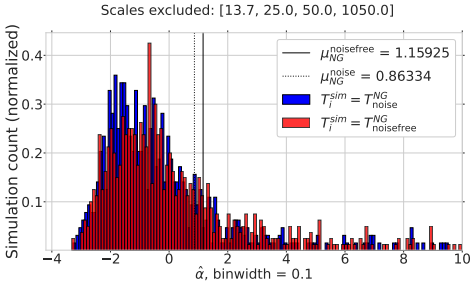


(b) Non-Gaussian

Figure B.6: Scales excluded: $[13.7, 25.0, 50.0, 75.0, 100.0, 150.0, 200.0]$

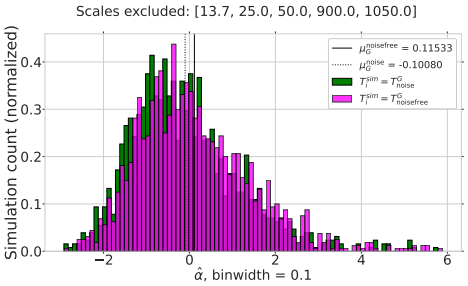


(a) Gaussian

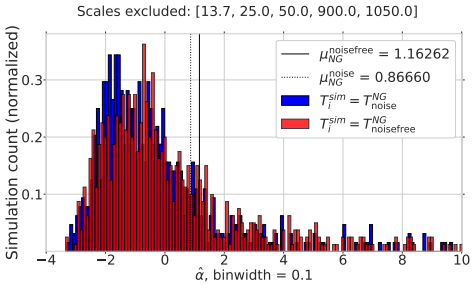


(b) Non-Gaussian

Figure B.7: Scales excluded: $[13.7, 25.0, 50.0, 1050.0]$

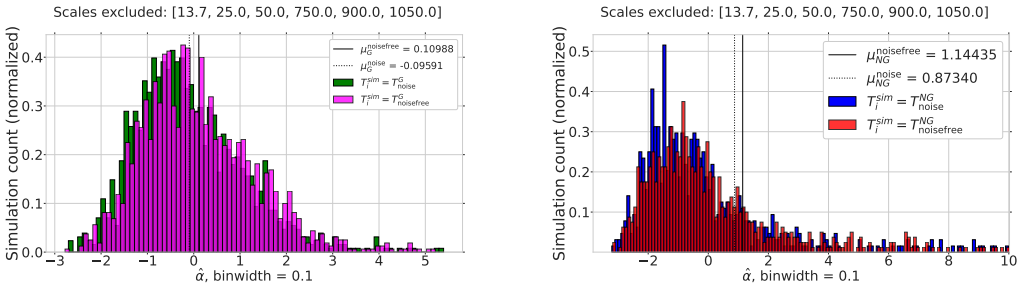


(a) Gaussian



(b) Non-Gaussian

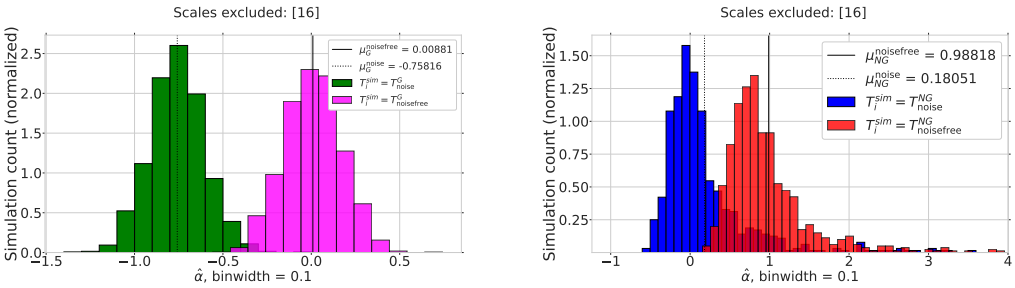
Figure B.8: Scales excluded: $[13.7, 25.0, 50.0, 900.0, 1050.0]$



