Tests of the statistical isotropy and Gaussianity of the Cosmic Microwave Background Radiation

by

BARTOSZ LEW

THE GRADUATE SCHOOL OF SCIENCE
NAGOYA UNIVERSITY
JAPAN

SUPERVISOR:

NAOSHI SUGIYAMA

DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL
OF SCIENCE OF THE NAGOYA UNIVERSITY
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF SCIENCE

9 SEPTEMBER 2008,
NAGOYA, JAPAN

To my wife Ola, who continuously,
thoughtfully and patiently supported me
form the beginning to the end of this work.

Acknowledgements

This dissertation summarizes the last 3.5 years of work pursued during my stay in Japan. My stay in Japan was generously promoted by the Japanese Government via the Monbukagakusho scholarship, granted by the Ministry of Education, Culture, Sports, Science and Technology, and it is hereby greatly and sincerely appreciated. In spite of securing my own, and my wife's long term stay in this amazing, and entirely new to us land of friendly, kind and gentle people, it also enabled me to pursue cosmological studies, at first in Mitaka (Tokyo) in the National Astronomical Observatory of the Graduate University for Advances studies, and latter at the Nagoya University. In order to ease my first steps through the, how enigmatic in the eye of a common foreigner, world of Japanese writing and speech, the scholarship package also provided with an introductory, intensive Japanese language course at the Tokyo University of Foreign Studies in Fuchu (Tokyo).

It was a great, colorful, sometimes immensely enduring, but always horizons-broadening journey for both of us, and it was filled many important moments spent with people whom we have met, and marked by friendships which have lasted over years; moments of witnessing a birth of our daughter, who joined us somewhere on the way, and times of continuous looking after and learning to be a father. All of that together left beautiful memories, which will always remain associated with Japan, which we have discovered and have learned to understand every day ever since the first step on the Japanese soil.

There are many people who I wish to thank. In particular I would like to express my uttermost gratitude to my supervisor Naoshi Sugiyama, who have continued to support me, and my progress in variety of ways, and who I want to thank for all of the help we received throughout our stay in Japan. Also I would like to thank Shioko Izumi-san for all the kindness I received, and for help with all of the paperwork that needed to be done during my stay in Mitaka. I wish to thank Sachiko Kuroyanagi and appreciate her help on the implementation of the final, last-minute corrections to the manuscript during the rush predeparture time. I also thank all other members of the "At lab" of the Nagoya University for enjoyable times spent together. I also wish to send special thanks to Olivier Doré for his involvement in helping to advance some of the work presented in this dissertation and for his encouragement and advice. I wish to thank Boudewijn Roukema for his careful reading and

correcting of some of my manuscripts, which helped elevating my English writing skills to more understandable levels. Finally, I wish to thank my wife Ola, and my daughter Hania, for keeping me socially active throughout the time, which given long hours spent by the lab-computer or laptop, is as much significant as anything else.

Contents

Chapter 1

Introduction

Progress during the last several years in precision and quality of cosmological data acquisition, as well as development of paralleled, dedicated numerical cosmological codes, and a significant increase in the accessible computing power opened up an era, where a consistency between theoretical predictions of a given cosmological model and the observations can be sought and statistically tested to a high accuracy. Today the wealth of the cosmological information, with a great model predictive power, consists of the observations of so called standard candles and standard rulers that help measure distances over cosmological scales. The cornerstone of these observations includes: the supernova explosion light curves measurements (Perlmutter et al., 1998; Riess et al., 2004; Astier et al., 2006; Astier et al., 2006; Wood-Vasey et al., 2007), the baryon acoustic oscillations features found in the correlation functions of the large scale galaxy distribution (Eisenstein et al., 2005; Gaztanaga et al., 2008), and the measurements of the primordial acoustic oscillation fluctuations imprinted in the cosmic microwave background radiation (CMBR) (Hinshaw et al., 2008; Komatsu et al., 2008; Dunkley et al., 2008; Sievers et al., 2003; Reichardt & ACBAR collaboration, 2007). These observables strongly depend on cosmological parameters, and probe completely different ranges of redshifts and scales, providing thereby a strong leverage for cosmological model selection. There are also constrains coming from the particle physics: most notably from the theory of the big bang nucleosynthesis (BBN) that provides tight limits on the primordial light elements abundances as a function of the baryon to photon number densities ratio (Cyburt et al., 2003; Kawasaki et al., 2005), which in turn is very precisely fixed by the current CMB observations (Coc, 2008). Additionally the measurements of the cluster abundance via cluster mass functions (Bahcall et al., 2003), measurements of the Ly_α absorption line widths in the spectra of the distant quasars, which are dubbed the *Lyman-α forest* (Jannuzi et al., 1998; Weymann et al., 1998; Bechtold et al., 1994; Cristiani et al., 1997; Lu et al., 1996) as well as the observations of the weak gravitational lensing of the CMBR on the large scale structures (LSS) (Bacon et al., 2003; Bartelmann & Schneider, 2001; Lewis & Challinor, 2006; Cooray & Hu, 2001), and the cosmic shear observations (Hoekstra et al., 2006; Bacon et al., 2003; Kaiser et al., 2000; Bacon et al., 2000) provide data sets that yield cosmological constraints which are broadly mutually consistent, and

provide altogether a proof for an isotropic and homogeneous and spatially flat cosmological model with considerable amount of so-called cold dark matter (CDM) and small amounts of baryonic matter, with nearly scale invariant spectrum of Gaussian, adiabatic, primordial perturbations (Gnedin & Hamilton, 2002; Tegmark et al., 2006; Lewis & Bridle, 2002; Sievers et al., 2003; Reichardt et al., 2008; Seljak et al., 2006; Dunkley et al., 2008). The observations are found to be vastly consistent with the ΛCDM cosmological model, with parameters tightly constrained in a number of studies relying on various compilations of these observations (Lewis & Bridle, 2002; Seljak et al., 2006; Tegmark et al., 2006; Nesseris & Perivolaropoulos, 2004). The constrained model is currently the simplest, most favoured cosmological scenario which, in spite of few cosmological and BBN conundrums (eg.Coc (2008); Cyburt et al. (2008); Kusakabe et al. (2008); Jittoh et al. (2007); Cumberbatch et al. (2007); Jedamzik et al. (2006); Primack (2004)), is capable to reconcile between the majority available cosmological data sets.

Independently of the technological and observational progress, the theory of inflation has been invented (Guth, 1981) and continues to be developed in a huge variety of models (Kim & Liddle, 2006; Kecskemeti et al., 2006; Bernardeau et al., 2006; Bartolo et al., 2004; Garriga & Mukhanov, 1999; Silverstein & Tong, 2004; Huang et al., 2008; Arkani-Hamed et al., 2004; Dvali et al., 2004; Bartolo & Liddle, 2002; Sasaki, 2008; Golovnev et al., 2008) while its original idea of a phase of accelerated expansion and generic predictions remain intact. These include the adiabatic, nearly Gaussian, curvature perturbations, imprinted over a spatially flat and isotropic background, with almost scale invariant power spectrum. These predictions remain in great consistency with a wealth of cosmological observations, and therefore the inflationary theories currently provide the basis for understanding of the history of early Universe, and the structure formation (Alabidi & Lidsey, 2008; Gordon & Lewis, 2003; Kinney et al., 2004; Spergel et al., 2007; Komatsu et al., 2008).

There are however number of observations that apparently stand out from predictions of the standard model (eg.Cyburt et al. (2008); Primack (2004), or exhibit evidence of incompatibility with other data-sets (eg. Diego et al. (2004)). In particular, in spite of great improvements in the resolution and sensitivity since the observations of the COBE satellite (Boggess, 1992; Bennett et al., 1992), some anomalies continue to exist in the currently accessible, full sky maps of the CMBR accumulated during years of observations of the WMAP satellite (Hinshaw et al., 2008), and some new and unexpected peculiarities have been found both in the temperature angular power spectrum, and in the topological properties of the CMB maps (eg. Gaztañaga et al. (2003); Cline et al. (2003); de Oliveira-Costa & Tegmark (2006); Efstathiou (2003b); Diego et al. (2004); Eriksen et al. (2004a, 2007); Land & Magueijo (2005c); Rakic & Schwarz (2007); Cruz et al. (2007); McEwen et al. (2006a); Covi et al. (2006) and references therein). Currently, these anomalies only tentatively elude the predictions of the simplest ΛCDM model, but they may hint on variety of processes that may take place beyond the standard framework, as well as help concretize some of the viable possibilities that still fit within the standard scenario (eg. Niarchou &

Jaffe (2007, 2006); Weeks & Gundermann (2006); Kunz et al. (2006); Jaffe et al. (2005); Alabidi & Lyth (2006); Adams & Cresswell (2003); Wang & Kamionkowski (2000); Covi et al. (2006); Destri et al. (2008); Erickcek et al. (2008); Akofor et al. (2007); Inoue & Silk (2007)).

The motivation of the work pursued in this thesis are the anomalies observed in the data, that elude the predictions of the standard cosmological model. Amongst these, the Gaussianity and the statistical isotropy are two basic predictions, and currently accessible, and directly testable observables via the CMBR data. Constraints put on these observables with the advent of the forthcoming, next-generation experiments will also soon become useful probes of the physics of the early Universe with sufficient power to discriminate between families of inflationary models.

In this work we address some of the known anomalies of the CMB data, and tests them in a new, and independent way. Our aim is to conduct a battery of statistical tests of the CMB data in order to search for sings of any new departures from these two generic predictions. We also aim at quantifying the statistical significance of the well known anomalies, and to test their robustness and stability under different tests and different data processing pipelines.

There are many non-cosmological (astrophysical, or instrumental) sources that may violate Gaussianity, or break the statistical isotropy. Amongst these most importantly, the residual extended foregrounds dominating at large angular scales, and galactic and extragalactic point sources as well as the secondary effects (like weak lensing, Sunyaev-Zeldovich effects) dominating at small angular scales. It is therefore important to detect, analyze and localize these sources, and either eliminate them and/or exclude from the data, which otherwise could lead to a biased results of the analyses aiming at constraining the inflationary physics. If all other explanations fail, and the data will robustly and significantly violate the Gaussianity, or statistical isotropy, such detection will become an extremely useful observable, calling for explanation within some alternative available but currently indistinguishable, competing models.

The organization of the Thesis is as follows:

An introductory review of the most relevant anomalies found within the standard cosmological model with various implications for cosmology are given in chapter 2.

The main part of the analysis presented in the thesis devoted to various tests of Gaussianity and statistical isotropy of the Wilkinson Microwave Anisotropy Probe (WMAP) CMBR data is given in chapter 3. The results presented in there were published in Journal of Cosmology and Astroparticle Physics (Lew, 2008b).

In chapter 4 we focus on constraining parameters of one of the alternative models, involving a large scale violation of the statistical isotropy by a bipolar modulation field. The results presented there are now accepted for publication in Journal of Cosmology and Astroparticle Physics (Lew, 2008a).

In chapter 5 we test the Gaussianity of the CMB using Minkowski functionals and describe a significant detection of residual, large scale foreground contamination leading to systematical deviation from foreground-free simulations expectation in one the functionals, as well as in a systematical offset of the mean and in the negative skewness of the one-point temperature distribution, along a ridge adjacent to the galactic plane, outside the commonly used, conservative sky masks. The results presented in this chapter are now in preparation for submission to the Journal of Cosmology and Astroparticle physics.

In chapter 6 we introduce a new dedicated statistics for testing deviations from the statistical isotropy, which are generated by the apparent, tentative anomalies in the reconstructed angular power spectrum of the CMB at largest scales. In chapter 7 we present first results from an ongoing project, dedicated to tests of the violations of the statistical isotropy, observed in the alignments between low-ℓ multipoles. We introduce relevant statistics, test and compare range of available renditions of the foregrounds cleaned CMB maps, and discuss some of the resulting problems and propose few generalized statistics that could be interesting to implement in the further studies.

Finally, in chapter 8, we investigate the hypothesis of the multiply connected Poincaré dodecahedral space model, which we test against the generic ΛCDM simply-connected model using realistic Monte-Carlo CMB simulations. The results presented in chapter 8 were published in Astronomy and Astrophysics 2008A&A 482...747L (Lew & Roukema, 2008).

Chapter 2

Anomalies in the Cosmic Microwave Background observations and cosmological implications

In this chapter we review the most related work pursued by various authors, dedicated to exploration of consistency of the CMB data with the predictions of Gaussianity and statistical isotropy through a large number of statistical tests. These studies continue to be increasingly significant since the COBE detection of the primordial metric fluctuations (Bennett et al., 1994) which brought about the first evidences for the scale invariant power spectrum of initial perturbations, predicted by the simplest, single-field, slow-roll, inflationary models, and constituted a strong support for the isotropy, and homogeneity of the Universe, measured at a level of $\sim 10^{-5}$, and thereby provided with the calibration of the amplitude of the initial perturbations at the currently observable close-horizon scales.

The tests, optimized for either measuring consistency with Gaussianity, or statistical isotropy, or randomness of the underlying density field constitute one, and a very limited approach of the issue of testing the two predictions. Another, and more involving way of addressing the problem is to question the degree, type and shape of plausible non-Gaussianity, and/or non-randomness, since the latter two can be realized in an infinite number of ways, while former two can only be realized in one way. Therefore a non-detection of the non-Gaussianity, or non-randomness in one of the tests does not automatically imply that the field is Gaussian or random. An example of this is given in figure 2.1. It is therefore important to come up with many different statistical tests, which are maximally independent, and *orthogonal* to each other in order to better assess the question of Gaussianity or statistical isotropy. Consequently, over the period of last several years there has been put an incredibly large effort in devising a whole battery of cosmological, statistical tests, whose only purpose was to address these two generic predictions, of the inflationary Universe. Insofar though, the properly foregrounds cleaned and/or masked CMB data generally comply with

these assumptions, however there are few or several exceptions. In this chapter we will focus on these exceptions, and discuss possible implications for cosmology by analysing viable models that are capable of reproducing them.

Non-Gaussianity vs non-randomness It is important to note that testing statistical isotropy and Gaussianity cannot entirely be separated from each other. This is because one has to deal with only a single realization of the surface of last scattering. Given no prior knowledge on any possible underlying temperature template, which leads to violations of the statistical isotropy, like in eg. Bianchi type models, it is not possible to ascribe any plausible detections of non-Gaussianity only to the brake down of Gaussianity, because it is possible that the deviation might be caused by an underlying anisotropic template upon which a Gaussian temperature field is generated. Since the underlying template could only be constrained via statistical measurements, using many different field realizations, there's a strong interplay between the two predictions. In the common day practise, it is widely assumed to consider the CMB field as random realization, and referring to deviations from GRF as to detections of non-Gaussanity, or statistical anisotropy depending on scales at which the detection occurred. As such the detections at small scales are often referred as to non-Gaussianity, whereas large scale anomalies are rather referred to as violation of statistical isotropy.

In figure 2.1 we plot an example of Gaussian random field (left-hand side plot) and Gaussian non-random field (right-hand side plot). The latter was obtained from the former only via reorganization of the pixels in the map (and recalibration of the power spectrum due to the smoothing effects which were destroyed upon pixel rearrangement). The two maps have exactly the same variance and power spectrum. They also have almost identical mean and kurtosis, and hence cannot be distinguished via these estimators. Performing a Gaussianity test using Minkowski functionals (bottom panels) we discover a significant differences between the two maps via the circumference of an enclosed contour at a given temperature threshold, and in the genus statistic. It is interesting to note, that instead of the signs of non-Gaussianity we actually detected the violation of the statistical isotropy of the right-hand side map, by performing tests, which one might think are a Gaussianity tests. Of course, in order to put it quantitatively correct a significance analysis would be in order, however qualitatively it is easy to see that it is possible to violate the statistical isotropy of the map by an arbitrary amount, and thereby reach an arbitrary level of deviation from the Gaussian random field, and therefore qualitatively the same results would hold.

It is interesting to note that in this case, the Minkowski area functional is actually helpless in detecting the difference between the two maps, as is the kurtosis, or the mean[1]. This exactly embodies the aforementioned need for implementation of many different Gaussianity tests in order to pin down the nature of possible violation of the Gaussianity or statistical isotropy.

[1]Note that given that the power spectrum is exactly preserved in the right-hand side map, any differences in the mean of the maps should results only from special arrangements between the phases in the non-random map.

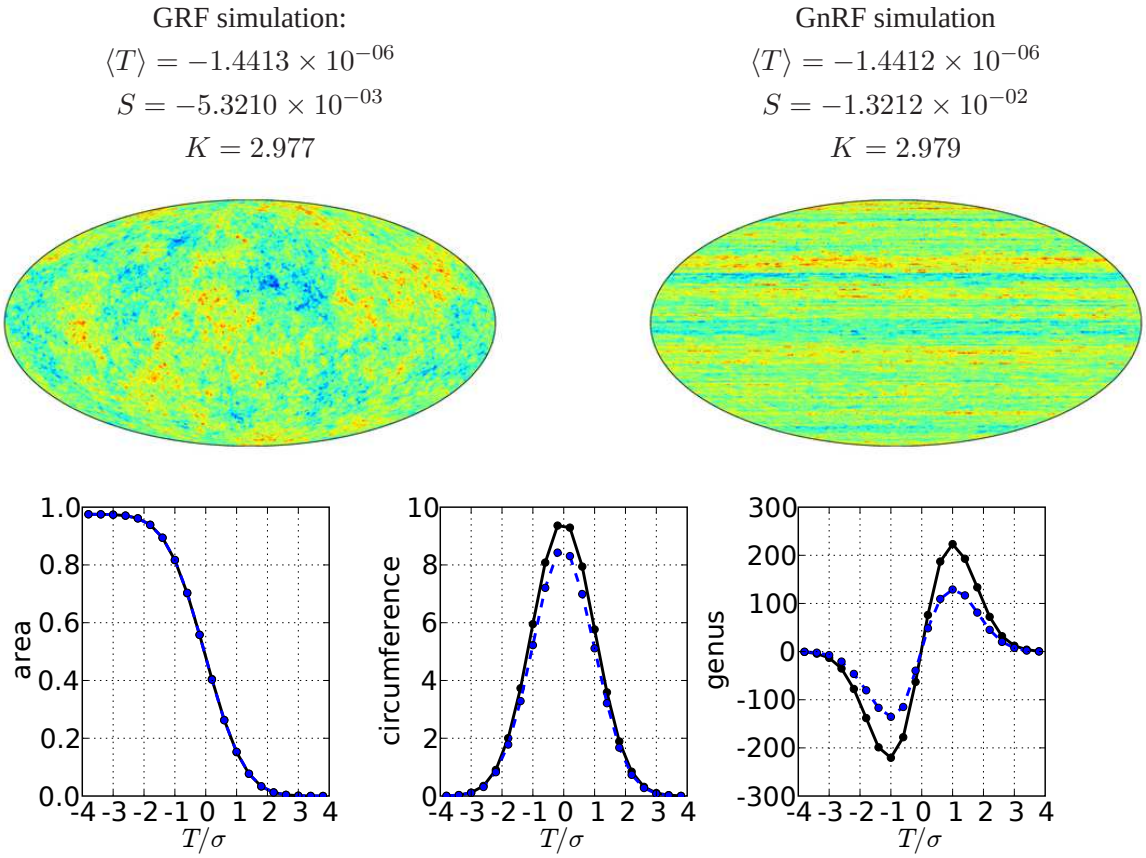


Figure 2.1: An example of a Gaussian random field (GRF) (top-left panel) and Gaussian non-random field (GnRF) (top-right panel). The non-random simulation was obtained from the Gaussian one by reordering the map pixels by treating the Healpix nested ordered map as Healpix ring ordered map and correcting for the power spectrum change due to beam effects. Consequently the two maps have exactly the same power spectrum (and hence the variance). Yet the map on the right hand side appears to be clearly non random due to horizontal stripes. The two maps have almost exactly the same mean, and kurtosis as well. In the bottom panel we plot the three Minkowski functionals for the left-hand side map (black) and the right-hand side map (blue). Since the right-hand side map still remains a Gaussian map (reorganization of pixels not alter the temperature statistics and so does not violate the Gaussianity), the difference results only due to break down of the statistical isotropy.

2.1 Hemispherical power asymmetry

2.1.1 Search and Evidence

The statistical isotropy has been tested by various authors using a whole spectrum of different methods. Some of the latest works based on the WMAP CMB data include Souradeep et al. (2006); Hajian & Souradeep (2006); Gordon (2007); Armendariz-Picon & Pekowsky (2008); Samal et al. (2008); Hansen et al. (2004a,b); Donoghue & Donoghue (2005); Bernui

et al. (2007b,a); Rakic et al. (2006); Gordon et al. (2005); Eriksen et al. (2004a, 2007, 2008); Hajian & Souradeep (2005, 2003); Bernui & Hipólito-Ricaldi (2008) and references therein.

In particular Hansen et al. (2004a) pursued tests of the statistical isotropy based on local measurements of the power spectrum in circular patches, equally distributed in the sky. The analysis was performed using the power spectrum reconstructions, based on pseudo-C_ℓ method, introduced in Hansen et al. (2002), and these local C_ℓ measurements were analysed in various multipole ranges and in patches of various angular sizes. Consequently several interesting anomalies have been found. Most notably, the evident suppression of power in the northern galactic (and ecliptic) hemispheres as compared to the GRF realizations was realized, spanning over the multipole range from $\ell = 5$ to $\ell = 40$. The hemispherical power spectra measurements proved the asymmetry to be maximized along the axis $(l, b) = (237^\circ, -10^\circ)$ in galactic coordinates with significance varying from less than one percent up to several percent as compared with GRF simulations, depending on the particular multipole bin.

Also, it was tentatively found, that the apparent outliers in the reconstructed power spectrum released by the WMAP team, and available at
http://lambda.gsfc.nasa.gov/data/map/dr3/dcp/wmap_tt_spectrum_5yr_v3p1.txt, seen at multipoles $\ell = 21$ and $\ell = 39$, are possibly associated with northern and southern hemispheres respectively and exclusively. Additionally, it was noted that the sharp local decrease, seen in the power spectrum around the first acoustic peak ($\ell \approx 182$), referred to as the "dent", was associated with few circular discs directly adjacent to the galactic plane, and hence the residual foregrounds have been suggested as a possible source for that feature. Interestingly, it was also found that the power asymmetry orientation is somewhat scale dependent, with larger scales preferring larger galactic latitudes and smaller scales preferring smaller galactic latitudes.

In the same year Hansen et al. (2004b) carried out another test of the statistical isotropy using regional, hemispherical measurements, based on the Hessian matrix calculations, of the local minima, maxima and saddle points, of the WMAP CMB V+W inverse noise co-added maps, and their simulated distributions, for few tested smoothing scales ranging from 0° to 15°. By varying the orientations of the hemispherical regions they found the maximal asymmetry orientation in particular, as far as the hemispherical ratios of number of lakes and hills are concerned, and they found the northern hemisphere to be anomalous at 2σ to 3σ confidence level at scales ranging from 1° to 5°. Although the authors do not mention the exact orientation of the hemispheres that maximize the asymmetry, they do mention that the results are widely consistent with another related work by Eriksen et al. (2004a), who also measured, and detected the hemispherical power asymmetry by deriving ratios of power spectra estimated on two opposing hemispheres, and quantified that the anomaly significance is as large as 99.7% for the multipole range $\ell \in [2, 40]$ in co-added V and W channel maps of the WMAP with Kp2 sky mask applied. The authors also noted that their results are stable with respect to the galactic sky cut and frequency band.

In Park (2004) a genus statistics has been utilized for stereographic projected WMAP

CMB maps from the first-year release, in order to address the power asymmetry question, and although the asymmetry has been consistently found, its significance was estimated as of somewhat lower confidence level of about 99%, which was further decreased down to only 95%, when the galactic sky mask was extended up to $|b| < 30^{\circ}$, which apparently contradicts the results of Eriksen et al. (2004a), who points out that the significance of the asymmetry is robust under extended, to the very same extent ($|b| < 30^{\circ}$), sky cuts.

Yet another independent estimator of the statistical anisotropy has been introduced and applied by Bernui et al. (2007b), based on histograms of angular separations of points lying within certain temperature threshold. By measuring the threshold averaged residual (difference between the measured and expected) distributions, referred to as *EPASH*, within specifically selected circular regions in the sky, the method is capable of tracing local deviations from the statistical isotropy. By calculating the variance of the *EPASH* estimators, and computing its spatial, angular distribution over the sky, utilizing 12288 uniformly distributed caps and comparing the power spectrum of such produced anisotropy map with the fiducial simulated analogical power spectra, the authors estimated the significance of the "north-south" asymmetry at the level of 97% in the dipole component of the anisotropy map, and even more significant: 99% and 99.9% for the quadrupole and octupole components respectively. While the authors arrive at similar conclusions in their follow-up paper Bernui et al. (2007a), it is not clear whether significance assessment is robust since the authors do not provide the details on their simulations approach, where as the earlier paper suggests that the power spectrum properties of the studied map (ILC map) were not simulated consistently with the observations.

Gordon & Trotta (2007) investigated the significance of the asymmetry via the Bayes factor analysis - applied also for quantifying the significance of the scale invariance of the primordial power spectrum. While the authors found the odds of the tilt of the power spectrum to be of order 49:1 (roughly 2.3σ), in favour of non-vanishing tilt, the CMB power asymmetry was found to be realized in odds 9:1 (roughly 1.7σ) under the null hypothesis (corresponding to the lack of any anisotropic, possibly modulation-like signals). This result is consistent with the results presented in chapter 3 in section 3.5.3 and 3.6.2 and those in chapter 4. The significance estimated in Gordon & Trotta (2007) clearly contradicts the results reported by Eriksen et al. (2007).

Eriksen et al. (2007) has performed an analysis of the hemispherical power asymmetry working within a frame of an anisotropic modulation model by Gordon et al. (2005); Gordon (2007), and constrained the relevant bipolar modulation field parameters: i.e. its orientation and amplitude, using a maximum likelihood and a maximum posterior methods, and implementing the Metropolis-Hastings algorithm for multi-dimensional posterior integrals. While we will describe the modulation model in more detail in the latter chapters, we note that Eriksen et al. (2007) estimated the significance of the asymmetry via non-vanishing multiplicative modulation CMB component of amplitude $A = 0.114$ and excluding the isotropic Universe value of $A = 0$ at the confidence level in excess of 99%, with small dependence on the data-set analysed. The best fit modulation field orientation was found to

be $(l, b) = (225^\circ, -27^\circ)$, and since the performed analysis operated on very low resolution maps the effective multipole up to which the analysis was sensitive was $\ell \lesssim 40$, at which the signal-to-noise ratio reached unity. We note that the data-set analyzed in that work was simply downgraded from high resolution maps, scrambling through all scales, rather than filtering the low-ℓ part of the signal form the initial maps. This approach contrasts the method used in this thesis (in chapter 4) where we apply a band-pass filters in the spherical harmonic space in order to select the preferred part of the input signal, and thereby better control the scales upon which a potential modulation works.

While more references to this work will be done in chapter 4, given that the evidence for the asymmetry ranges from about 2σ to 3σ depending on the author and pursued method and the statistical significance assessment, there has been already significant effort put in finding possible theoretical explanations for this kind of effect, because the CMB asymmetry could be a manifestation of, and an interesting window onto, a possibly new physics. In the following we mention some of the plausible proposed explanations.

2.1.2 Implications for cosmology

A range of possible explanations have been proposed for the observed power asymmetry, which clearly manifests a break down of the statistical isotropy, at least at large angular scales. The break down of the statistical isotropy is highly undesirable, because of the tremendous success of the consistency of all other CMB features with the predictions of the inflationary Universe: most notably, the near scale-invariance of the power spectrum of the primordial perturbations of the gravitational potential Φ, their Gaussianity, and also flatness of the spatial sections of space, as well as solutions of the long-standing cosmological horizon and flatness problems. Therefore dropping the idea of inflationary period in the early evolution of the Universe would rise a need for alternative explanations, which would need to deal with all of these problems jointly or independently. Currently, no scenario is equally successful as the inflationary. Since the standard, single-field inflationary solutions (eg. based on chaotic inflation) do not produce significant hemispherical anisotropy, solutions were devised that explain the asymmetry within the somewhat modified or more complicated models.

Recently Erickcek et al. (2008) proposed that the curvaton scenario (Lyth & Wands, 2002; Bartolo & Liddle, 2002) could give rise to such asymmetry under certain implementation. The generic curvaton models assume that apart from the inflaton field, which primarily contributes to the energy density of the Universe during inflation, there's another minimally coupled scalar field - curvaton - which does not contribute significantly to the total energy density during inflation, nor does it drive the inflation, and hence need not to yield the slow-roll conditions, and consequently need not become very homogeneous during inflation, so as to meet the constraints from the CMB measurements. In its minimal version the potential can be assumed as: $V(\phi, \sigma) = \frac{1}{2}M^2\phi^2 + \frac{1}{2}m^2\sigma^2$ where σ is the curvaton field and ϕ the inflaton field and m and M are their masses respectively. The curvaton is assumed to roll-

fast (faster than the inflaton) to sufficiently small values of σ so as to prevent the stage of second inflation after the inflaton decay. This value is assumed to be roughly constant over the period of inflation: eg. $\sigma = \sigma_\star$ and the curvaton is assumed to stay there until long after the inflation is finished. Typically the inflaton curvature perturbations are assumed to be small (as compared to the level inferred from the CMB observations) while the isocurvature perturbations of the curvaton at horizon exit can be large: $\delta\sigma \sim H_\star/2\pi$ over a one Hubble time. Gradually as the $H(t) \sim m$ the curvaton begins to decay into standard model particles (eg. photons), reheating the Universe and beginning the onset of the standard hot big-bang scenario[2]. Hence the inflaton is being released from responsibility of generating curvature perturbations. Instead, the iso-curvature perturbations are converted into curvature ones as the curvaton decays and thermalizes in some processes proposed by eg. Mollerach (1990); Lyth & Wands (2002). This model is not only capable of having some isocurvature component in the final power spectrum, proportional to $\mathcal{P}_{\Phi,\sigma} \sim \left(\frac{H_\star}{2\pi\sigma_\star}\right)^2$, but also significant non-Gaussianity: $f_{\rm NL} \approx 5\xi^2/(4R)$, where ξ is the fraction of the perturbations due to curvaton in the total curvature metric perturbations, generated due to both: the curvaton and the inflaton together; and $R = \rho_\sigma/\rho_{\rm tot}$ is the ratio of the energy density due to curvaton decay to the total energy density. In Erickcek et al. (2008) it was assumed that the curvaton actually does not contribute significantly to the total energy density: i.e. $R \ll 1$ which will grant the sufficient (controllable amount of) homogeneity of the gravitational potential perturbations Φ, even if the fluctuations in ρ_σ are large $\mathcal{O}(1)$. Then it was hypothesized that the observed asymmetry in the CMB is a consequence of the spatial variation of the $\sigma_\star$ value in the observed patch of the Universe, due to some large super-horizon mode. Constraints were then put on the allowed values of R and ξ from the observations of: the level of primordial non-Gaussianity $f_{\rm NL} \lesssim 100$ (Komatsu et al., 2008; Komatsu et al., 2003; Yadav & Wandelt, 2008), the CMB quadrupole value and the degree of the power asymmetry in the CMB (eg. Eriksen et al. (2007)).

However it has been shown that the modulation signals in the CMB are generally extending only up to $\ell_{\rm max} \lesssim 40$ (eg. Hansen et al. (2004a), see. also chapter 4), while the asymmetry generated in the above model would be scale independent and hence some generalizations would be required in the curvaton power spectrum in order to fit to this observation. In that case it's possible that some fine tuning would arise. It was however suggested that it will be useful to utilize the polarization signals, induced from quadrupole charge configurations at the surface of last scattering, to possibly reinforce the evidence in favour of the cosmological origins of the power asymmetry.

These issues were analytically and numerically studied by Dvorkin et al. (2008) who also pointed out that given that the temperature observations have already reached the precision limited only by the cosmic variance (at the relevant scales) the situation can be improved by precise, large scale polarization measurements - the quest to be challenged by the

[2]Note that after the inflaton decay, the Universe becomes already radiation-dominated, as the curvaton field does not contribute significantly to the total energy density; which otherwise could trigger another inflationary stage.

PLANCK surveyor mission.

Very recently yet another possibility has been investigated (Bernui & Hipólito-Ricaldi, 2008) that the primordial, cosmological, magnetic fields with magnitude of order $B \sim 10^{-9}$G could be responsible for generation of both, the planar and aligned low-order CMB multipoles (see next section) and the hemispherical power asymmetry. A range of the magnetic induction values were considered ($0 < B < 30$ nG) and it was found that stronger alignments are obtained for larger field values. However in the earlier work by Kahniashvili et al. (2008) a scale dependent constraints on the amplitude of the primordial magnetic field have been derived, for a range of primordial magnetic field power spectrum spectral indexes resulting from "plausible magnetic field generation mechanisms". The constraints were based on the recent limits on the B-polarization mode of the CMB power spectrum Komatsu et al. (2008) and compared with generated theoretical B-mode power spectra for the primordial magnetic field induced Faraday rotation, converting the E-polarization modes generated at last scattering into B-polarization modes. In particular, for the scale of 1 Mpc, B is constrained to be $0.4 < B < 30$ nG, depending on the spectral index of the primordial magnetic field power spectrum, whereas at the 100 Mpc scale, the constraint yields $0.7 < B$ nG, for all considered spectral indexes $-2.9 < n_B < -1$, and yet even less for the Giga-parsecs scales. The estimates of Kahniashvili et al. (2008) seem as an improvement to the previous independent works by Barrow et al. (1997) who also estimate the primordial magnetic field at the level of few nG: $B < 3.4 \times 10^{-9} (\Omega_0 h_{50}^2)^{1/2} f^{1/2}$ G where f is a factor constrained as $0.6 < f < 2.2$. More recently Chen et al. (2004) obtained a 3σ upper limits on the strength of the magnetic field of $B < 15$ nG for vector perturbation power spectrum spectral index $n_B = -5$, and $B < 1.7$ nG for $n_B = -7$. These results therefore could make it difficult to reconcile the explanation of the low-ℓ multipole alignments and north-south power asymmetry within the ranges of magnetic field amplitude considered in Bernui & Hipólito-Ricaldi (2008), since for example it was suggested that in order to accommodate for the power asymmetry present in the WMAP data, one would need a primordial magnetic field of order ~ 15 nG. However we note that since the primordial magnetic field generates non-Gaussianity in the temperature maps (Naselsky & Kim, 2008) possibly the required levels of B could be cross-checked with the future constraints on the primordial non-Gaussianity.

2.2 Power spectrum anomalies and implications for cosmology

2.2.1 Glitches in the power spectrum

Low quadrupole Independently from the power distribution in the CMB maps, there have been detected anomalies in the angular power spectrum of the CMB, some of which were well known since the COBE data release and letter confirmed by the WMAP experiment.

Most notably the apparently low, with respect to the best fit concordance cosmological ΛCDM model, quadrupole value, as derived from the different renditions of the foregrounds

cleaned maps[3], or as estimated from the cut-sky analysis using the pseudo-C_ℓ based methods of reconstruction of the underlying power spectra(eg. Jing & Fang (1994); Hansen et al. (2002); Hivon et al. (2002); Taylor et al. (2007)). The low-ℓ range of the CMB power spectrum is depicted in figure 6.1. Is it easy to calculate that the reported by the WMAP team value of the amplitude of the quadrupole $\Delta T_2^2 = \ell(\ell+1)C_\ell/2\pi = 236\mu\mathrm{K}^2$ is compatible with the best fit ΛCDM power spectrum within 95% CL limits[4], however this value is a subject to slight variations depending on the reconstruction method and data set used, and over years the quoted values ranged from $\Delta T_2^2 = 123\mu\mathrm{K}^2$ to $\Delta T_2^2 = 250\mu\mathrm{K}^2$, which as compared to the theoretical value of $\Delta T_2^2 = 1252\mu\mathrm{K}^2$ is still quite low.

The significance of the large scale anomalies has also been addressed via the two-point correlation functions by Spergel et al. (2003) where based on S statistic:

$$S = \int_{-1}^{0.5} [C(\theta)]^2 \mathrm{d}\cos\theta \tag{2.1}$$

it was found that only about 0.15% of simulations realize a lower value of S than measured in the data, and with respect to the best fit cosmological model, as inferred from the WMAP+CBI+ACBAR+2dFGRS data combined, the CMB quadrupole has been found unusual at the level of 0.7%. More recently Copi et al. (2008) have pointed out that the situation is even exacerbated against the standard ΛCDM model (i.e. in sense of vanishing the correlations over the scales of $\theta > 60^\circ$) when considering the regions away from the galactic plane.