(a) Gaussian (b) Non-Gaussian

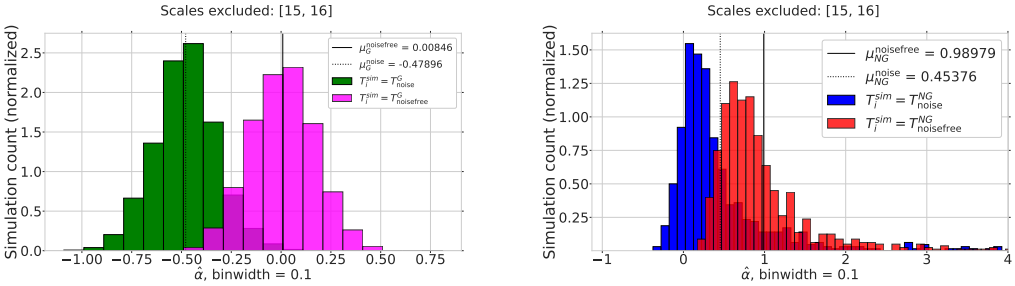
Figure B.9: Scales excluded: $[13.7, 25.0, 50.0, 750.0, 900.0, 1050.0]$

B.2 Standard needlets



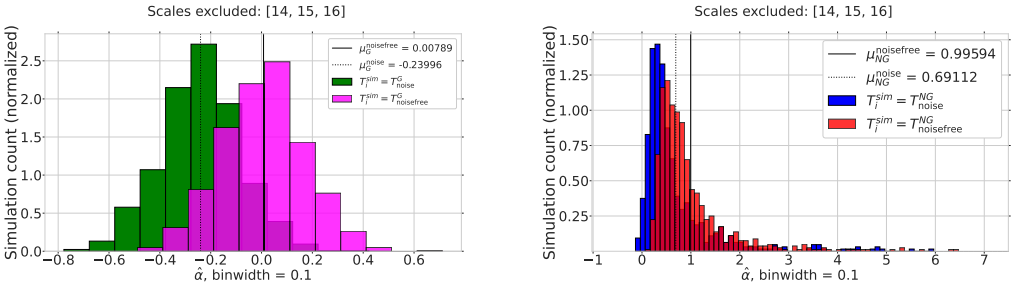
(a) Gaussian (b) Non-Gaussian

Figure B.10: Scales excluded: 16



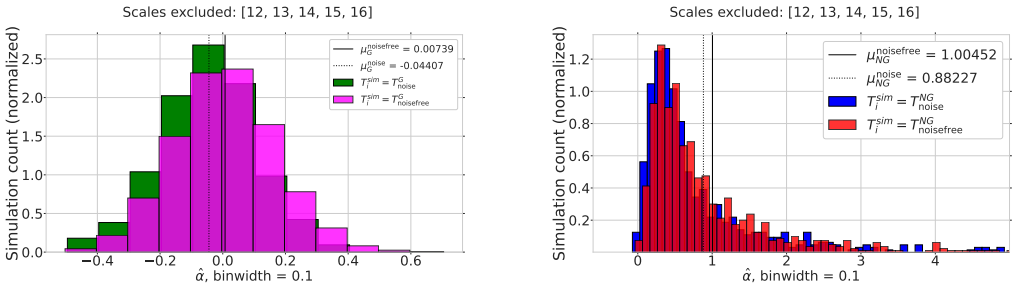
(a) Gaussian (b) Non-Gaussian

Figure B.11: Scales excluded: $[15, 16]$



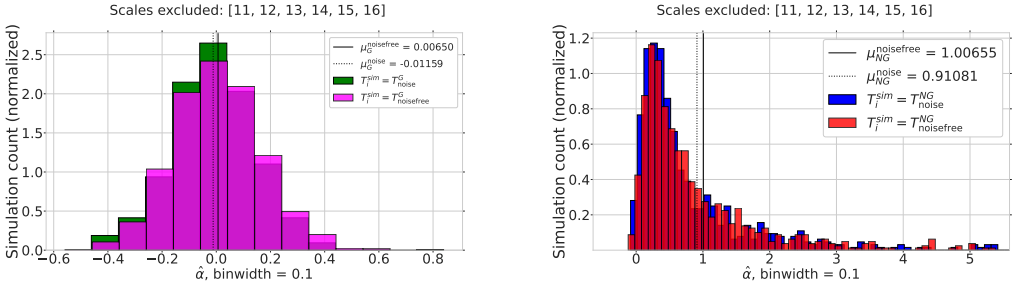
(a) Gaussian (b) Non-Gaussian

Figure B.12: Scales excluded: $[14, 15, 16]$