Independently, Efstathiou (2003b) performed a frequentist and Bayesian significance analyses, and based on a standard cosmological fiducial model, arrived at the significance of rejecting the quadrupole as inconsistent with that model no larger than 98.7%. However it seems that the analysis did not attempt to marginalize over the unknowns, resulting from the uncertainties in the cosmological parameters estimations, which given the previously mentioned results seems to be important. Secondly, the analysis was based on the fiducial model with large optical depth to reionization ($\tau = 0.17$), which given the improved, three and then five-year release of the WMAP data, enriched by the polarization power spectra, is currently ruled out. In the latter analysis (Efstathiou, 2004) involving both the pseudo power spectra estimators, and the reconstructed spectra estimators and involving various sky-cuts and data sets the same author has confirmed the previous results estimating the CMB quadrupole being realized in at least few per cent of simulations. More recently, Magueijo & Sorkin (2007) performed an approximated Bayesian analysis of the significance of the quadrupole anomaly of different renditions of the first and the three-year WMAP data, and concluded that the evidence in favour of hypothesis violating the scale invariance of the primordial power spectrum, leading to suppression of low-ℓs in the CMB power spectrum, is insignificant, and does not justify introduction of additional parameter into the standard model, according to the various information criteria (such as Akaike and

[3] See more details on different renditions of the cleaned CMB maps in section 7.3.1.

[4] An estimate inferred from 1000 realizations of the best fit power spectra as found in Dunkley et al. (2008) for the mean likelihood values.

Bayes). However, again, it was not mentioned which particular cosmological model is taken as an reference null hypothesis model. Also there's a possibility that relying on evidence estimates, rather than on explicitly integrated Bayesian evidences, might lead to a so-called posterior bias, since the information criteria base only on the maximum likelihood values and do not account for the full distribution shape.

Although the evidence for the anomalous quadrupole is still only tentative, rather than decisive, a hindering promise of an exciting new window onto the plausible new physics responsible for the generation of the observed suppression, inspired a number of possible theoretical explanations. For example Efstathiou (2003a) proposed a slightly closed Universe ($\Omega_k \approx 0.05$) with corresponding curvature radius $R_c = c/H_0|\Omega_k|^{-1/2}$ and an "admittedly speculative" exponential suppression in the primordial power spectrum at the curvature scale. In this model the discrete spectrum of perturbation wave numbers k, related to the discrete spectrum of the eigenvalues β of the Laplacian in curved space via $\beta^2 = (1 + k^2 R_c^2)$, is truncated at the curvature scale, leading to the apparent power suppression at the horizon scales. However as it was pointed out, this model suffers at least two problems. One related to the fine tuning of the primordial power spectrum suppression to be somehow related to the curvature scale of the Universe today, and more importantly, in models with even slight - at the level of now days measurement precision - positive curvature, it is difficult to explain the generation of perturbations via standard inflationary mechanism even for reduced number of e-folds, as the perturbation spectra will become too red, with perturbations amplitude of order $\mathcal{O}(1)$ at the close horizon scales (Linde, 2003). Consequently, contrary to the standard models of eg. chaotic inflation with potential $V \sim m^2\phi^2/2$, in which typical number of e-folds is of order 10^{12}, the model with small number of e-folds (of order ~ 60) would have become anisotropic, contrary to already verified CMB observations, and would also suffer a very strong fine tuning of the exact number of e-folds to fit into the requested curvature of the today's Universe. Therefore the prediction of flatness of the Universe is still a very important observable strongly constraining theoretical models. Once, and if, the future experiments allow to reach the accuracy allowing to rule out the flatness of space at high significance, it will automatically rule out whole families of standard inflationary models, and pose a serious challenge to explain the perturbations generation within the inflationary paradigm, matching the already constrained properties of the primordial power spectrum, *and* maintain the non-zero spatial curvature. Unfortunately the CMB data alone constrain the curvature of space only to within about 10% accuracy, and this implies a need for using external priors from other cosmological data-sets that might be a subject to systematical biases (like de-reddening dust corrections in the supernova data or biases due to gray extinction on large dust grains). However only a combination of the CMB with estimates of the Hubble constant will eventually allow reach sufficient accuracy to possibly push the curvature scale far outside the now days casual horizon scale and rule out these possibilities.

Instead, explanations involving a non-trivial topology of space have become popu-

lar Aurich (1999a); Lehoucq et al. (1999); Inoue (1999); Levin et al. (1998); Cornish et al. (1996); de Oliveira-Costa et al. (1996); Stevens et al. (1993); Aurich (1999b); Starkman (1998); Uzan (1998); Cornish et al. (1997); Luminet & Roukema (1999); Cornish & Spergel (1999); Roukema et al. (2004); Weeks et al. (2003); Niarchou & Jaffe (2007, 2006); Kunz et al. (2008); Weeks & Gundermann (2006); Gausmann et al. (2001); Riazuelo et al. (2004b); Aurich et al. (2004); Luminet et al. (2003); Mota et al. (2004); Weeks (2003); Gomero et al. (2002a); Gomero & Rebouças (2003); Riazuelo et al. (2004a), in explaining the apparent $\ell = 2$ suppression, and number of applied tests have been devised and applied to the CMB, first with the COBE data, then WMAP data, yet with somewhat unsuccessful, null only results (Cornish et al., 1998b; Roukema, 2000; de Oliveira-Costa et al., 2004; Spergel et al., 2003; Roukema et al., 2004; Phillips & Kogut, 2006; Riazuelo et al., 2006; Key et al., 2007; Caillerie et al., 2007; Lew & Roukema, 2008). Attempts were also made to unveil the shape of space via cosmic crystallography of compact spaces in different spatial curvatures by exploring the quasars and galaxy surveys and utilizing observations in frequencies from ratio to X-rays Roukema & Blanlœil (1998a); Uzan et al. (1999b); Roukema & Blanlœil (1998b); Uzan et al. (1999a); Gomero et al. (2002b, 2000); Weatherley et al. (2003), although the scales probed by these data-sets (ranging from 10 Mpc to 100 Mpc) would imply a rather small compact Universe, and naturally could not be related the to suppression detected in the CMB.

However as pointed out in Uzan et al. (2003) the curvature of space, if once determined to differ at the level of a few per cent from unity, will have a serious consequences on the dynamics of the early Universe. In particular, it will pose a serious constraint on the number of possible e-folds that might have taken place before the current horizon scales left the horizon during inflation (problem related to horizon problem and the unwanted relic (moduli) problems) (Lyth & Stewart, 1996). A possible workarounds were suggested via temporal violations of the slow roll conditions by making the inflaton field potential very steep at the onset of inflation, in order to help keep the level of primordial perturbations at largest scales consistent with observed limits.

Tests of multi-connectedness of space may be a very interesting measure of cosmological models by utilizing the so-called identified circles principle with the CMB data under a proposed multiply-connected model. If a significant correlations were found, a curvature associated with that model would be preferred, which could have serious implications on inflationary scenario. However given the infinite number of possible non-flat manifolds, and that the general search, utilizing the identified circles principle is extremely cpu-intensive, as well as that the available high resolution full sky CMB data is drowned in the instrumental noise at small scales, while the topological signature signals additionally obscured by acoustic oscillations scale physics, even in case if the Universe actually *was* multiply-connected, the detection of this fact would be challenging. In chapter 8 we perform a topological test of the CMB data for one previously proposed topological model.

An interesting conception was also put forth by Contaldi et al. (2003) and Germani & Liguori (2008) where models with large scale power suppression were studied. In the

earlier work it was suggested that it is relatively straight forward to obtain the required features in the CMB spectrum, as was previously mentioned, by manipulation of the scalar field potential during inflation. Since the amplitude of the primordial metric perturbations Φ at the horizon crossing ($k \approx a(t)H_\star(t)$) within the slow-roll approximation is:

$$k^3\Phi(k) \sim \frac{V(\phi)^{3/2}}{V'(\phi)} \sim \frac{H_\star^2}{\dot{\phi}} \tag{2.2}$$

inversely proportional to the speed of the scalar field roll, where $H_\star$ denotes the Hubble constant during inflation. It is easy to see that the increase of the potential derivative ($V'(\phi)$) will lead to a desirable suppression of the corresponding k modes, because with such temporarily increased speed of inflation, the relevant perturbations are faster being removed outside the horizon before the fluctuations manage to grow. This is possible to obtain using even a generic polynomial potentials as was discussed in detail in Hodges et al. (1990). It was explicitly shown that this model actually suppresses the large scales, and the fit to the WMAP data has been found. In this model however, the rapid cut-off in the power spectrum is directly followed by a series of oscillations about the standard power-law power spectrum, which are quickly decaying for larger k/h mode values: i.e. for smaller comoving scales. Interestingly such feature has actually been already found by Shafieloo & Souradeep (2007), who directly reconstructed the primordial power spectrum $\mathcal{P}(k)$, using a deconvolution method, directly out of the CMB power spectrum convolved via:

$$C_\ell = 4\pi \int \frac{\mathrm{d}k}{k} \mathcal{P}(k)\Delta_\ell^2(k) \tag{2.3}$$

where the $\Delta_\ell^2(k)$ is the radiation transfer function. A sharp horizon-scale cut-off was found, followed by an increase of power - possibly related to the aforementioned oscillations. This feature has then been studied by Sinha & Souradeep (2006) via number of previously proposed theoretical models. Of course the model with a cut-off requires tuning of the potential to the scales at which the suppression is observed. These happen to be of order of the current horizon scale and the corresponding cut-off scale in the primordial spectrum, and using the WMAP data has been constrained to be $k_c = 4.9^{1.3}_{-1.6} \times 10^{-4}\mathrm{Mpc}^{-1}$ at 68% CL. However, it was also mentioned that the need for the cut-off (and the penalty of the related additional parameter in the standard model) is not highly preferred (of order 2σ). This was also realized by Magueijo & Sorkin (2007) using Bayesian analysis.

However, interestingly it was also pointed out, that the observation of the suppression at the scales comparable to the horizon scale today ($\sim H^{-1}$) as well as the onset of the domination of the cosmological constant energy density is an intriguing coincidence apart from an independent coincidence with the matter energy density, which is also known as the cosmological "coincidence problem". This motivated yet another possible explanation of the low quadrupole problem: i.e. by the effective interplay and, partial cancellation between the standard Sachs-Wolfe contribution (Sachs & Wolfe, 1967) due to the gravitational potential perturbations Φ: $\Delta T/T(\hat{\mathbf{n}}) \sim \Phi(\hat{\mathbf{n}})/3$ with the scale invariant power spectrum of $k^3|\Phi(k)|^2 = A^2$, and the integrated Sachs-Wolfe contribution $\Delta T/T = -2\int_{\eta_{SLS}}^{\eta_0} \Phi(\eta)'\mathrm{d}\eta$

(Contaldi et al., 2003) accumulated long the line of sight from the last scattering (η_{SLS}) until now, where η_0 is the conformal time today (this is also referred to as the *late* integrated Sach-Wolfe effect).

Outliers and "dents" in the angular power spectrum Since the first release of the WMAP data, there have been known few tentative anomalies in the CMB power spectra, which correspond either to outliers, having large χ^2 per degree of freedom value (Hinshaw et al., 2007), or extended local "dents", found eg. around multipole $\ell = 180$ or $\ell = 113$. The latter case of "dent" were analysed by the WMAP team without conclusive results and it was mentioned that similar features also appear in Monte-Carlo simulations (Peiris et al., 2003), while Hansen et al. (2004a) found some tentative dependence of the feature on the analyzed hemisphere as discussed earlier in section 2.1. The earlier case of outliers, being localized at multipoles $\ell = 21$ and $\ell = 39$, was also analyzed in the aforementioned work of Hansen et al. (2004a) as well as in eg. in Covi et al. (2006) and also in Peiris et al. (2003). In the latter works such feature was explained and fitted, within models involving discontinuities in the inflaton potential. These discontinuities are associated with, a well physically motivated models of spontaneous symmetry breaking during the inflationary phase, during which, given the step-like nature of the feature, the inflation is not interrupted, but instead an oscillatory feature is imprinted on the spectrum of primordial perturbations at the corresponding scale (Adams et al., 2001; Adams & Cresswell, 2003). Within this model the potential $V(\phi)$ of the inflaton field ϕ, with the step was parametrized by three parameters: $\phi_{\rm step}$,c and d, corresponding to step location, amplitude, and gradient respectively, as

$$V(\phi) = \frac{1}{2}m^2\phi^2\big[1 + c\tanh\big(\frac{\phi - \phi_{\rm step}}{d}\big)\big] \tag{2.4}$$

where m is the mass of the inflaton. It was shown that (Adams et al., 2001) the WMAP data, as well as the large galaxy surveys, place already strong constraints on such step-like features in the inflaton potential via the size of the oscillations arising in the primordial spectrum that can contradict these observations. It was pointed out that such features are neither unexpected, due to frequent symmetry breaking phase transitions in fields coupled to the inflaton and related to a sudden mass change, nor they require any fine tuning, since many symmetry breaking (every 10-15 e-folds in this model) are expected, based on some arguments from supergravity theories. In particular in the aforementioned work Covi et al. (2006) fitted the data with the χ^2 improvement by about $\Delta\chi^2 \sim 5$ over the "vanilla" simplest flat ΛCDM model. Similar conclusions were obtained in Peiris et al. (2003). Although the CMB spectra is better fitted within the toy model with the step, it was noted that the WMAP data do not really require such model in terms of Bayesian analysis, however it is an intriguing possibility for measuring the processes of the very early Universe, once when the error bars of the measured data will allow to strongly discriminate between the models.

2.2.2 Power spectrum anomalies and induced symmetries in the CMB sky

As a last note in the list of known anomalies in the CMB power spectrum, but still somewhat incomplete list, we mention a tentative oddity in the low-ℓ range of the angular power spectrum, which was previously studied by Land & Magueijo (2005c). It appears that the reconstructed spectrum in that range exhibits an unusual suppression of even multipoles with respect to the odd multipoles (figure 6.1). As will be discussed in chapter 6 such suppression, if true, would be very difficult to explain on grounds of some residual foregrounds contamination, as it introduces a point-symmetries in the CMB sky (see. Appendix A-4), wile the residual foregrounds would rather induce a mirror like symmetries about the galactic plane which in turn would affect both the even and odd multipoles. Nevertheless the mentioned work quantifies the significance of the effect at about $< 97\%$ CL, maximized at cumulative multipole number of $\ell_{\max} = 19$. We will pursue an alternative statistic in chapter 6 to independently quantify the oddity of these features.

Certain plane symmetries in the sky have also been sought in de Oliveira-Costa et al. (2004) using a difference statistic of:

$$S(\hat{\mathbf{n}}_k) = \left\langle \left[\frac{\Delta T}{T}(\hat{\mathbf{n}}_i) - \frac{\Delta T}{T}(\hat{\mathbf{n}}_j) \right]^2 \right\rangle_{N_{\rm pix}} \tag{2.5}$$

where the $\hat{\mathbf{n}}_j$ direction is defined as $\hat{\mathbf{n}}_j = \hat{\mathbf{n}}_i - 2(\hat{\mathbf{n}}_i \hat{\mathbf{n}}_k)\hat{\mathbf{n}}_i$ and the $\hat{\mathbf{n}}_k$ is the direction perpendicular to the plane with respect to which the symmetry is being measured. The average is taken over all directions (pixels). It is interesting to note a great similarities between such derived $S(\hat{\mathbf{n}}_k)$ maps in that work for quadrupole and octupole, and results presented in our analysis (figure 7.5) but derived from completely different statistic: i.e. the statistic probing the percentage of power accumulated by a given $a_{\ell m}$ mode towards a given direction. The similarity shows that in fact for the orientation where there is a significant S symmetry, the amount of power that could be accounted to a_{21} coefficient of the spherical harmonics analysis, is small, which can be perfectly understood since that the shape of Y_{21} harmonic exhibits an exact plane antisymmetry, and hence these harmonics for large symmetry value directions must be cancelled out.

2.3 Large scale multipole alignment anomalies

One of the most outstanding apparent anomalies in the CMB sky are the strong alignments between the quadrupole and octupole components of the full sky cleaned maps. This has been first realized by de Oliveira-Costa et al. (2004) and quantified as an effect occurring roughly once in every 60 GRF realizations. It was also independently measured that the low quadrupole value is in odds as 1 in 20 as is the, yet another, low-ℓ multipole anomaly: the planar octupole (figure 2.2).

Indeed, interestingly the octupole of the WMAP data seems extremely planar: i.e. having apparently most of the power aligned within a plane in all renditions of the WMAP data as showed in figure 2.2. Apart from the low quadrupole value, the planarity of the octupole

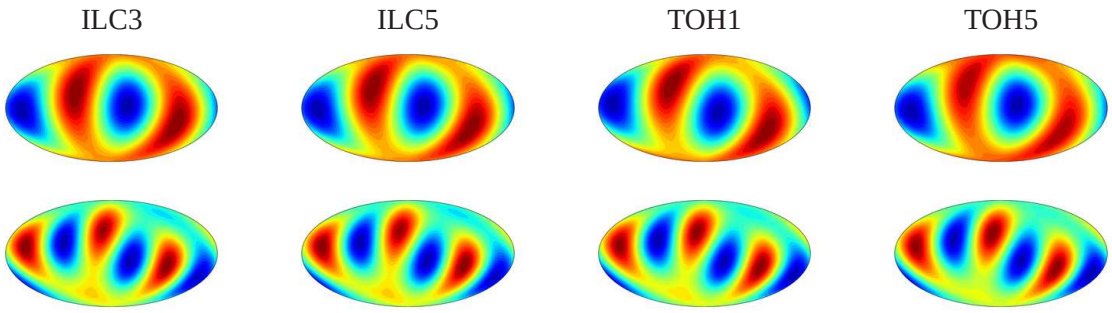


Figure 2.2: Quadrupoles (top row) and octupoles (bottom row) reconstructed from three and five year version of the ILC map (first two from the left) and from first year and five year version of the TOH cleaned map by Max Tegmark (first two from the right) (see section 7.3.1). Note slight differences in quadrupole orientation in the TOH data due to different renditions of different year data sets, which is not that much apparent in case of the ILC data. Note also the apparent alignment between the quadrupole and octupole and the planarity of octupole.

fed the speculations about the non-trivial topology of space, with finite dimension oriented at the direction of the apparent suppression of power in the octupole, coinciding with the preferred maximum-momentum axis, derived via statistic defined in equation 7.1 applied to the quadrupole and octupole independently, which orientations, as already noted, also mutually well coincide. Consequently the direction, roughly estimated at $(l, b) = (260^\circ, 60^\circ)$ in Galactic coordinates, which also roughly coincides with the kinetic CMB dipole direction, and is roughly perpendicular to the direction of the north ecliptic pole, was dubbed the "axis of evil". While the suppression via planarity is well confirmed, no significant evidence was found in favour of the compact (toroidal as originally proposed) topology.

These, and other anomalies, in particular in phase space, have been also investigated by many authors such as: Copi et al. (2006a); Bielewicz et al. (2004); Abramo et al. (2006); Bielewicz et al. (2004, 2005); Land & Magueijo (2005d); Copi et al. (2006a,b); Jaffe et al. (2005), and various and sometimes contradictory statements, regarding the relation of the apparent alignments with the residual galactic foregrounds, have been favoured. Nevertheless, regardless of their origin, the matter of fact consistently remains, that these features are robust and continue to exist since the first releases of the COBE data.

Insofar the origin of the $\ell = 2, 3$ alignments with the ecliptic plane as revealed in the multipole vector analysis by Copi et al. (2006a,b) remains a tantalizing puzzle.

It remains interesting though to try to explain these features in terms of some physical model that would decrease the level of the anomaly from the alarming significance of $\sim 99\%$ to a more acceptable levels. In this spirit Inoue & Silk (2006, 2007) proposed a local voids as a possible explanation of their origin. They argued that two, dust-filled voids, at redshift $z \lesssim 1$ and with radius of $\sim 300h^{-1}$Mpc, separated by roughly 50° apart from the direction $(l, b) = (-30^\circ, -30^\circ)$ along the meridian in a reference frame that maximizes the angular momentum of the low-ℓ multipoles, and of the density contrast $\delta \sim -0.3$, can explain the strong quadrupole-octupole alignments as they would contribute to the temper-

ature anisotropy at the level of $\Delta T/T \sim -10^{-5}$. Such an locally inhomogeneous Universe would lead to a local spatial variation of the Hubble parameter of order of few to several per cent. Future, precise measurements of the Hubble constant as well as large scale structure will put constrains on these possibilities. It was also noted that the chance of finding such a large void is rather unlikely (not mentioning two of them) however is was suggested that the problem could possibly be circumvented via the percolation of many smaller (of order $10h^{-1}$Mpc) voids into a larger void. This scenario was also proposed to help explain the non-Gaussian cold spot (see the next section) and also the north-south power asymmetry discussed earlier.

Other possibilities of explaining these anomalies included less fashionable models like for example the proposed Bianchi type VIIh model tested by Land & Magueijo (2006), Within the models that yield $\Omega_{\rm tot} = 1$, it was argued that the template could provide with an explanation for both: the low quadrupole and the alignments problem, but the model would however suffer from some need of tuning so as to produce certain anti-correlation between the template, which tends generally to add power to the model rather than subtract it, and the GRF CMB. The template itself was not found to be significantly detected, however apparently it helped in solving the strong anomalies. In a related work (Bridges et al., 2007) a whole family of Bianchi type VIIh models were studied in and MCMC parameter estimation search using the WMAP data, and it was found that the only viable solution included the low density cosmology with $\Omega_{\rm tot} \approx 0.43$ and the Hubble parameter of $h = 0.32$, which is currently a disfavoured scenario, most prominently due to the supernovae observations and possibly due to strong detections of the ISW effect in cross-correlations studies with LSS (Giannantonio et al., 2008; Ho et al., 2008; Rassat et al., 2006). This model has also been studied by Jaffe et al. (2006) and subsequently ruled out at high significance level.

Yet another explanation proposal was given by Vale (2005) who suggested that the alignment can arise from weak lensing of the CMB dipole arising in the local neighborhood, in our super cluster, an explanation which was originally motivated by the fact that the preferred axis of evil is also closely aligned with the kinetic dipole of the CMB. The main line of argumentation is that the kinetic dipole, which is two to three-orders of magnitude larger than the primordial perturbations, is weak-lensed and thereby distorted by the inhomogeneous super structures of the local Universe, such as the Great Attractor, of the Shapley super-cluster, and as such, the resulting mixing of the low-ℓ multipoles leads to spill-over of the kinetic signal onto the higher multipoles: most notably the quadrupole and octupole making them aligned and spatially correlated with the dipole orientation. In this context the measures of accounting of for the non-cosmological dipole during the cleaning process are not sufficient to completely remove the kinetic terms from the resulting map. However is was also showed that while this model sufficiently and naturally mitigates the alignment problem, it exacerbates the problem of the low power of quadrupole, as the kinetic, non-cosmological lensing power is added to the low-ℓ multipoles, rather than subtracted, and

hence the level of the primordial fluctuations in $\ell = 2$ should be considered as yet even lower than currently measured. In this context, the low quadrupole science sector becomes even more exciting, however the quantitative assessment of the magnitude of the additional lensed power onto $\ell = 2$ has not been investigated.

There also have been many other speculations on the origins of the alignment problem, ranging from instrumental effects like (gain models of the WMAP, beams of scanning systematical effects), to astrophysical unknown foregrounds related to the ecliptic plane, and finally to the primordial magnetic fields (Bernui & Hipólito-Ricaldi, 2008), capable of generating both: the phase correlations and the hemispherical power asymmetry.

2.4 The non-Gaussian cold spot and other non-Gaussianities

A non-Gaussian feature, extending over an angular size of about 10° towards $(l, b) = (209^\circ, -57^\circ)$ has been detected by many authors via different wavelet decompositions followed by Gaussianity tests of the convolution coefficients (eg. Vielva et al. (2004); Mukherjee & Wang (2004); Cruz et al. (2005, 2007); Pietrobon et al. (2008)). Subsequently, a closer inspection revealed that the cold spot is unusual due to non-Gaussianity at the level of $\lesssim 3\sigma$, that it is a circular-like feature and that has no spectral dependency. Inoue & Silk (2006) then proposed that given a dust-filled void of proper size and under-density would yield an observed temperature deficit, which would also help mitigate the hemispherical power asymmetry, while Cruz et al. (2008) came up with an exotic scenario claiming that the cold spot is actually preferably associated with texture-like topological defect; being an artifact from the super-symmetry phase transitions in the early Universe.

Interestingly, the cold spot was sought in the NVSS radio source 1.4 GHz continuum survey by Rudnick et al. (2007) and a dip in the brightness intensity has been found in the cold spot direction, and a hence a corresponding detection was claimed. This however was recently refuted by Smith & Huterer (2008) as an incorrectly treated striping systematical effect of the NVSS survey.

We will pursue such a general search for non-Gaussianity from the point of view of the real-space oriented statistic in chapter 3.

2.5 Primordial non-Gaussianity

It is currently commonly recognised that the primordial non-Gaussianity will soon become an interesting tool for testing inflationary models (Bartolo et al., 2004). This is because, generically, the single field inflation models predict very small, non-detectable amounts of non-Gaussianity due to negligible self-coupling of the scalar field. This is the case for example with a standard chaotic inflation model (Linde, 1982). However many other models like hybrid inflation models with many scalar fields Linde (1994), models with features, glitches, steps in the potential Wang & Kamionkowski (2000); Contaldi et al. (2003), k-inflation

models with non-standard kinetic terms in the equation of motion Garriga & Mukhanov (1999) or string theory inspired models of inflation with Dirac-Born-Infeld terms in action Silverstein & Tong (2004); Huang et al. (2008), or curvaton scenario models Lyth et al. (2003), as well as some of the ekpyrotic models Koyama et al. (2007); Koyama & Wands (2007) and many others, typically are capable of generating large values of non-Gaussianity of either sign and in general of various shape, classified via k-mode configurations in n-point correlation functions in spherical harmonic space. This of course motivates pursuit of experimental measurement of the primordial non-Gaussianity via estimates of the second-order corrections to the gravitational potential fluctuations imprinted in the cosmic microwave background radiation.

From the observational side, no significant proof for the primordial non-Gaussianity hasn't been insofar found, apart a tentative report by Yadav & Wandelt (2008) who claim a detection of positive non-Gaussianity at a confidence level exceeding 99%. These results however need further confirmations, and of course higher significance level, in order to concretize the inflationary model and/or rule out other competing scenarios of structure formation. A satisfactory confidence level would be for example the level at which today the fact of reionization is confirmed ($> 5\sigma$) Dunkley et al. (2008) and well accepted. Naturally large doze of hopes is given to PLANCK surveyor in regard of detecting the primordial non-Gaussianity signals. However given that the non-Gaussianity is related only with the second, or higher order, effects in the perturbation theory, with the level of anisotropies detected by COBE $\sim 10^{-5}$ the higher-order effects will be very challenging to detect, given all sorts of foreground astrophysical and instrumental effects to deal with.

With currently accessible data the best constraints on the non-Gaussianity gravitational potential quadratic term coupling coefficient $f_{\rm NL}$ yield $-9 < f_{\rm NL}^{\rm local} < 111$ at 95% CL, and $-151 < f_{\rm NL}^{\rm equilateral} < 254$ at 95% CL (Komatsu et al., 2008) using bispectrum method, where "local" and "equilateral" correspond to configurations of models yielding $k_1, k_2 >> k_3$ and $k_1 \sim k_2 \sim k_3$ respectively. It was pointed out that a full-sky, high resolution observations of the soon-to-be-launched PLANCK satellite, given that the secondary effects (like weak lensing, or S-Z effects) are well under control, it will be possible to significantly shrink these limits down to values of order of *few*. This in turn should be sufficient to already rule out many models only due to the measurement of sing of the non-linear coupling parameter.

2.6 Motivation and problem statement

The list of the anomalies, found in the WMAP data and the related research, mentioned in the previous sections, is by no means complete. Given a huge cosmological implications, the studies devoted to (non-)Gaussianity and statistical isotropy, have now become an important, well motivated and advanced science with large predictive power.

There are few very important ways in which the Gaussianity and the statistical isotropy

analyses are going to be useful in cosmology. As is well known, non-Gaussianity is going to become a very important discriminant of the viable inflationary models. In this context the constraints of level and shape of the primordial non-Gaussianity are of main interest. From this standpoint, tests of consistency of the CMB maps with Gaussianity serve as a complementary science, helping to reduce the impact of foregrounds inevitably biasing the estimated level of the primordial non-Gaussianity. As such it is useful to perform these tests both: in multipole space, due to scale dependence of various non-Gaussianity sources, and due to signal-to-noise ratio issues; as well as in real space, due to the need for precise localization of the non-cosmological non-Gaussianity contributions. Statistical isotropy tests are, of course, related since one cannot separate these from the tests of Gaussianity as long as the underlying temperature or polarization template is unknown, as was already discussed.

Following the standard lore of the modern cosmology, such as the assumptions of Gaussianity and statistical isotropy, in recognition of the importance of experimental study of these assumptions we carry out a series of independent statistical tests, aiming at cross-verification, in variety of ways, of some of the well known, tentative and strong anomalies described in this chapter. By devising and implementing new methods and tests we further explore their properties, and verify their significance.

We also aim on testing different statistical methods themselves prior to any cosmological applications. Given an infinite number of statistical tests one could possibly come up with, and also finite amount of work that has already been pursued, there always remain space for generalizations, and improvements. In the following chapters we present some of the results from the projects that has been already finished, and some results from the still ongoing projects, and we suggest a few, possibly-interesting improvements, that could be addressed in the future similar statistical tests.

Chapter 3

Real space tests of the statistical isotropy and Gaussianity of the Wikinson Microwave Anisotropy Probe Cosmic Microwave Background data

The material presented in this chapter was published in Journal of Cosmology and Astroparticle Physics (Lew, 2008b). It has been submitted on 6th of May 2008 and accepted on 9th of July 2008.

3.1 Abstract

We introduce and analyze a method for testing statistical isotropy and Gaussianity and apply it to the Wilkinson Microwave Anisotropy Probe (WMAP) cosmic microwave background (CMB) foreground reduced, temperature maps. We also test cross-channel difference maps to constrain levels of residual foregrounds contamination and systematical uncertainties. We divide the sky into regions of varying size and shape and measure the first four moments of the one-point distribution within these regions, and using their simulated spatial distributions we test the statistical isotropy and Gaussianity hypotheses. By randomly varying orientations of these regions, we sample the underlying CMB field in a new manner, that offers a richer exploration of the data content, and avoids possible biasing due to a single choice of sky division. In our analysis we account for all two-point correlations between different regions and also show the impact on the results when these correlations are neglected. The statistical significance is assessed via comparison with realistic Monte-Carlo simulations.

We find the three-year WMAP maps to agree well with the isotropic, Gaussian random field simulations as probed by regions corresponding to the angular scales ranging from 6° to 30° at 68% confidence level.

We report a strong, anomalous (99.8% CL) dipole "excess" in the V band of the three-year WMAP data and also in the V band of the WMAP five-year data (99.3% CL).

Using our statistic, we notice the large scale hemispherical power asymmetry, and find that it is not highly statistically significant in the WMAP three-year data ($\lesssim 97\%$) at scales $\ell \leq 40$. The significance is even smaller if multipoles up to $\ell = 1024$ are considered ($\sim 90\%$ CL). We give constraints on the amplitude of the previously-proposed CMB dipole modulation field parameter.

We find some hints of foreground contamination in the form of a locally strong, anomalous kurtosis-excess in the Q+V+W co-added map, which however is not significant globally.

We easily detect the residual foregrounds in cross-band difference maps at rms level $\lesssim 7\mu$K (at scales $\gtrsim 6^\circ$) and limit the systematical uncertainties to $\lesssim 1.7\mu$K (at scales $\gtrsim 30^\circ$).

3.2 Introduction

Observational cosmology has established the flat ΛCDM model with nearly scale invariant initial density perturbations as the standard model of modern cosmology (e.g. Riess et al. (2004); Astier et al. (2006); Eisenstein et al. (2005); Cole et al. (2005); Hinshaw et al. (2007); Page et al. (2007); Spergel et al. (2007); Tegmark et al. (2006)). These observations seem consistent with the simplest predictions from inflation theory. Amongst those predictions, one consequence from the cosmological principle, the statistical isotropy (SI), and one generic consequence from inflation theories, the Gaussianity (to leading order) of the cosmic microwave background (CMB) temperature fluctuations, have received a lot of attention with the release of the first year of observations of the WMAP satellite. The relevant statistical analyses either aimed at detecting small amounts of non-Gaussianity (NG), that stems from non-linear effect even within inflation theories (Bartolo et al., 2004), or looked for any anomalous signal that would challenge this standard model.

However, separating SI from Gaussianity is a delicate task when making such a test, since one has to deal with only one realization of the CMB, that is considered in this context to be a random field. SI and NG have been tested in variety of ways and some "anomalies" have been reported. In particular, using tests optimized for SI, in spherical harmonic (SH) phase space (de Oliveira-Costa et al., 2004, 1996) an unusual alignment (98% CL) at low multipoles have been found and confirmed (e.g.Copi et al. (2006a); Land & Magueijo (2005c)). Number of other tests and statistical tools and estimators have been devised and used to constrain SI and/or NG. Among others, these include: bi-polar power spectrum (Hajian & Souradeep, 2006), phase correlations tests (Naselsky et al., 2005), higher order correlations in SH space (bi/tri-spectrum) e.g.(Ferreira et al., 1998; Magueijo & Medeiros, 2004; Cabella et al., 2006, 2005), n-point real space statistics: (Durrer et al., 2000; Gaztañaga et al., 2003; Gaztañaga & Wagg, 2003), morphological estimators (like Minkowski functionals) (Shandarin, 2002; Wu et al., 2001; Park, 2004), multipole vectors (Copi et al.,

2006a; Land & Magueijo, 2005d; Schwarz et al., 2004; Copi et al., 2004) higher order correlation functions (Gaztañaga & Wagg, 2003), phase space statistics (Naselsky et al., 2005; Chiang et al., 2003), wavelet space statistics (McEwen et al., 2006a; Cruz et al., 2007; Vielva et al., 2004; McEwen et al., 2006b), higher criticism statistic (Cayón et al., 2005), pair angular separation histograms (Bernui et al., 2007b) and also various real-space based tests eg: Hansen et al. (2004a); Eriksen et al. (2004a, 2007); Hansen et al. (2004b) In particular a dedicated tests of hemispherical power asymmetry have been reported by many authors and found anomalous at confidence levels ranging from $\sim 2\sigma$ to $\sim 2.6\sigma$ ($95\%\mathrm{CL} \sim 99\%$ CL)

In this work, we measure regional one-point statistics in the WMAP data and in simulations in order to test the SI and Gaussianity hypotheses. We mean to extend and generalize the previous similar works in three ways.

Firstly, we show that the result of the analysis strongly depends on the way in which the sky is partitioned into regions for the subsequent statistics, and we circumvent this problem by relaxing the constraints on the shape and the orientation of a chosen sky pixelization by considering many randomly oriented sky regionalizations. This allows us to avoid a possible bias in such regional analysis that is constrained only to a single choice of pixelization scheme.

Secondly, we relax the constraint on the size of the regions, thereby statistically probing features at different angular scales.

Thirdly, we account for all correlations between different regions, resulting from the well known two-point correlations (or possible higher-order correlations) using multivariate full covariance matrix calculus for more robust estimation of the statistical significance of local departures from Gaussian random field (GRF) simulations.

We will assess the statistical significance of our results in three different manners so as to avoid the standard pitfalls of such an analysis and will rely heavily on realistic simulations to either probe the underlying distributions or to test the sensitivity of our statistic.

The chapter is organized as follows: in Sect. 3.3 we introduce the data sets that are being tested, and provide details of the simulations. In Sect. 3.4 we describe the details of our statistical approach for regional statistics. We then test and illustrate the sensitivity of our statistics via Gaussian and non-Gaussian simulations in Sect. 3.5 before presenting the results in Sect. 3.6 and discussing them in Sect. 3.7. We conclude in Sect. 3.8.

3.3 Data and simulations

For the main analysis in this chapter we use the WMAP three-year foreground reduced temperature maps from differential assemblies (DA) Q1, Q2, V1, V2, and W1, W2, W3, W4, pixelized in the *HEALPIX* sphere pixelization scheme with resolution parameter $N_s = 512$. We co-add them using inverse noise pixel weighting (Eq. 3.1) and form either individual frequency combined maps (Q, V, W) or an overall combined map (Q+V+W) to increase the

signal to noise ratio according to:

$$T_i = \frac{1}{W_i} \sum_{j=j_{st}}^{j_{en}} w_{ji} T_{ji} \tag{3.1}$$

where $W_i = \sum_{j=j_{st}}^{j_{en}} w_{ji}$ and $w_{ji} = N_{ji}/\sigma_{0j}^2$ and σ_0 is the noise rms for a given DA and N_{ij} is the number of observations of the ith pixel for jth DA (Hinshaw et al., 2007). The sum over j iterates the DAs whose maps are co-added (in numbers, $\{2, 2, 4, 8\}$ respectively for Q, V, W and all channels). We will refer to those datasets as Q, V, W and INC (inverse noise co-added map) respectively and define a data set vector $d \in \{\mathrm{Q}, \mathrm{V}, \mathrm{W}, \mathrm{INC}\}$ for further reference.