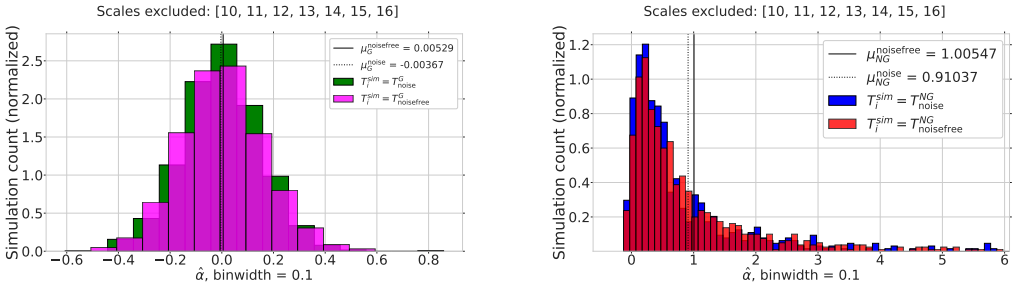
(a) Gaussian (b) Non-Gaussian

Figure B.13: Scales excluded: $[12, 13, 14, 15, 16]$



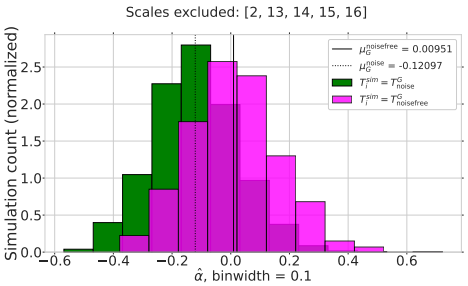
(a) Gaussian (b) Non-Gaussian

Figure B.14: Scales excluded: $[11, 12, 13, 14, 15, 16]$

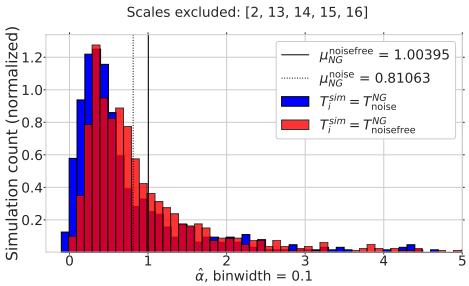


(a) Gaussian (b) Non-Gaussian

Figure B.15: Scales excluded: $[10, 11, 12, 13, 14, 15, 16]$

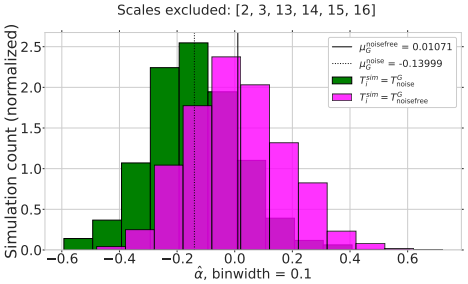


(a) Gaussian

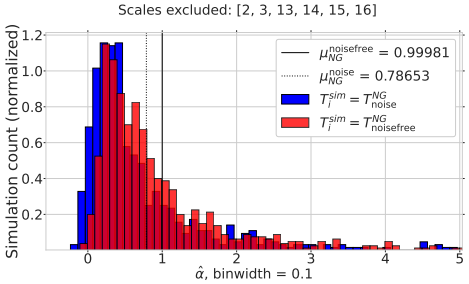