We also consider a difference maps between different channels to independently test the residual foregrounds and to cross-check with the results obtained from the single band NG analysis. We consider a single band difference maps (e.g. Q1-Q2, V1-V2) as well, since nearly identical frequency difference maps have a negligible amount of CMB or foreground signal[1] and these are used to test the consistency of our white noise realizations against the pre-whitened $1/f$ pink noise of the WMAP data and constrain the systematical uncertainties. Details of this check is given in appendix A-3. We will refer to these maps as QV, QW or VW for cross-band difference maps and Q12, V12 etc. for an individual differential assembly difference maps.

As an extension to the main analysis we also test the five-year WMAP data set from the V channel and refer to it as V5. For this purpose the WMAP five-year simulations are used and preprocessed in the same way as in case of the WMAP three-year data except for the sky-mask, which here we choose to be KQ75.

The residual monopole, measured outside the three-year release of the Kp0 (hereafter the Kp03) sky mask, is removed from each map by temperature shift in real space. The Kp03 sky mask (including galactic region and bright point sources) is applied and no downgrading is performed at this level. We will use $N_{sim} = 10^4$, realistic, full resolution simulations to test our statistics and to assess confidence thresholds (see Appendix A-1 for details and basic tests).

3.4 Directional statistics

If the CMB sky is a realization of a multivariate Gaussian random field (GRF), then statistics of any linear statistical estimator should not deviate from Gaussianity within any arbitrary region in the sky. Otherwise - in case of non-linear estimators - in general deviations from Gaussian statistics are expected, hence MC approach for assessing limits on consistency with Gaussianity is used.

[1] The non-vanishing CMB or foregrounds content, even in the single band differential maps, comes from slight differences in the effective working frequencies of the differential assemblies (DAs) and also from slightly different beam profiles. While in case of the single band difference maps (e.g. Q1-Q2) the residual rms signal is weaker than the noise by more than two orders of magnitude, in case of the different frequency bands (e.g. Q-V) the residual CMB rms signal is about one order of magnitude weaker than the noise.

In order to test the *stationarity* and Gaussianity of the temperature fluctuation hypotheses we use two independent sphere pixelization schemes to define sky divisions and consequently a set of adjacent, continuous regions.

3.4.1 Sky pixelizations

The first pixelization scheme (hereafter referred to as *HP*) is an independent implementation of the *HEALPIX* pixelization scheme (Górski et al., 2005) and its resolution is parametrized by the n_s parameter (Fig. 3.1 top-left). The total number of pixels for a given n_s is $r = 12n_s^2$. We will use three different resolutions (n_s) as specified in Table 3.1.

The second one (hereafter called *LB*) covers the sphere by dividing it along lines of parallels (iso-latitude) and meridians (iso-longitude) to obtain arbitrarily elongated pixels, generally of varying angular sizes (Fig. 3.1 top-right). This results in the total number of pixels $r = N_l N_b$ (where N_l and N_b are the numbers of longitudinal and latitudinal divisions). The three different resolutions used in the analysis defined by these parameters are specified in Table 3.1. Further flexibility is allowed by rotating the polar axis by three randomly chosen Euler angles.

Since there is no reasonable, physical motivation for preferring any particular sky pixelization over another from the standpoint of testing a GRF hypothesis, we consider $N_m = 100$ random orientations for each of the six types of pixelization schemes, which altogether yields 600 different sky pixelizations with a total number of $280\,000$ regions of different shapes and sizes probing different angular scales (Fig. 3.2). We therefore draw the three Euler angles used to define the axis position and pixelization scheme orientation about this axis from a uniform distribution.

All sky pixelizations are subject to the Kp03 (three-year Kp0) galactic/point sources cut which masks $\sim 23\%$ of the sky. In practice there is no lower bound for the size of a region due to its random orientation with respect to the Kp03 sky cut. However for the sake of numerical stability, when computing the inverse covariance matrix (see below), we only consider regions that happen to have $N_{pix} > N_{pixth} = 100$, where N_{pix} refers to the number of pixels of the original $n_s = 512$ map falling into this particular region.

Hereafter we refer to a particular random realization of a pixelization scheme (a random set of regions covering the full sky and merged with Kp03 sky mask) as a *multi-mask*, since it uniquely tags sky regions and allows to pursue statistics exclusively within them (see Fig. 3.1 bottom-left and bottom-right). Of course different *multi-masks*, even defined from a similar pixelization scheme, may have a different number of regions due to the random orientations with respect to the Kp03 sky mask. We define $N_{reg}(r, m)$ as the number of regions of a *multi-mask* as a function of initial resolution parameter r and *multi-mask* ID number $m \in \{1..N_m\}$. As an illustration, the two lowest resolution, pixelization schemes and two examples of *multi-masks* are shown in Fig. 3.1. We will also use additional sets of *multi-masks* to complete and extend the main part of the analysis in a few selected cases.

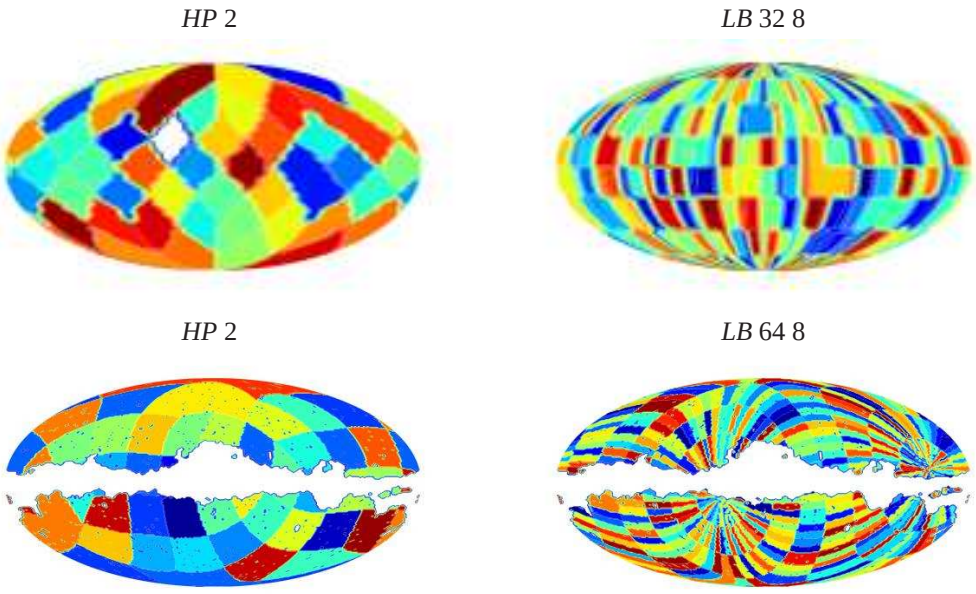


Fig 3.1: In the first row, two lowest-resolution pixelization schemes – *HP* 2 (top-left) and *LB* 32 8 (top-right) are shown. In the second row, we present an examples of two *multi-masks* actually used in analysis. Pixelization schemes are rotated to a random orientation, with Kp03 sky mask applied. These are *HP* 2 (lower-left) and *LB* 64 8 (lower-right) respectively. Values in all regions were randomized for better visualization.

3.4.2 One-point statistics

In each of the defined regions of each *multi-mask*, the first four central moments (i.e. mean (m), standard deviation (σ), skewness (S), and kurtosis (K)), of the underlying temperature fluctuations are computed for the data and for all $N_{\rm sim} = 10^4$ simulations. Together this yields 2.8×10^9 regions for assessment of uncertainties. As will be shown in the Sect. 3.5, allowing for arbitrary orientations of *multi-masks* has an impact on the results and yield a more stringent test on stationarity. The fact that we choose to work in real space allows for a good localization of deviations in the sky.

Table 3.1: Summary on the *LB* and *HP* pixelization schemes and resolutions used in the main analysis, given explicitly for quick reference. The columns abbreviations are as follows: (1) pixelization scheme reference name, (2) resolution parameter value, (3) approximated angular size of regions, (4) number of regions in pixelization scheme.

HP				*LB*					
(1)	(2)	(3)	(4)	(1)	(1)		(3)		(4)
Ref.	Res.	Ang.size [deg]	regs.	Ref.	Res.		Ang.size [deg]		regs.
name	n_s	$\Omega_{\rm reg}$	r	name	N_l	N_b	Δ_l	Δ_b	r
HP 2	2	29.3	48	*LB* 32 8	32	8	11.3	22.5	256
HP 4	4	14.6	192	*LB* 64 8	64	8	5.6	22.5	512
HP 8	8	7.3	768	*LB* 64 16	64	16	5.6	11.3	1024

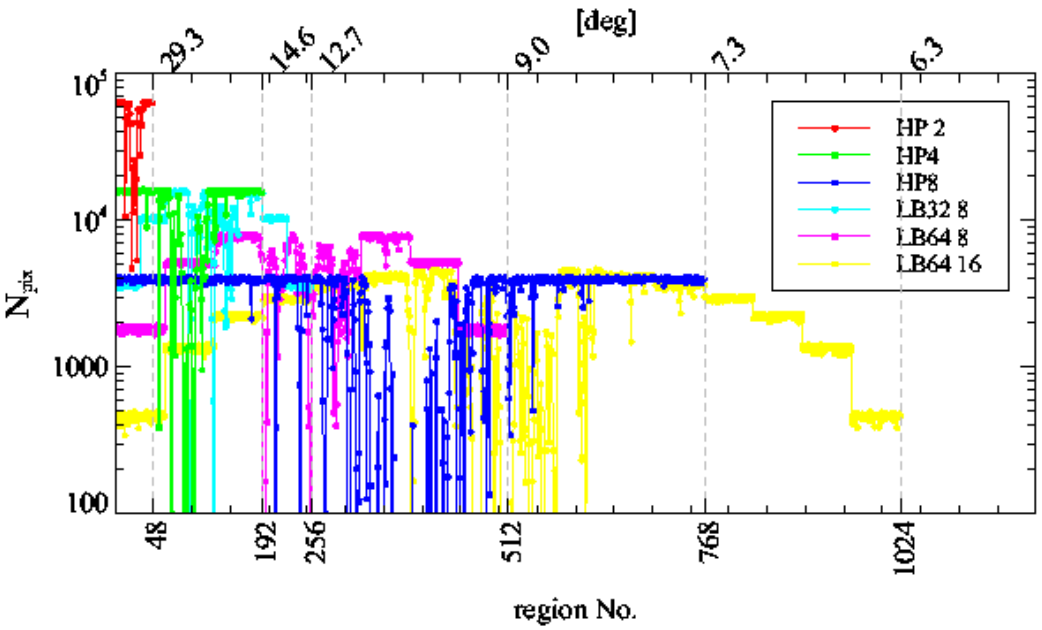


Fig 3.2: Number of points, $N_{pix}(k)$, in regions for all 6 types of pixelization schemes used in their initial position, after masking by Kp03 sky mask, in function of region number. All regions with $N_{pix} < 100$ were not considered in the analysis as detailed in the text and treated as masked. The central parts ($\sim r/2$) of a given pixelization are strongly covered by the Kp03 mask (this is true only for the particular initial orientation of a *multi-mask*). At the top abscissa we give the approximate angular scales probed by pixelization schemes with the corresponding total number of regions indicated in the bottom abscissa.

The presence of extended, residual foregrounds or unremoved, unresolved point sources, will affect the local central moments distributions. In particular the mean of the fluctuations will tend to be up-shifted with respect to simulations if diffuse foregrounds are present or down-shifted if they are over-subtracted. Also, depending on the amplitude of the residual foregrounds the local variance will also be altered. Looking jointly at the distribution of these moments on large scales might also provide a handle on the large scale distribution of power via the off-diagonal terms of the inverse covariance matrix. The physical extent and position of the regions where particular type of deviation occurred can provide a clue to the possible nature of the foregrounds causing it (see Sect. 3.5).

3.4.3 Assessing statistical significance

Since our measurements are statistical, a crucial stage remains in probing their statistical significance. Our approach relies on a detailed comparison between the measurements performed on real data with the distribution of the same measurements performed on simulations.

We consider three different ways to address the significance of these measurements. Each step involves one extra-level of generality and will shed light on the subtleties of such

an assessment.

At first, we look at individual regions, ignore their correlations and compare them with the simulations. We call this approach the "individual region analysis". It is the simplest approach one can consider.

Secondly, we compute the overall statistical significance per *multi-mask*, by taking into account the two-point correlations between moments of distributions (MODs) measured in regions of the same *multi-mask* via the full covariance matrix. We call this approach the "multi-region analysis". The resulting probabilities $P\left(\chi^2_{\mathbf{q}}\right)$ (Eq. A-7) are the joint probabilities of exceeding a certain confidence threshold as a function of pixelization scheme (r), *multi-mask* ($m \in \{1..N_m\}$), MOD ($X \in \{\mathrm{m}, \sigma, \mathrm{S}, \mathrm{K}\}$), and dataset ($d \in \{\mathrm{Q}, \mathrm{V}, \mathrm{W}, \mathrm{INC}\}$) configured by a parameter vector $\mathbf{q} = \{X, r, m, d\}$. This analysis extends the information from the single region analysis by testing the consistency of the data with the simulations via standard multivariate calculus.

Finally, we combine all the information probed by different *multi-masks* to find the joint cumulative probability of rejecting the GRF hypothesis as a function of pixelization scheme (r), MOD (X), and dataset (eg. frequency) (d). We call this approach the "all *multi-masks* analysis".

We remind that the statistical significance of any real data measurement, at any stage of the analysis, is always assessed by a comparison to the set of the same measurements performed using GRF simulations. The exact details of the analysis at each step are given in Appendix A-2.

3.4.4 Visualizing the results

To visualize our results from the single-region analysis, or multi-region analysis at certain confidence level, we proceed the following way. For individual region statistics, for each region of each *multi-mask* we define n_σ as

$$n_\sigma = \sqrt{2}\mathrm{erf}^{-1}(1 - P(X)) = \mathrm{cdf_G}^{-1}(P(X)/2) \tag{3.2}$$

where $P(X)$ is the quantile probability derived according to Eqs. A-3 and A-4. The n_σ thus defined is the Gaussian number of σs by which a region, defined by a given *multi-mask*, deviates from simulation average. We then produce maps of n_σ estimator, for data processed through each of the 600 generated *multi-masks* and for each MOD. Then, to present all the results in a compact way, we scramble these maps within the same MOD. We over-plot the individual pixels from regions with the strongest deviations from the underlying pixels. Positive n_σ values correspond to excessive value of a given MOD in a region, and negative value correspond to its suppression. For clarity, we use a threshold $|n_{\sigma,th}| = 3$ to produce maps with only the strongest (3σ) detections.

For the joint multi-region statistics we produce maps (as detailed above) using only those *multi-masks* that yield $|P(\chi^2_{\mathbf{q}})| \leq P_{th}$ (Eqs. A-3, A-4) revealing detections at the statistical significance $1 - P_{th}$ for a given MOD, *multi-mask* resolution r and dataset d,

i.e. for a given parameter $\mathbf{q}$. Within this notation, single region statistics corresponds to $P_{th} = 1$.

While the "n-sigma" maps are easy to read when looking at distributions of the anomalies at a given local significance, they are unitless and cannot be directly linked with quantities that are physically measured. We therefore also consider a difference maps (Δ maps[2]) of regional departures in individual MODs between datasets and averaged simulation expectation: i.e. for ith region of a given *multi-mask* we plot $\Delta_i = X_i - \langle X_i \rangle_N$, where $\langle\rangle_N$ stands for average over N simulations.

3.5 Tests of the simulation and measurement pipeline

In order to validate the statistical tools introduced above, test the sensitivity and the correctness of the numerical code, we performed a set of experiments using both simulated WMAP CMB data and data with either simulated violation of the large scale statistical isotropy or localized NG features.

3.5.1 Consistency check with GRF simulations

We first test self-consistency by generating 10 additional INC CMB data sets and carry out for each of them the single-region, joint multi-region and all *multi-masks* statistics. As expected, we found that the simulated datasets yield a good consistency with the simulations (at 68% CL).

3.5.2 Sensitivity to local NG

For a statistically isotropic Gaussian process the kurtosis is expected to be exactly equal to $\mathrm{K} = 3$, which translates into a kurtosis excess $\mathrm{KE} = K - 3 = 0$. Violation of either of the assumptions can lead to a positive or negative KE.

At first we simulate what could be a residual component, resulting from subtraction of an non-ideal foreground template, extended over an area of 10° angular, centered at $(l, b) = (50^\circ, 50^\circ)$, whose signature would be a non-vanishing KE. Such residuals must be expected to be small in the foreground-cleaned maps, and they could be either positive – resulting from foregrounds undersubtraction – or negative ones – resulting from foregrounds oversubtraction.

Therefore, the introduced NG component is drawn from a normal distribution with variance σ^2 and with a mean increasing uniformly across the patch when going north-south following Healpix ring ordering inside the spot. From the northern point of the patch down to the southern part of the patch, the mean will shift by $2n\sigma_{\mathrm{CMB}}$.

We introduce in this way a gradient in the noise and thus a non-zero local negative KE (Fig. 3.4) but still preserve a vanishing (within the spot) skewness. The σ value is chosen

[2]We will hereafter refer to maps so produced as Δ maps to make a clear distinction from the difference maps obtained by differentiation of temperature maps from different frequencies (e.g. QV, QW, etc...)

to be 1% of the underlying CMB rms ($\sigma_{\rm CMB}$). To test the sensitivity of our estimator we consider n to be either 1 or 2, above which the NG template starts to be visually noticeable due to edge discontinuities. Note that, with a so defined anomaly, the pixels of the spot that are close to its horizontal diameter will have the least impact on the underlying CMB field distortions.

The choice of the n parameter values corresponds to the NG signals of the rms amplitude $\sim 50\mu$K and $\sim 100\mu$K for $n = 1$ and 2 respectively within the spot. We note that the rms. value within the spot of the same size, in the foregrounds reduced difference VW map of the WMAP3 data, yields $> 100\mu$K depending on the exact location of the spot in the sky, hence our choice of NG signals amplitude aim at detection of relatively small anomalies as compared to the WMAP3 noise specifications.

We find that for a single NG spot of radius 10°, the multi-region analysis does not return any significant detection for $n = 1$ in any of the MODs, but the single region analysis finds the contaminated regions unusual at $n_\sigma \approx 2.5$. In case of $n = 2$ we detected a 3σ local deviation in kurtosis, while in all-*multi-masks* analysis we reject Gaussianity at 99.8% CL (*HP* 8) due to variance distributions[3].

We rerun the test for the same type of NG templates but replicated in 3 disjoint spots at different directions in the sky for the same values of n parameter (Fig. 3.3-a).

The choice of the position and the size of the spots is most relevant to the results presented in Section 3.6. The results of the single region analysis is shown in Fig. 3.3(b)-(d). Note how different *multi-masks* trace the locally introduced anomaly. Depending on the orientation of the *multi-mask* and its regions around the directions of the NG spots, the returned n_σ values differ. In the overall multi-region analysis this naturally leads to a distribution of probabilities which strongly depend on how the features of the map are split and probed by different regions.

Note that some of the *multi-masks* also return an $n_\sigma > 3$ detections even in a template-free regions. It is therefore clear that use of many differently oriented *multi-masks* helps to investigate the statistical significance of local anomalies.

The multi-region analysis in case of $n = 1$ return no significant detections in any of the MODs, but very significant deviations were detected for the case $n = 2$ in all-*multi-masks* analysis (Table 3.2 in section "KE-"), again only in the variance distributions.

Consequently, we find that for the unfiltered maps the distributions of variances are actually more sensitive to this kind of simulated anomalies, rather than higher order MODs. We note, however, that measuring a local sign of KE may be a hint of the nature of the foregrounds signals as the kind of template used in this example introduces locally only the negative KE as shown in Fig. 3.4. Similar dependences are obtained for templates of different shapes and sizes and combinations of these factors. In the limit of a flat template ($n = 0$) the field becomes Gaussian, as expected.

It is possible to introduce a non-vanishing skewness by making the template unsym-

[3]Of course, the estimated rejection confidence thresholds given in Table. 3.2 based only on a single-simulation measurements may be somewhat biased depending on particular realization of the GRF simulation.

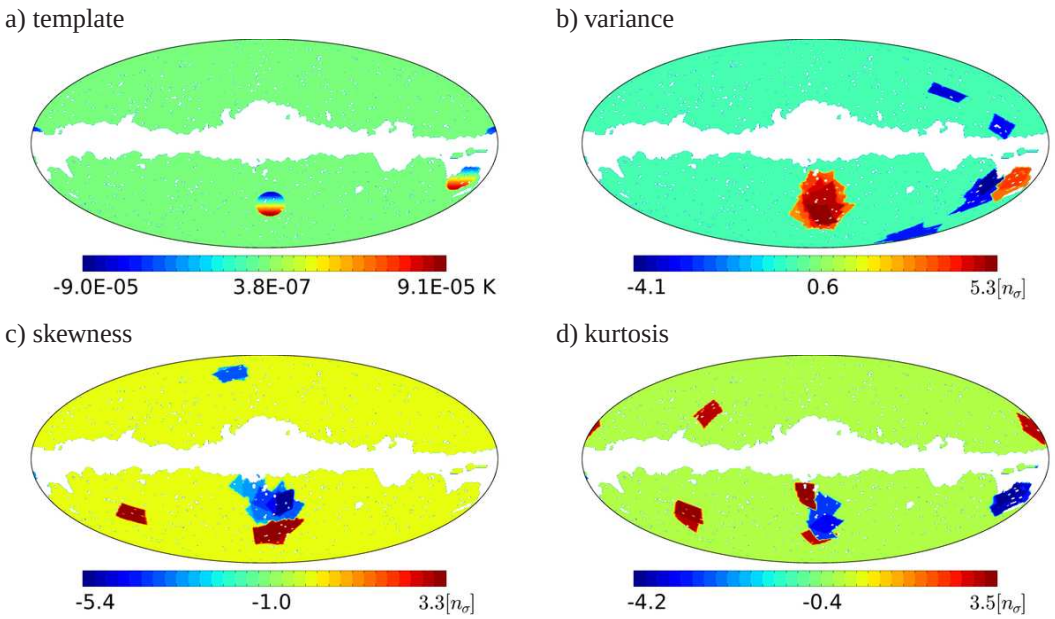


Fig 3.3: *a)* A Non-Gaussian temperature gradient template, leading to locally negative kurtosis excess. The parameter value of $n = 2$ is used (see the main text for details). Note that one of the spots is practically removed by the galactic sky cut. *b-d)* The results of the single-region analysis using the 100 *multi-masks* of the *HP* 4 pixelization scheme of one of the simulated INC data and contaminated with the template. The scrambled n_σ detection map in variance, skewness and kurtosis, thresholded at $n_{\sigma,th} = 3$, is plotted. Note the strong, template-induced, local anomalies detections traced by different *multi-masks*, as well as some other, but somewhat weaker, detections of $|n_\sigma| > 3$ regions. In particular, the template leads to the local kurtosis suppression and strong local excess of the variance.

metrical, by considering shifts in the mean which are unsymmetrical about zero. A similar effect is possible by considering regions of *multi-masks* that only partially overlap with the area of the NG spot (Fig. 3.3c). Note that the example from Fig. 3.4 does not include the effects of the non-uniform noise component which is included in our tests and also that the confidence level contours were derived assuming the Gaussian error statistics; therefore it is not straightforward to extrapolate the strong detections expected from Fig. 3.4 for $n = 2$ onto the full sky, locally templated, signal and noise realizations, subject to subsequent regional statistics.

Note that the assumed size of local deviations is small relative to the full sky one, and hence their global impact is reduced accordingly. Larger, in sense of area, deviations will be of course easier to detect. Also a specific pre-filtering in the spherical harmonic (SH) space, of the data prior to the test may help to expose the most relevant scales to the test. In this test we focused on testing unfiltered maps; therefore, necessarily the strength of the detections must be suppressed.

It is not possible to obtain a locally positive KE with the above-described template. Such

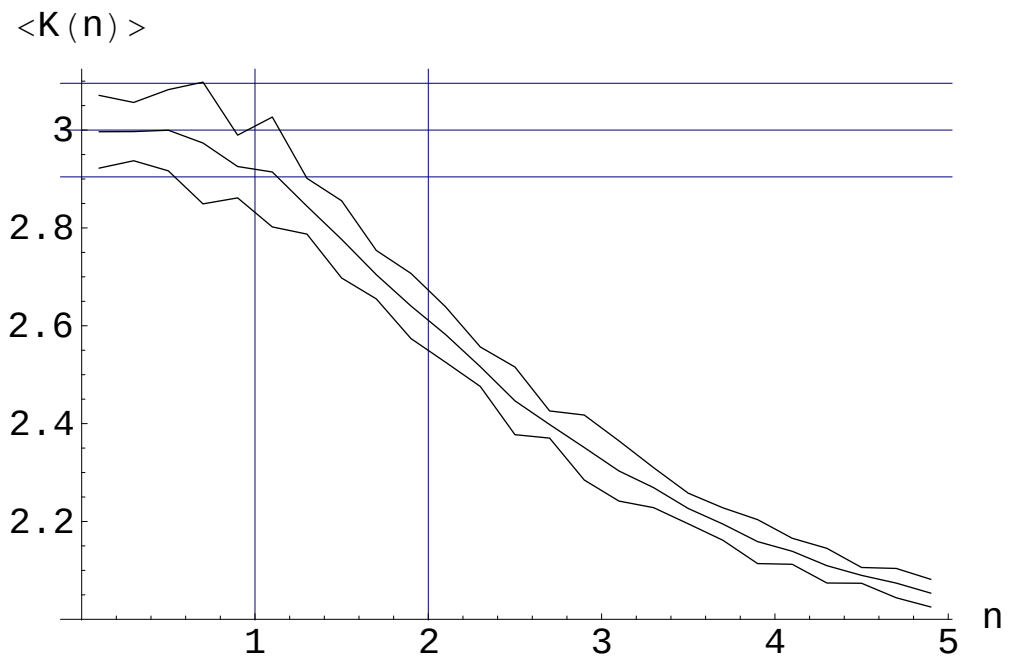


Fig 3.4: Negative departure of the kurtosis in the NG spot as a function of n parameter. The plot shows the kurtosis of a sum of GRF and the generated NG template averaged over 10 random realizations of the two. The variance of the GRF is 100 larger than the variance of the NG template and the field size is chosen to correspond to the NG spot size as described in the text. The 3σ confidence thresholds were quoted around the expected values for the GRF only and for the GRF+NG template cases. Vertical lines indicate the chosen values of the n parameter used for the templated maps generation for the test.

deviation could however be the signature of the unresolved point source contribution[4], or an unknown and localized noise contribution.

Again, for qualitative studies only, we simulate the point source component in the full sky by adding random numbers drawn from a distribution which is an absolute value of the Gaussian distribution with zero mean and random variance $n\sigma_{\rm CMB}$ parametrized by parameter n and uniformly distributed within range $[0, n\sigma_{\rm CMB}]$, where $\sigma_{\rm CMB}$ is the rms value of the underlying CMB fluctuations.

From our simulations however, it appears that it is difficult to detect a significant contribution due to point source contamination since such signal is significantly smeared by the instrumental beam even for n as large as 6. Even when $\mathrm{KE} \gtrsim 6$ before beam smearing, the variance response is much stronger than the KE response leading to inconsistencies with simulations in the total power of the map as measured by e.g. full sky variance distribution (Fig. A-2). We therefore conclude that it is unlikely to detect any point source contribution to KE in this test which is not surprising since we work at fairly low resolution which dilutes the point source signal.

As already mentioned, a locally generated noise-like component in the map, in prin-

[4]Although this would have a specific frequency dependence, we ignore this fact here.

ciple, could generate an non-negligible positive KE[5] as it would not be processed by the instrumental beams. However since such noise is not well motivated physically due to non-local properties of the TOD data and scanning strategy of the WMAP, and also since the noise properties are well constrained, therefore we do not consider such case.

We conclude that small and single (compared to full sky observations) localized NG features will be difficult to detect via higher order MODs in the joint multi-region and all-*multi-masks* analysis due to their small statistical impact on the overall statistics. However a single region statistics carried out first might be a rough guide in selecting a possibly interesting foreground NG signals. If these indicate regions with significant deviations in variance and possessing a negative KE it would hint on residual, large scale foreground contamination.

Stability of results in function number of *multi-masks*

As different *multi-masks* probe the underlying data differently, the joint-probabilities per *multi-mask* differ and lead to a distribution that typically covers a wide range of possible probability values. As such, the multi-region analysis (see sec. A-2.2 for details) can be used to find the orientation of the most unusual regions in the *multi-masks* that yield the smallest probability as comparred to GRF simulations.

Since from the point of view of statistical isotropy all *multi-masks* are equivalent, in the all-*multi-masks* analysis (see sec. A-2.3 for details) we integrate the infomation from all *multi-masks* within a pixelization scheme to obtain an average level of consistency of the data with GRF simulations for that pixelization scheme. This approach also provides a conservative way of averaging over a possibly-spurious detections that could be a fluke, due to some accidental arrengement between a *multi-masks* and a data set. If the anomalous feature in the map is strong enough to be detected in many *multi-masks* then this will also result in a detection in the joint all-*multi-masks* analysis. Conversly, if only one or few *multi-masks* result in a very small probability the overall impact will not be large due to stability of the median estimator with respect to the distribution outliers. However, the investigation of the most anomalous *multi-masks* may help in selecting the deviating regions for further analyses.

In this section we show how the convergence to the results of the all-*multi-masks* analysis is reached in function of number of *multi-masks* used to derive the median χ^2 value and the corresponding median χ^2-distribution leading to the all-*multi-masks* probability. As shown in Fig. 3.5 the convergence of the joint all-*multi-masks* probability to the reported value for $N_m = 100$, is in this case (*HP* 2 pixelization scheme) quite fast; however in general it depends on particular properties of the map as well as on the set and ordering of the *multi-masks* used. The speed of the convergence and the robustness of the final value is as good as the convergence of the unbiased mean estimator - i.e. the median - to the intrinsic mean value as the number of random trials (corresponding to the number of *multi-masks*)

[5] This is most easily seen in the contribution of the anisotropic noise of the WMAP to the kurtosis of the signal only simulated map, which induces a significant positive overall KE value in the simulation.

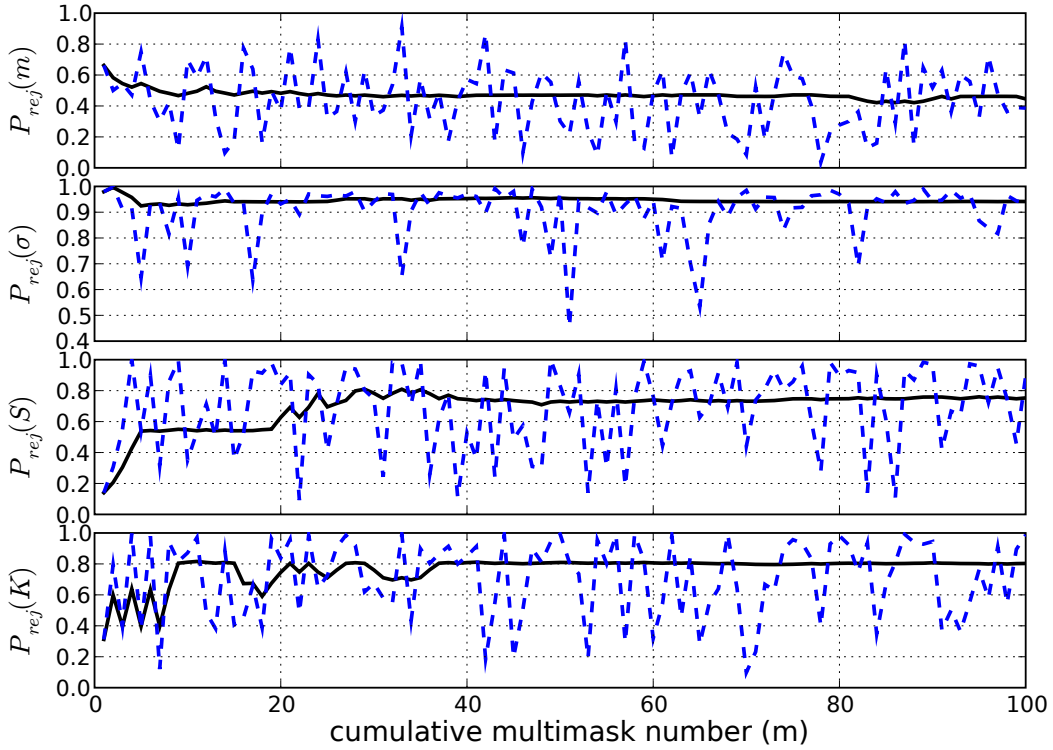


Fig 3.5: Convergence of the joint all *multi-masks* probability to the value reported in Table 3.2 (in section (2KE-) for $n = 2$) for the case of *HP* 2 pixelization scheme in function of the cumulative number of *multi-masks* used to derive it (black solid lines) for mean (top panel), variance, skewness and kurtosis (bottom panel). Additionally we overplot the joint multi-region probabilities per *multi-mask* for each *multi-mask* number (blue dashed lines).

probing the underlying distribution increases.

Note how strongly the joint probability per *multi-mask* depends on the orientation of the *multi-mask* (blue-dashed lines in Fig. 3.5). In particular for the variance case, the probabilities of rejection per *multi-mask* range from 45.7% for *multi-mask* $m = 51$, to 99.8% for *multi-mask* $m = 2$. Yet the, reported in all-*multi-masks* analysis, median value is very stable with respect to these variations.

The shape of *multi-mask* regions.

Naturally, the *multi-masks* having large regions will be insensitive to the small scale map features, while the *multi-masks* having small regions will be insensitive to the large scale structures. This motivates the usage different number of regions to probe different scales of the map.

Theoretically there are infinitely many ways of defining the shape of regions of *multi-masks*, and of course, our choice of the shape of pixelization schemes and the *multi-masks* is somewhat arbitrary; however the motivation for using different shapes of regions is straight-

forward: to probe the data using different binning techniques since the GRF statistically should not depend on it. However if the data turns out to be non-random it is possible that the non-randomness will be explored differently by different region shapes. A loose analogy to the real-space *multi-mask* region shape, which is used to derive a local value of an estimator, over the defined area for a given orientation of *multi-mask*, is in the wavelet space the shape of the mother wavelet, which is used to obtain the local convolution coefficients for the input map. In this analogy the size of the region correspons to the wavelet scale.

3.5.3 Sensitivity to the large scale phase anomalies

We test the sensitivity of our method to the well known large scale anomalies found in the WMAP data: i.e to the aligned and planar low multipoles $\ell = 2$ and $\ell = 3$. In order to test such anomalies we generate two GRF CMB simulations. In both simulations we use the same realization of the power spectrum and phases as in the case of the first GRF simulation (Sect. 3.5.1).

In the first simulation we enforce large scale phase correlation by introducing an "m-preference" in the power distribution in the quadrupole ($\ell = 2$) and octupole ($\ell = 3$). We choose the "sectoral" spherical harmonic coefficient a_{ll}, to carry all the power of the multipoles. In the second simulation we extend this modification up to $\ell = 5$. The GRF signal simulations are rotated to a preferred frame prior the introduction of the planar multipoles. The signal maps are then rotated back to the original orientation so that the maximal momentum axis was located at $(l, b) = (260^\circ, 60^\circ)$ before adding noise.

As shown in Table 3.2 such anomalies have not been significantly detected at tested scales. Although it is expected and observed that considering unfiltered maps (containing all multipoles information mixed together) there will be a little overall impact on the statistics we note a higher sensitivity would be obtained is a prefiltered in SH space data were used.

3.5.4 Sensitivity to the large scale power anomalies

We give a special attention to testing the sensitivity of the method for detecting and quantifying the previously reported large scale anomaly in the power distribution in the sky (Eriksen et al., 2004a, 2007). We create a simulated CMB maps where the CMB signal is modulated according to:

$$\begin{aligned} T(\hat{\mathbf{n}}) &= T_{CMB}(\hat{\mathbf{n}})\big(1 + M(\hat{\mathbf{n}})\big) \\ M(\hat{\mathbf{n}}) &= A_{\ell_{\max}} \hat{\mathbf{n}} \cdot \hat{\mathbf{d}} \end{aligned} \tag{3.3}$$

where $\hat{\mathbf{n}}$ is a unit vector and M is a bipolar modulation field, oriented in direction $\hat{\mathbf{d}} = (225^\circ, -27^\circ)$ with amplitude $A_{\ell_{\max}} \in \{0.114, 0.2\}$ which modulates the CMB component up to the maximal multipole of $\ell_{\max} = 1024$.

The result of the test with such modulated simulation is given in Table 3.2 (in section "M"). As the amplitude of $A = 0.114$ has been previously claimed to be preferred for the CMB data (Eriksen et al., 2007) we process five additional, full resolution V simulations, modulated with this amplitude, to reduce a potential biases from a single draw of a random

simulation, and report (Table 3.2 in section "$\langle M \rangle$") the average rejection thresholds as a function of the pixelization scheme.