(b) Non-Gaussian

Figure B.16: Scales excluded: $[2, 13, 14, 15, 16]$

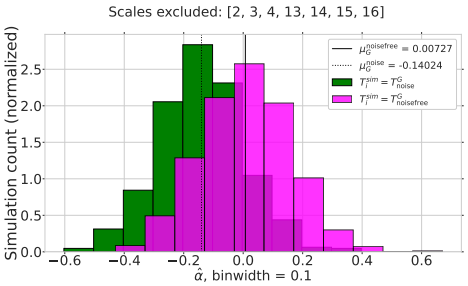


(a) Gaussian

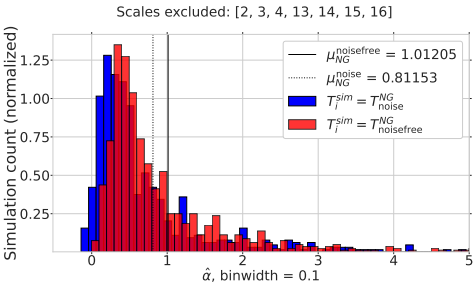


(b) Non-Gaussian

Figure B.17: Scales excluded: $[2, 3, 13, 14, 15, 16]$



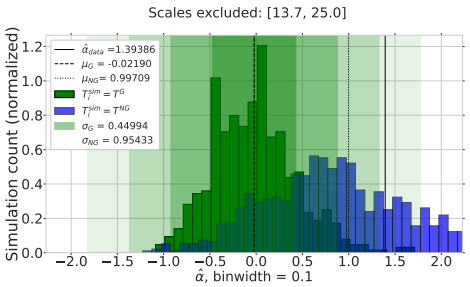
(a) Gaussian



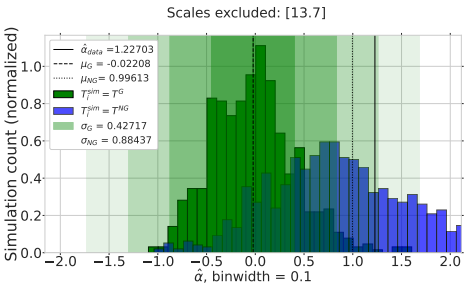
(b) Non-Gaussian

Figure B.18: Scales excluded: $[2, 3, 4, 13, 14, 15, 16]$

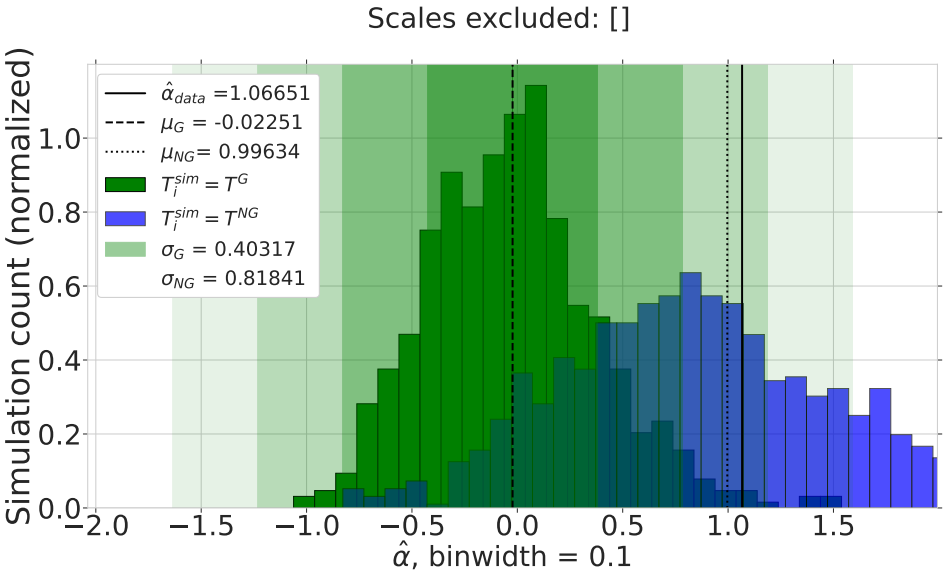
B.3 Re-inclusion of scales for $\hat{\alpha}$