We also process five additional modulated simulations, and apply the modulation only to the range of multipoles $\ell \leq 40$, leaving higher multipoles unmodulated, since the aforementioned work operated at much lower resolution.

The test is able to reject the modulation of the CMB of the amplitude $A_{1024} = 0.114$ at a very high confidence level (99.9%) depending on the pixelization scheme. Note however, that this model modulates all scales equally.

Although in principle, the modulation will change the underlying power spectrum at scales where it was applied, we estimate (Appendix A-1.1) that any such effect for the modulation of $A_{1024} = 0.114$ still remains in greatly consistent with the non-modulated simulations' power spectrum, and hence the results given in Table 3.2 do not result from the underlying power spectrum discrepancies.

We are unable to reject the possibility of modulation with such amplitude applied only to the large scales ($\ell \leq 40$). Such modulation is therefore consistent with the GRF field or unnoticed by the test (Table 3.2). However, according to the best-fit ΛCDM model (assuming even a noiseless observation), the multipoles $\ell \leq 40$ carry only about 24% of the map's power. The possible modulation signals at these scales must also be more difficult to constrain as these are dominated by the cosmic variance uncertainty.

To investigate this further, we test simulations with only the large scales being modulated according to $A_{40} = 0.114$ along direction $(l, b) = (225^\circ, -27^\circ)$. We filter out these scales up to $\ell_{\max}= 40$ (using the Kp03 sky cut) in SH space and downgrade the map to $n_s = 64$, and process these using a new set of *multi-masks* of type: *LB* 1 2 - i.e. having only two regions, each covering a hemisphere. We use 96 such *multi-masks* with orientations defined by the centres of pixels of the northern hemisphere in the ring ordering of the Healpix pixelization scheme of resolution $n_s = 4$. We prepare a set of $1\,000$ modulated simulations treated as data and use $1\,000$ independent GRF simulations to test the consistency with SI. We split the GRF simulations into two sets of 500 simulations each, to derive the covariance matrix, and probe the underlying χ^2 distribution.[6] We carry out the multi-region analysis using 96 *multi-masks LB* 1 2, and record the values of minimal probabilities (per modulated simulation) and the corresponding orientation of the *multi-mask*. The spatial distribution of these orientations defines the accuracy to reconstruct the correct intrinsic modulation field orientation at these scales, under cut sky and negligible amounts (at these scales) of noise.

The result is plotted in Fig. 3.6 (top-left). It is easily seen that while the direction is quite correctly found, the dispersion of the directions within even 50% CL contour (Fig. 3.6 top-right) is quite large which precludes a very precise determination of the modulation axis. We find, that statistically $\sim 8\%$ (the probability corresponding to the peak value in the bottom plot in Fig. 3.6) of Gaussian simulations, to which we compare the modulated simulations treated as a data, exhibit a more unusual configurations of hemispherical variance

[6] Note that with the *LB* 1 2 *multi-masks* having only two regions it is not necessary to process a very large number of simulations to assess a good convergence.

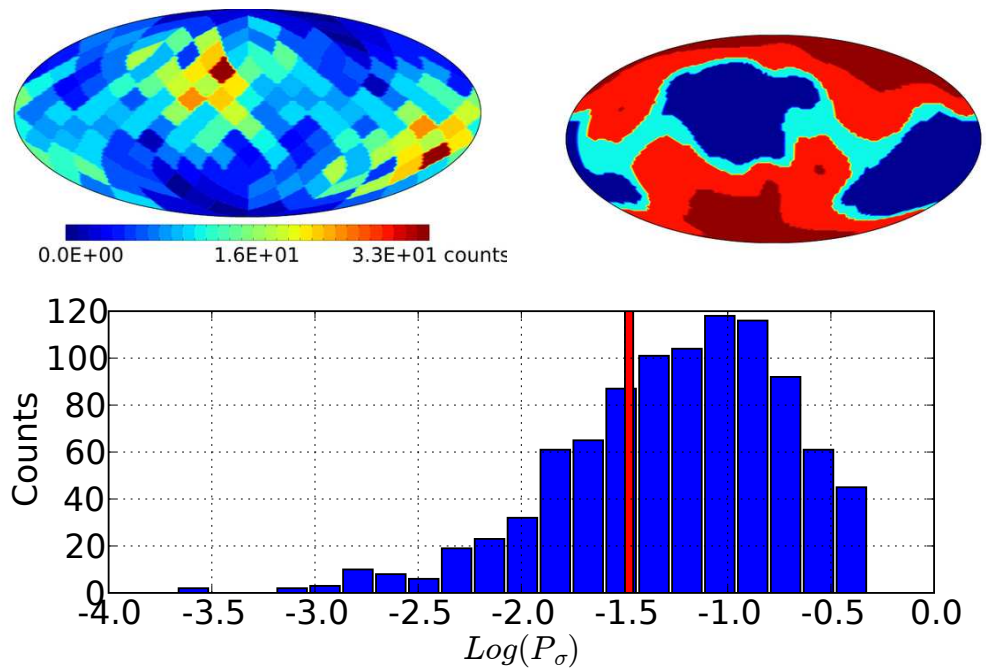


Fig 3.6: Histogram of the reconstructed modulation axis orientations as measured via the minimal joint probability $P\left(\chi^2_{\mathbf{q}}\right)$ (Eq. A-7) in the set of 96 *LB* 1 2 *multi-masks* from 1000 modulated simulations (top-left) visualized using Healpix grid of resolution $n_s = 4$. The corresponding, reconstructed 50% (dark blue), 68% (light blue) and 95% (light red) confidence level contours obtained after smoothing the histogram with a Gaussian beam of FWHM$= 7^\circ$ (top-right). The reconstructed distribution is normalized on a hemisphere. In the bottom plot: the distribution of minimal log-probabilities ($\log P\left(\chi^2_{\mathbf{q}}\right)$) obtained from 1000 simulations modulated with amplitude $A_{40} = 0.114$ along with the value obtained from the V map of the WMAP3 data.

distribution.

Consequently, we report this median rejection threshold of $\sim 92\%$, as the statistical sensitivity of the method for detecting the large scale modulation (i.e. in the filtered maps modulated with $A_{40} = 0.114$). We consider this result - i.e. the low rejection confidence level - to be penalized mostly by the cosmic variance uncertainty and freedom of phases to assume an arbitrary orientations with respect to the (unsymmetrical) sky cut. Consequently, we note that due to these uncertainties, it may be difficult to increase this rejection level for scales of $\ell_{\max}\sim 40$ and amplitude $A = 0.114$.

Table 3.2: All-*multi-masks* analysis results from tests of the simulated datasets. The columns content is as follows: (1) pixelization scheme, (2..4) "probability of rejecting" the consistency with GRF simulations (Sect. A-2.3) for each MOD. The abbreviations of the datasets are: 1KE-: INC simulation with one NG spot (three NG spots) leading to KE< 0 and $n = 1$ (see text Sect. 3.5 for details), 2KE-: INC simulation with one NG spot (three NG spots) leading to KE< 0 and $n = 2$ (see text Sect. 3.5 for details), A2..3: INC simulation with aligned multipoles $\ell = 2$ and $\ell = 3$, A2..5: INC simulation with aligned multipoles from $\ell = 2$ to $\ell = 5$, M: INC dipole modulated simulation with dipole amplitude of $A_{\ell_{\max}}$. $\langle M \rangle$: V simulation with CMB signal fully (partially) modulated according to parameter $A_{\ell_{\max}}$. The average confidence thresholds are given. The values are rounded to integer percentiles in case of probabilities $\leq 99\%$ CL. The saturated values are marked with $\star$.

Reg.sk.	P^m_{rej}[%]	P^σ_{rej}[%]	P^S_{rej}[%]	P^K_{rej}[%]
(1)	(2)	(3)	(4)	(4)
		$n = 1$: 1KE- (1KE-):		
		no significant detections (no significant detections)		
		$n = 2$: 2KE- (2KE-):		
HP 2	42 (45)	82 (94)	58 (75)	49 (80)
HP 4	43 (69)	97 (> 99.98)	31 (66)	39 (62)
HP 8	8 (70)	99.8 (99.8)	5 (32)	23 (41)
LB 32 8	28 (58)	88 (99.0)	29 (72)	27 (53)
LB 64 8	34 (64)	69 (96)	20 (60)	18 (34)
LB 64 16	18 (69)	98 (99.5)	10 (37)	28 (44)
		A2..3 (A2..5):		
		no significant detections (no significant detections)		
		M: $A_{1024} = 0.114$ ($A_{1024} = 0.2$)		
HP 2	39 (36)	94 (99.99)	42 (50)	32 (35)
HP 4	43 (47)	99.6 ($> 99.99^\star$)	22 (25)	31 (32)
HP 8	11 (19)	99.96 ($> 99.99^\star$)	7 (8)	18 (20)
LB 32 8	30 (35)	97 ($> 99.99^\star$)	20 (22)	22 (23)
LB 64 8	34 (38)	91 ($> 99.99^\star$)	15 (17)	14 (14)
LB 64 16	22 (30)	99.4 ($> 99.99^\star$)	12 (13)	23 (25)
		$\langle M \rangle$: $A_{1024} = 0.114$ ($A_{40} = 0.114$)		
HP 2	58 (59)	99 (56)	52 (44)	54 (44)
HP 4	77 (56)	99.87 (57)	54 (37)	50 (38)
HP 8	79 (62)	99.97 (55)	64 (29)	54 (52)
LB 32 8	78 (57)	99 (55)	54 (38)	51 (47)
LB 64 8	78 (58)	97 (55)	59 (38)	54 (43)
LB 64 16	76 (61)	99.7 (54)	64 (32)	55 (46)

3.6 Application to WMAP three-year data

We now present the results of the statistics described in Sect. 3.4 applied to the three-year and five-year WMAP data.

3.6.1 Individual region statistics

The individual region statistics as described in Sect. 3.4 find numerous regions amongst our many pixelization schemes, which deviate by more than certain $|n_{\sigma,th}|$ in all MODs. Table 3.3 gives an incomplete list of some of the strongest detections. In Fig. 3.7 we plot the detected regions at $|n_{\sigma,th}| = 3$ in the individual region statistics of the INC data.

Note that with our approach the so-called "cold spot" is not detected at 3σ level from being cold (i.e. via distribution of means) at all tested resolutions (Table 3.3). It appears at about $\sim 2.7\sigma$ around galactic coordinates $(l, b) = (211^\circ, -57^\circ)$. Excessively "cold" or "hot" deviations in all MODs are detected in general with large values of n_σ.

As expected the n_σ map of the means in Fig. 3.7 shows that the strongest deviating regions are directly close to the galactic plane cut off by the Kp03 mask, thus hinting at foreground residuals.

The variance n_σ map shows local strong anomalies with the extended variance suppression in the northern hemisphere towards $(l, b) = (67^\circ, 19^\circ)$ and with an extended variance excess towards $(l, b) = (199^\circ, -55^\circ)$. We note that these localized anomalies must, at least in a part, make up for the hemispherical power asymmetry.

Skewness and kurtosis maps consistently indicate strong local deviations from GRF simulations towards $(l, b) = (193^\circ, -26^\circ)$ and $(l, b) = (356^\circ, -36^\circ)$. While some regions appear in all three maps, some appear only in one of the moments therefore the correlation between those results is not obvious.

Table 3.3: An incomplete list of strongest deviation directions from maps in Fig 3.7 at $|n_{\sigma,th}| = 3$ in the individual region analysis for means, variances, S and K, sorted in galactic longitude ascending order. Notice that with our simulations we directly probe the $\pm 3.7\sigma$ PDF region without need for extrapolations.

mean		variance		skewness		kurtosis	
(l,b)	n_σ	(l,b)	n_σ	(l,b)	n_σ	(l,b)	n_σ
157, -29	-3.07	63, 28	-4.21	173, -73	-4.50	84, -31	-4.38
211, -57	**-2.78**	67, 19	-4.21	193, -26	5.06	195, -27	5.20
241, 42	-2.36	199, -55	4.13	209, 8	-3.50	**212,-55**	**-3.77**
265, -21	3.81	319, -27	3.40	217, 35	4.00	309, 59	-3.25
318, -9	-3.69			225, -20	-3.60	312, -21	3.65
167, 79	3.06			241, -52	-3.51	356, -36	4.25
				311, -20	3.86		
				357, -35	4.15		

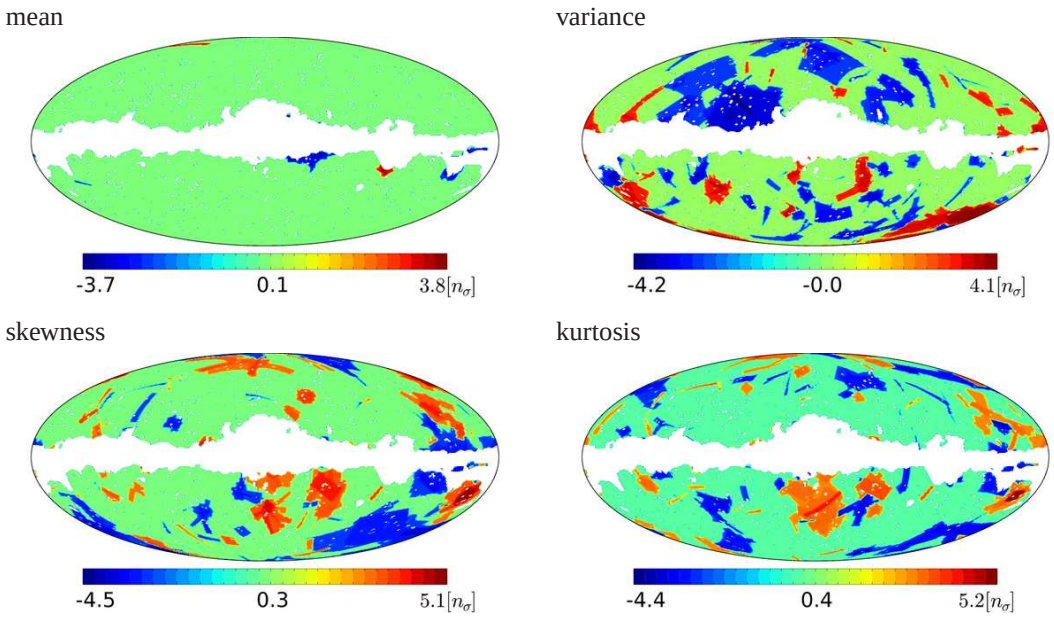


Fig 3.7: Results of the single-region analysis visualized in the composite n_σ maps of the INC data, for each of the MODs. The threshold of $n_{\sigma,th} = 3$ (as defined in Sect. 3.4.4) is used. Mollweide projection and the galactic coordinates are used. The origin of the coordinate system $(l, b) = (0, 0)$ is in center of the plots and the galactic longitude increases leftwards. Regions around the Galactic plane are partially removed by the Kp03 sky mask. The same convention is kept throughout the rest of this chapter.

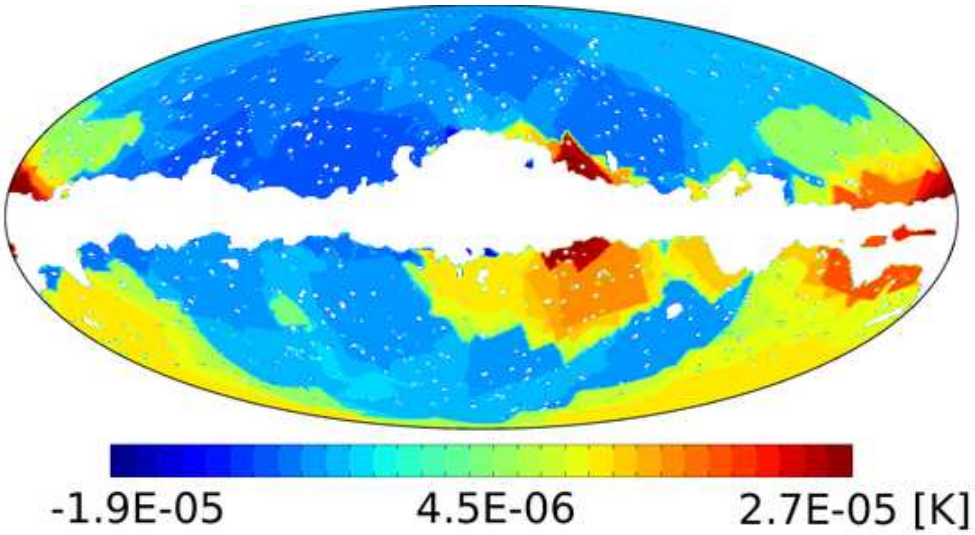


Fig 3.8: Results of the single region analysis. Residual difference ($\Delta = \sqrt{\sigma^2} - \langle\sqrt{\sigma^2}\rangle$) scrambled map of the variance distribution in the INC map processed with 100 *multi-masks* of the *HP* 2 pixelization scheme. The well know, large scale hemispherical power distribution asymmetry is clearly seen as is the distribution of foreground residuals along the galactic cut.

In order to investigate the spatial distribution of this high n_σ regions, we plot in Fig. 3.8 the Δ map, of differences between the variance distribution measured in each region of *multi-masks* of the *HP* 2 pixelization scheme, and the simulations average, $\Delta \equiv d_i = \sqrt{\sigma_i^2} - \langle\sqrt{\sigma_i^2}\rangle$. The map is obtained by scrambling 100 difference maps from 100 different *multi-masks* (as described in Sect. 3.4.4). This map can be seen as a residual map for the local variance. This residual map exhibits a well known power asymmetry (Eriksen et al., 2004a, 2007). The dipole ($\ell = 1$) component of this residual masked map (ignoring the effect of the mask) is aligned along axis $(l, b) = (237°, -44°)$ with power excess in the southern hemisphere. In order to probe this direction further, for each individual *multi-mask* we also produced an n_σ estimator maps and checked the orientation of the dipole axis. Fig. 3.9 shows the PDFs obtained for the orientation of the dipole asymmetry in galactic longitude and latitude. Interestingly, we notice that the orientation of the axis of the hemispherical power asymmetry has some scale dependence. When using smaller scales with finer pixelization schemes, the orientation of the power asymmetry dipole shifts from larger galactic latitudes (roughly from the position of the cold spot, with the mean PDF value $(l, b) = (218°, -43°)$ – see also Table 3.3) for the *HP* 2 resolution to smaller latitudes $(l, b) = (206°, -18°)$ for the *HP* 8 pixelization scheme. The dependence of the power asymmetry orientation in function of the pre-filtered in SH space data have previously been tested by Hansen et al. (2004a). While we will return to the power asymmetry issue in the next subsections, we note that the medians of the dipole axis distributions of other MODs are not correlated with the dipole axis orientation of the variance map (Fig. 3.9) and generally point at some other locations.