(a) 2 exclusions

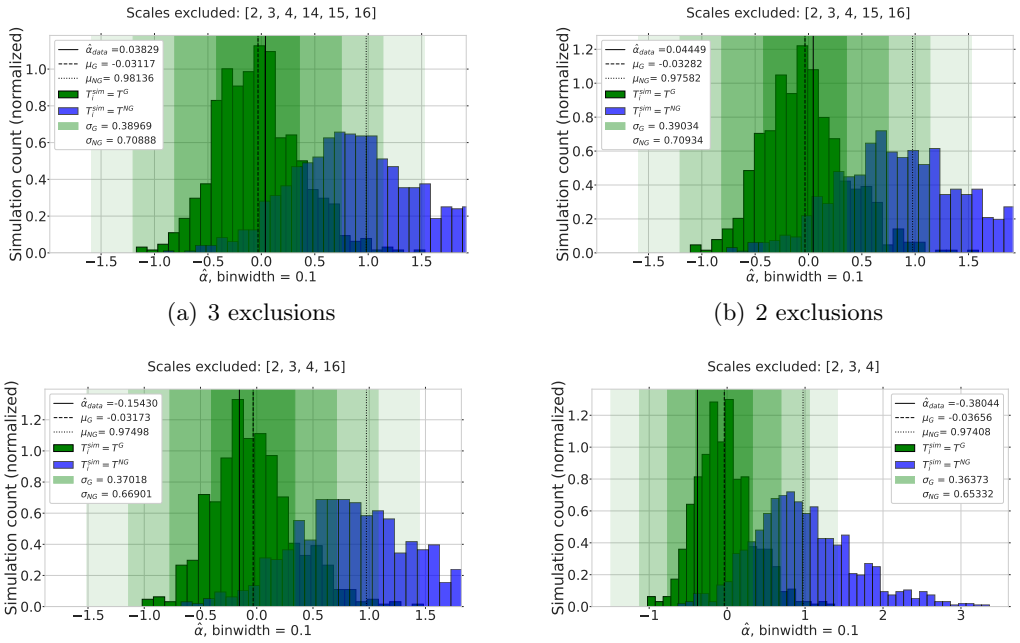


(b) 1 exclusion



(c) No exclusions

Figure B.19: $\hat{\alpha}$ shown for inclusion of the previously omitted SMHW scales, with the optimal trispectrum estimator.



(a) 3 exclusions

(b) 2 exclusions

(c) 1 exclusion

(d) no exclusions

Figure B.20: $\hat{\alpha}$ shown for inclusion of the previously omitted needlet scales, with the optimal trispectrum estimator.

Bibliography

[1] F. K. Hansen et al. "Isotropic non-Gaussian g_{NL}-like toy models that reproduce cosmic microwave background anomalies". In: 626, A13 (June 2019), A13.

[2] K. M. Górski et al. "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere". In: 622.2 (Apr. 2005), pp. 759–771.

[3] Martin B. Sweatman and Alistair Coombs. *Decoding European Palaeolithic art: Extremely ancient knowledge of precession of the equinoxes*. 2018.

[4] WMAP Science Team. *Cosmology: The Study of the Universe. NASA's Wilkinson Microwave Anisotropy Probe*. Available at `http://map.gsfc.nasa.gov/universe/`, last modified June 6, 2011.

[5] Albert Einstein. "Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]". German. In: 322.10 (1905), pp. 891–921.

[6] Albert Einstein. "Die Feldgleichungen der Gravitation. (German) [The Field Equations of Gravitation]". German. In: (1915), pp. 844–847.

[7] E. P. Hubble. "Extragalactic nebulae." In: *The Astrophysical Journal* 64 (Dec. 1926), pp. 321–369.

[8] E. P. Hubble. "A spiral nebula as a stellar system, Messier 31." In: 69 (Mar. 1929), pp. 103–158.

[9] Helge Kragh. *Masters of the Universe: Conversations with Cosmologists of the Past*. eng. Oxford: Oxford University Press, 2014.

[10] A. A. Penzias and R. W. Wilson. "A Measurement of Excess Antenna Temperature at 4080 Mc/s." In: *The Astrophysical Journal* 142 (July 1965), pp. 419–421.

[11] R. H. Dicke et al. "Cosmic Black-Body Radiation." In: *The Astrophysical Journal* 142 (July 1965), pp. 414–419.

[12] Alan H. Guth. "Inflationary universe: A possible solution to the horizon and flatness problems". In: 23.2 (Jan. 1981), pp. 347–356.

[13] Daniel Baumann. "TASI Lectures on Inflation". In: *arXiv e-prints*, arXiv:0907.5424 (July 2009), arXiv:0907.5424.

[14] ESA. *How many stars are there in the Universe?* Available at `https://www.esa.int/Science_Exploration/Space_Science/Herschel/How_many_stars_are_there_in_the_Universe`, accessed 17/10-2020.

[15] Planck Collaboration et al. "Planck 2013 results. XV. CMB power spectra and likelihood". In: 571, A15 (Nov. 2014), A15.

[16] Planck Collaboration et al. "Planck 2018 results. VI. Cosmological parameters". In: 641, A6 (Sept. 2020), A6.

[17] N. W. Boggess et al. "The COBE Mission: Its Design and Performance Two Years after Launch". In: 397 (Oct. 1992), p. 420.

[18] J. C. Mather et al. "Measurement of the Cosmic Microwave Background Spectrum by the COBE FIRAS Instrument". In: 420 (Jan. 1994), p. 439.

[19] C. L. Bennett et al. "The Microwave Anisotropy Probe Mission". In: 583.1 (Jan. 2003), pp. 1–23.

[20] C. L. Bennett et al. "Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results". In: 208.2, 20 (Oct. 2013), p. 20.

[21] Planck Collaboration et al. "Planck 2013 results. I. Overview of products and scientific results". In: 571, A1 (Nov. 2014), A1.

[22] Planck Collaboration et al. "Planck 2013 results. XII. Diffuse component separation". In: 571, A12 (Nov. 2014), A12.

[23] Planck Collaboration. "Planck 2018 results. VII. Isotropy and statistics of the CMB". In: 641, A7 (Sept. 2020), A7.

[24] Planck Collaboration. "Planck 2018 results. IX. Constraints on primordial non-Gaussianity". In: 641, A9 (Sept. 2020), A9.

[25] Planck Collaboration. "Planck 2018 results. IV. Diffuse component separation". In: 641, A4 (Sept. 2020), A4.

[26] Planck Collaboration. "Planck 2018 results. I. Overview and the cosmological legacy of Planck". In: 641, A1 (Sept. 2020), A1.

[27] Sean M. Carroll. *Spacetime and geometry. An introduction to general relativity.* 2004.

[28] Scott Dodelson. *Modern Cosmology.* Amsterdam: Academic Press, 2003.

[29] Planck Collaboration. "Planck 2013 results. XXII. Constraints on inflation". In: 571, A22 (Nov. 2014), A22.

[30] Alejandro Gangui et al. "The Three-Point Correlation Function of the Cosmic Microwave Background in Inflationary Models". In: 430 (Aug. 1994), p. 447.

[31] Licia Verde et al. "Large-scale structure, the cosmic microwave background and primordial non-Gaussianity". In: 313.1 (Mar. 2000), pp. 141–147.

[32] Limin Wang and Marc Kamionkowski. "Cosmic microwave background bispectrum and inflation". In: 61.6, 063504 (Mar. 2000), p. 063504.

[33] Eiichiro Komatsu and David N. Spergel. "Acoustic signatures in the primary microwave background bispectrum". In: 63.6, 063002 (Mar. 2001), p. 063002.

[34] Takemi Okamoto and Wayne Hu. "Angular trispectra of CMB temperature and polarization". In: 66.6, 063008 (Sept. 2002), p. 063008.

[35] Uros Seljak and Matias Zaldarriaga. "A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies". In: 469 (Oct. 1996), p. 437.

[36] Paniez Paykari and Jean-Luc Starck Starck. "Cosmic Microwave Background Data Analysis". In: *Advances in Machine Learning and Data Mining for Astronomy*. Ed. by Michael J. Way et al. 2012, pp. 55–87.

[37] Planck Collaboration. "Planck 2015 results. X. Diffuse component separation: Foreground maps". In: 594, A10 (Sept. 2016), A10.

[38] E. M. Leitch et al. "An Anomalous Component of Galactic Emission". In: 486.1 (Sept. 1997), pp. L23–L26.

[39] B. T. Draine and A. Lazarian. "Diffuse Galactic Emission from Spinning Dust Grains". In: 494.1 (Feb. 1998), pp. L19–L22.

[40] M. Gervasi et al. "The Contribution of the Unresolved Extragalactic Radio Sources to the Brightness Temperature of the Sky". In: 682.1 (July 2008), pp. 223–230.

[41] R. A. Sunyaev and Ya. B. Zeldovich. "The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies". In: *Comments on Astrophysics and Space Physics* 4 (Nov. 1972), p. 173.

[42] Planck Collaboration. "Planck 2013 results. XVI. Cosmological parameters". In: 571, A16 (Nov. 2014), A16.

[43] Planck Collaboration. "Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters". In: 594, A11 (Sept. 2016), A11.

[44] Planck Collaboration. "Planck 2015 results - XXII. A map of the thermal Sunyaev-Zeldovich effect". In: *A&A* 594 (2016), A22.

[45] H. K. Eriksen et al. "Power Spectrum Estimation from High-Resolution Maps by Gibbs Sampling". In: 155.2 (Dec. 2004), pp. 227–241.

[46] H. K. Eriksen et al. "Joint Bayesian Component Separation and CMB Power Spectrum Estimation". In: 676.1 (Mar. 2008), pp. 10–32.

[47] Soumen Basak and Jacques Delabrouille. "A needlet internal linear combination analysis of WMAP 7-year data: estimation of CMB temperature map and power spectrum". In: 419.2 (Jan. 2012), pp. 1163–1175.

[48] Soumen Basak and Jacques Delabrouille. "A needlet ILC analysis of WMAP 9-year polarization data: CMB polarization power spectra". In: 435.1 (Oct. 2013), pp. 18–29.

[49] S. M. Leach et al. "Component separation methods for the PLANCK mission". In: 491.2 (Nov. 2008), pp. 597–615.

[50] R. Fernández-Cobos et al. "Multiresolution internal template cleaning: an application to the Wilkinson Microwave Anisotropy Probe 7-yr polarization data". In: 420.3 (Mar. 2012), pp. 2162–2169.

[51] Jean-François Cardoso et al. "Component Separation With Flexible Models—Application to Multichannel Astrophysical Observations". In: *IEEE Journal of Selected Topics in Signal Processing* 2.5 (Nov. 2008), pp. 735–746.

[52] Krzysztof M. Gorski et al. "The HEALPix Primer". In: *arXiv e-prints*, astro-ph/9905275 (May 1999), astro–ph/9905275.

[53] Planck Collaboration. "Planck 2018 results. II. Low Frequency Instrument data processing". In: 641, A2 (Sept. 2020), A2.

[54] H. T. Ihle et al. "BeyondPlanck VI. Noise characterization and modelling". In: *arXiv e-prints*, arXiv:2011.06650 (Nov. 2020), arXiv:2011.06650.

[55] C. L. Bennett et al. "First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results". In: 148.1 (Sept. 2003), pp. 1–27.

[56] Eric Hivon et al. "MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets". In: 567.1 (Mar. 2002), pp. 2–17.

[57] L. Cayón et al. "Spherical Mexican hat wavelet: an application to detect non-Gaussianity in the COBE-DMR maps". In: 326.4 (Oct. 2001), pp. 1243–1248.

[58] J. -P. Antoine and P. Vandergheynst. "Wavelets on the n-sphere and related manifolds". In: *Journal of Mathematical Physics* 39.8 (Aug. 1998), pp. 3987–4008.

[59] E. Martínez-González et al. "The performance of spherical wavelets to detect non-Gaussianity in the cosmic microwave background sky". In: 336.1 (Oct. 2002), pp. 22–32.

[60] S. Scodeller et al. "Introducing Mexican Needlets for CMB Analysis: Issues for Practical Applications and Comparison with Standard Needlets". In: 733.2, 121 (June 2011), p. 121.

[61] F. J. Narcowich, P. Petrushev and J. D. Ward. "Localized Tight Frames on Spheres". In: *SIAM Journal on Mathematical Analysis* 38.2 (2006), pp. 574–594.

[62] F. Narcowich, P. Petrushev and J. Ward. "Decomposition of Besov and Triebel-Lizorkin spaces on the sphere". In: *Journal of Functional Analysis* 238.2 (2006), pp. 530–564.

[63] P. Baldi et al. "Asymptotics for spherical needlets". In: *The Annals of Statistics* 37.3 (June 2009), pp. 1150–1171.

[64] P. Baldi et al. "Subsampling needlet coefficients on the sphere". In: *Bernoulli* 15.2 (May 2009), pp. 438–463.

[65] D. Marinucci et al. "Spherical needlets for cosmic microwave background data analysis". In: 383.2 (Jan. 2008), pp. 539–545.

[66] N. Bartolo et al. "Non-Gaussianity from inflation: theory and observations". In: 402.3-4 (Nov. 2004), pp. 103–266.

[67] Xiohong Lan and Domenico Marinucci. "The needlets bispectrum". In: *Electronic Journal of Statistics* 2 (Jan. 2008), pp. 332–367.

[68] Øystein Rudjord. "Non-Gaussianity in the CMB: directional analysis and application of needlets". PhD thesis. University of Oslo, Jan. 2010.

[69] H. K. Eriksen et al. "Asymmetries in the Cosmic Microwave Background Anisotropy Field". In: 605.1 (Apr. 2004), pp. 14–20.

[70] F. K. Hansen, A. J. Banday and K. M. Górski. "Testing the cosmological principle of isotropy: local power-spectrum estimates of the WMAP data". In: 354.3 (Nov. 2004), pp. 641–665.

[71] Y. Akrami et al. "Power Asymmetry in WMAP and Planck Temperature Sky Maps as Measured by a Local Variance Estimator". In: 784.2, L42 (Apr. 2014), p. L42.

[72] Christopher Gordon et al. "Spontaneous isotropy breaking: A mechanism for CMB multipole alignments". In: 72.10, 103002 (Nov. 2005), p. 103002.

[73] H. K. Eriksen et al. "Hemispherical Power Asymmetry in the Third-Year Wilkinson Microwave Anisotropy Probe Sky Maps". In: 660.2 (May 2007), pp. L81–L84.

[74] J. Hoftuft et al. "Increasing Evidence for Hemispherical Power Asymmetry in the Five-Year WMAP Data". In: 699.2 (July 2009), pp. 985–989.

[75] F. K. Hansen et al. "Power Asymmetry in Cosmic Microwave Background Fluctuations from Full Sky to Sub-Degree Scales: Is the Universe Isotropic?" In: 704.2 (Oct. 2009), pp. 1448–1458.

[76] M. Axelsson et al. "Directional Dependence of ΛCDM Cosmological Parameters". In: 773.1, L3 (Aug. 2013), p. L3.

[77] Planck Collaboration. "Planck 2013 results. XXIII. Isotropy and statistics of the CMB". In: 571, A23 (Nov. 2014), A23.

[78] Planck Collaboration. "Planck 2015 results. XVI. Isotropy and statistics of the CMB". In: 594, A16 (Sept. 2016), A16.

[79] P. Vielva et al. "Detection of Non-Gaussianity in the Wilkinson Microwave Anisotropy Probe First-Year Data Using Spherical Wavelets". In: 609.1 (July 2004), pp. 22–34.

[80] Patricio Vielva. "A Comprehensive Overview of the Cold Spot". In: *Advances in Astronomy* 2010, 592094 (Jan. 2010), p. 592094.

[81] M. Cruz et al. "The non-Gaussian cold spot in WMAP: significance, morphology and foreground contribution". In: 369.1 (June 2006), pp. 57–67.

[82] D. N. Spergel et al. "First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters". In: 148.1 (Sept. 2003), pp. 175–194.

[83] Max Tegmark, Angélica de Oliveira-Costa and Andrew J. Hamilton. "High resolution foreground cleaned CMB map from WMAP". In: 68.12, 123523 (Dec. 2003), p. 123523.

[84] Jaiseung Kim and Pavel Naselsky. "Anomalous Parity Asymmetry of the Wilkinson Microwave Anisotropy Probe Power Spectrum Data at Low Multipoles". In: 714.2 (May 2010), pp. L265–L267.

[85] Jaiseung Kim and Pavel Naselsky. "Anomalous parity asymmetry of WMAP 7-year power spectrum data at low multipoles: Is it cosmological or systematics?" In: 82.6, 063002 (Sept. 2010), p. 063002.

[86] P. Naselsky et al. "Is the Cosmic Microwave Background Asymmetry due to the Kinematic Dipole?" In: 749.1, 31 (Apr. 2012), p. 31.

[87] Angélica de Oliveira-Costa et al. "Significance of the largest scale CMB fluctuations in WMAP". In: 69.6, 063516 (Mar. 2004), p. 063516.

[88] Neil J. Cornish et al. "Constraining the Topology of the Universe". In: 92.20, 201302 (May 2004), p. 201302.

[89] Kate Land and João Magueijo. "Is the Universe odd?" In: 72.10, 101302 (Nov. 2005), p. 101302.

[90] A. Gruppuso et al. "New constraints on parity symmetry from a re-analysis of the WMAP-7 low-resolution power spectra". In: 411.3 (Mar. 2011), pp. 1445–1452.

[91] Kate Land and João Magueijo. "Examination of Evidence for a Preferred Axis in the Cosmic Radiation Anisotropy". In: 95.7, 071301 (Aug. 2005), p. 071301.

[92] Antonino Troja et al. "The Needlet CMB Trispectrum". In: *Statistical Challenges in 21st Century Cosmology*. Ed. by Alan Heavens, Jean-Luc Starck and Alberto Krone-Martins. Vol. 306. May 2014, pp. 48–50.

[93] Planck Legacy Archive Wiki. *Simulation data*. `https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/Simulation_data`. Accessed: 2021-05-28. 2018.

[94] G. C. Wick. "The Evaluation of the Collision Matrix". In: *Physical Review* 80.2 (Oct. 1950), pp. 268–272.

[95] L. ISSERLIS. "ON A FORMULA FOR THE PRODUCT-MOMENT COEFFICIENT OF ANY ORDER OF A NORMAL FREQUENCY DISTRIBUTION IN ANY NUMBER OF VARIABLES". In: *Biometrika* 12.1-2 (Nov. 1918), pp. 134–139.