In the next section we quantify the statistical significance of these deviations.

3.6.2 Joint multi-region statistics

In Fig. 3.10 we plot a compilation of all joint "probabilities of exceeding", calculated with all datasets considered (Q, V, W, and INC) in all 600 *multi-masks*. In order to visualize the smallest probabilities logarithmic scale is used. Note that we sort these probabilities in each MOD and dataset so as to ease visualization, so the points with same abscissa in different MOD and data sets do not necessarily correspond to the same *multi-mask*.

Most of the results concentrate along the zero point of the joint log-probabilities, which indicates a good consistency of the data with the simulations at relatively high CL. (The white region in the Fig. 3.10 encompasses CLs of up to 3σ; 3σ and 4σ regions are shaded in red and yellow respectively).

It is important to note that within one pixelization scheme the dispersion of probabilities in the Fig. 3.10 results only from the orientation of the *multi-mask*. As a result, the statistical method involving many pixelization schemes help us obtain the unbiased results that one could get relying only using a single pixelization scheme. We also recall that for each plotted point, the statistic was also calculated for $N_{\mathrm{simPDF}(\chi^2_{\mathbf{q}})} = 1\,000$ simulations in order to probe the underlying PDFs. For each point the corresponding full covariance matrix was obtained from $N_{\mathrm{sim}(\mathrm{C}_{\mathbf{q}})} = 9\,000$ simulations (as described in Appendix A-2.2).

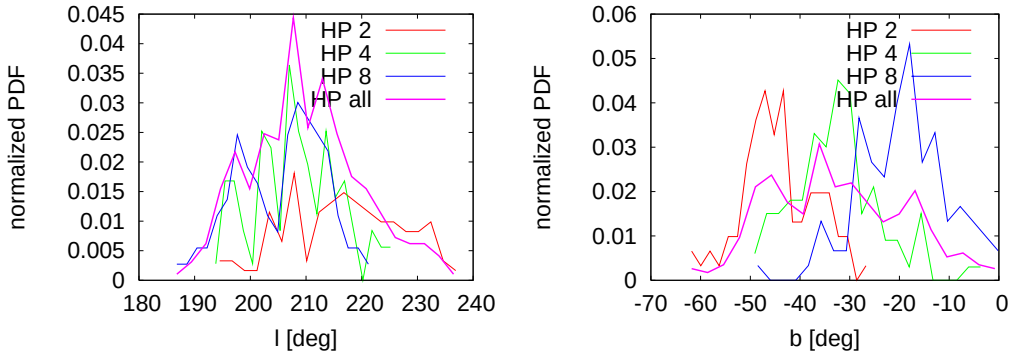


Fig 3.9: Orientation of the dipole component (galactic longitude - left panel and galactic latitude - right panel) of the n_σ maps for the INC map processed individually with all *HP* pixelization schemes. Each color corresponds to a different resolution of the pixelization scheme. While the longitudinal orientation of the dipole does not vary with resolution, the galactic latitude systematically shifts to lower galactic latitudes as resolution increases.

We now focus on three distinctive sets of results, based on the Fig. 3.10, and quantify the deviations in more detail. We detail on a tentative excess seen in kurtosis before focusing on the large scale power anomalies, and on unusually strong dipole contribution in the V channel of the WMAP. Then, we comment on the results from tests carried out with the difference map datasets.

Localized Kurtosis excess

In Fig. 3.10 there is one "3σ" detection in kurtosis in *HP* 4 pixelization scheme in the INC dataset (bottom second from the left panel in Fig. 3.10) – a result found using one in 100 of *multi-masks* probing these scales. Here we discuss this particular point as a tentative detection because although the *multi-mask* bins the data to create the most unlikely realization of the kurtosis, it lies in the low-end tail of the whole spectrum of equivalent measurements and hence its statistical impact cannot be large.

Table 3.4: Three σ NG detections in K multi-region statistics of the INC data using *multi-mask* resulting in joint probability $P\left(\chi^2_{\mathbf{q}}\right) < 0.0027$ for the resolution $n_s = 512$ and $P\left(\chi^2_{\mathbf{q}}\right) < 0.005$ for the resolution $n_s = 64$. The (l,b) field gives the galactic coordinates to the center of the region

	resolution ($n_s = 512$)		resolution ($n_s = 64$)		
region	N_{pix}	n_σ	N_{pix}	n_σ	(l,b)
160	2750	3.48	37	1.38	181 , 2
185	12610	3.21	199	3.26	199 , -23
104	16125	3.52	250	3.27	355 , -44

In Fig. 3.11a we plot, thresholded at "3σ", the n_σ map using only this particular *multi-*

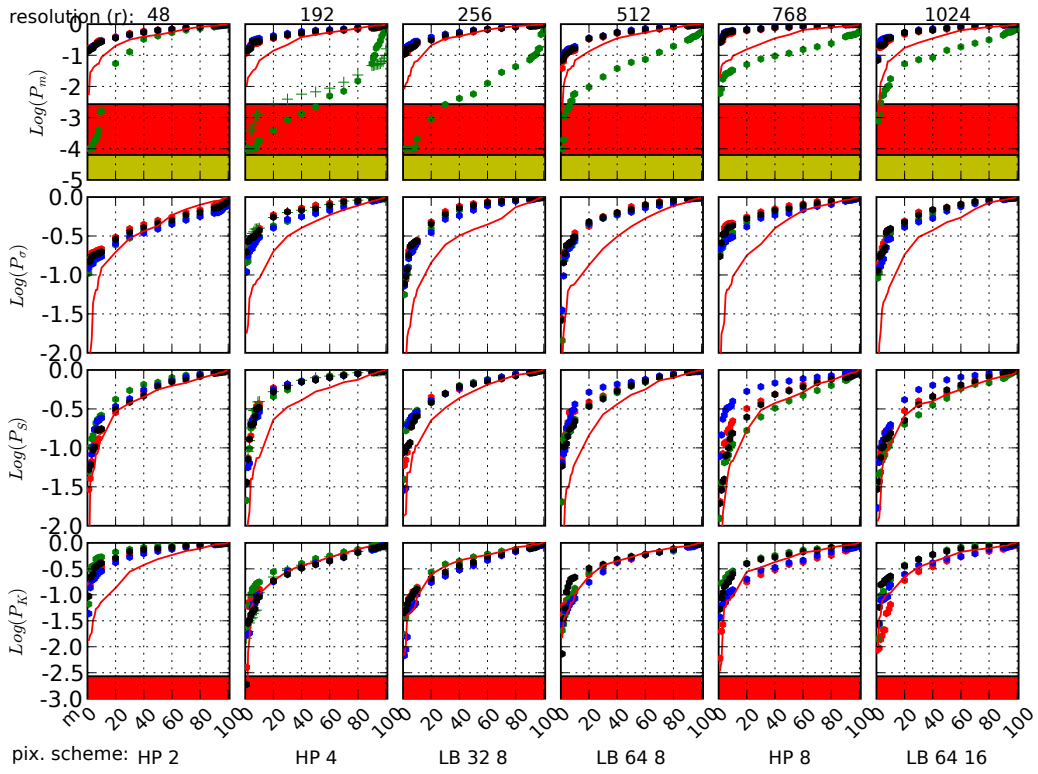


Fig 3.10: Results of the "multi-region" analysis. Dependence of joint probabilities of exceeding ($P\left(\chi^2_{\rm q}\right)$, Eq. A-7) as a function of *multi-mask* number for all pixelization schemes considered. The probabilities calculated with the spectral data maps from channels Q, V, W, INC are plotted in red, green, blue and black dots respectively. Each dot corresponds to the joint probability using one *multi-mask*. From the left to the right, the panels show results for increasing resolution of the pixelization scheme (see Table 3.1) with 100 different *multi-masks* (along abscissa) in each. From top to bottom the four rows correspond to the four MODs - i.e. mean, variance, S and K respectively. Probabilities corresponding to the WMAP5 V5 data for pixelization scheme *HP* 4 are plotted with green crosses ($+$). 3σ and 4σ confidence levels are shaded in red and yellow respectively. The joint probabilities were sorted in each dataset before plotting for better visualization; therefore the probabilities from different datasets generally do not correspond to the same *multi-mask* numbers and do not directly correspond to the unique reference numbers used in the analysis. Hence the most unlikely events are localized at the left side in each panel. Additionally we plot the thin red line, which indicates the distribution of probabilities obtained from 100 GRF simulations of the Q data, each of which was processed with one (different for each simulation) *multi-mask*. If the data follow the expectation of GRF then statistical departure of data from this line would manifest certain degree of correlation between probabilities obtained with different *multi-masks* for the same dataset (as discussed in Sect. 3.7). For better visualization in range $m \in [10, 90]$ we plot only every 10'th sorted probability value.

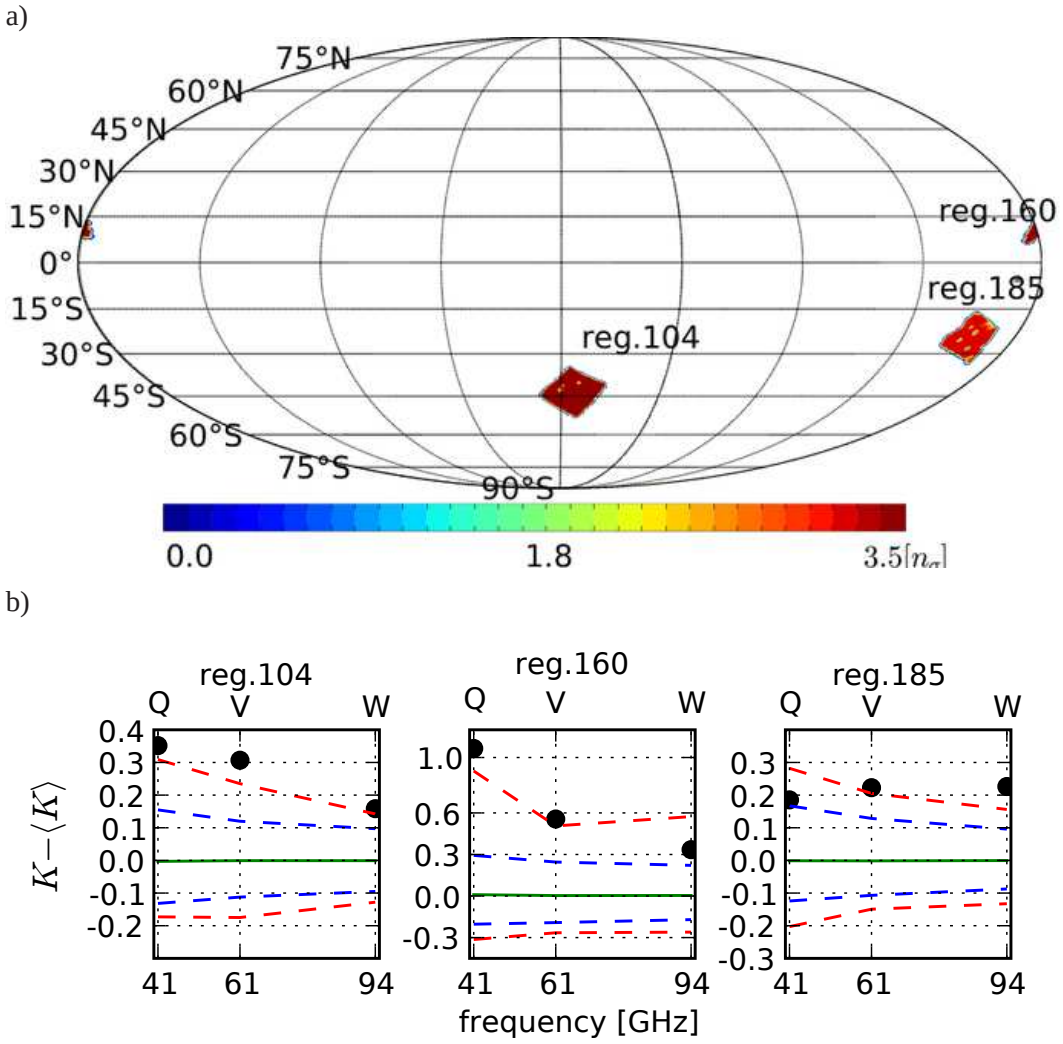


Fig 3.11: a) Kurtosis (thresholded at 3σ) n_σ estimator map from the multi-region analysis of the INC data, inconsistent with GRF hypothesis at joint probability $> 99.73\%$ CL. Only one multimask (*multi-mask no. 53* of the *HP* 4 pixelization scheme) is used for this map, since only the one out of $N_r N_m = 600$ yields $|P(\chi^2_{\mathbf{q}})| < P_{th} = 1 - 0.9973$. (See text and Table 3.4 for more details). b) The spectral dependence of the kurtosis in the depicted regions in the three WMAP frequency bands along with the two and three sigma contours, from 1000 simulations, and the simulations mean are also plotted.

mask. The details of the three most strongly deviating regions in this map are given in Table 3.4. In case of two other *multi-masks*, using which the INC data yield a detection at confidence levels of 99% and 98% in K (*LB* 64 8) and S (*HP* 8) respectively, the "3σ"-deviating regions turn out to be similarly located (like those in Fig. 3.11a). With these regions masked out from the analysis a good consistency with the GRF simulations is reached. Note that the region 160 is rather small – mostly removed by the galactic sky cut (see Fig. 3.11a and Table. 3.4 for precise coordinates and size) however masking it has a comparable effect on the joint probability increase as masking out the two other and much larger regions. The con-

sistency of the INC data with GRF simulations increases from 0.18% (without removal) to 1.1% and 0.9% with regions 160 and 104+185 removed respectively. Individual removal of regions 185 and 104 only increases the level of consistency by $\lesssim 0.5\%$. The simultaneous removal of all three regions increases the consistency up to 12%.

Dependence as a function of *multi-mask* To further test the robustness of this detection we have generated two other *HP* 4 pixelization scheme sets of $N_m = 100$ *multi-mask* each: one by simply choosing the three rotation angles with the prescriptions given earlier, and the other by focusing only on the region in the rotation angle parameter space within $\pm 5^\circ$ around the original orientation of the *multi-mask* leading to the 3σ detection.

With the first set we obtained results yielding a joint probability $P(\chi^2_{\mathbf{q}}) < 0.05$ with 3 *multi-masks*, while in the second we find that 25% of *multi-mask* yield $P(\chi^2_{\mathbf{q}}) < 0.05$, and 4% yield $P(\chi^2_{\mathbf{q}}) < 0.01$, with the strongest detection $P(\chi^2_{\mathbf{q}}) = 0.0035$, of which the n_σ map points to the same three regions as depicted in Fig. 3.11a. We note that the reported regions (160 and 185) are located in directions towards which the strongest deviations in the individual region statistics (Fig. 3.7) were found.

Dependence as a function of frequency and resolution In Fig. 3.11b we present the spectral dependence of kurtosis in the regions depicted in Fig. 3.11a. While there is a non-trivial spectral dependence in regions 160 and 185, with opposite tilt – red and blue respectively – there is almost no spectral dependence in region 104.

We also check the dependence on the S/N ratio in the selected regions. For this purpose we downgrade all datasets and simulations to resolution $n_s = 64$, which effectively increases the S/N ratio per pixel by a factor of 8. We redo the multi-region analysis lowering the minimal region pixel number threshold down to $N_{pix} > 10$ and find that the minimal probability per *multi-mask* ($P\left(\chi^2_{\mathbf{q}}\right)$) corresponds to a rejection threshold of 99.5% CL (Table 3.4). As seen in Table 3.4 the individual region response to the resolution change is strong only in case of region 160 while in the two other regions it is rather small. The result is robust under variations of region pixel number threshold and the number of simulations used to probe the underlying PDFs ($N_{\mathrm{simPDF}(\chi^2_{\mathbf{q}})} \in \{1000, 5000\}$). Masking-out regions 104 and 185 reduces the anomaly to $\sim 96\%$ CL and as expected in this case, removal of region 160 has basically no impact on this value.

Summary The non-trivial spectral dependence and close galactic-plane alignment in two of the three selected regions (160 and 185) suggests presence of some residual foreground anomalies. In case of the regions away from the galactic plane (104 and 185) since the local oddity is insensitive to the S/N ratio change it is also unlikely that an unknown instrumental noise fluke generates them. While we will return to the overall statistical significance of these findings in Sect. 3.6.3, we note that the positive KE seems to be inconsistent with the extended foregrounds interpretation of these detections, according to the results from Sect. 3.5. Also the 3σ detection of the multi-region analysis appears only in the INC data,

but is clearly weaker in other single band maps.

Given that this detection results from just one particular *multi-mask* and is selected from the lower-tail end of a whole distribution of equivalent measurements, it is inconclusive as regards indicating whether this detection is not just a fluke. Given that, we report in particular region 160, whose removal leads the overall significance to drop below 3σ CL, as a tentative detection noting that more sophisticated local statistical analyses (see Sect. 3.7) could be invoked to back these results up or refute them.

Variance large scale distribution

In Sect. 3.6.1 we analyzed the large scale power distribution as measured via *multi-masks* with regions of angular sizes ranging from 6° to 30° (Fig. 3.8 and Fig. 3.9).

In this section we focus on the joint multi-region analysis of the variance distribution. The corresponding results are illustrated in the second row of Fig. 3.10. The data remain in excellent consistency with the simulations.

In Sect. 3.5.4 we found that the modulation amplitude (defined by Eq. 3.3) of $A_{\ell_{\rm max}=1024} = 0.114$ would be rejected at $> 99.9\%$ CL, and we argued that the modulation parameter $A_{\ell_{\rm max}=40} = 0.114$ would statistically be difficult to exclude at CL higher than 92%. Using our main set of the *multi-masks* (Table 3.1) we fail to detect, in any of the data sets tested, any statistically significant anomaly, such as the claimed hemispherical power asymmetry (depicted in Fig. 3.8), as measured from the large scale variance distributions in the multi-region analysis.

As an extension to that, we repeat the analysis performed in Sect. 3.5.4 for the low resolution, filtered in SH space up to $\ell_{\rm max}= 40$, WMAP data, using the same set of *LB* 1 2 *multi-masks*, to test the variance distributions in the corresponding set of 96 differently oriented hemisphere pairs. We thereby extend the test for the largest possible scales of 180°. We merge the Kp03 sky mask with the *LB multi-masks* for the analysis of the WMAP3 data, and KQ75 sky mask for analysis of the WMAP5 data. The result of the multi-region (here only two region) analysis is plotted in Fig. 3.13 for the V data (left) and V5 data (right). The minimal "probabilities of exceeding" found are: $\min\left(P(\chi^2_{\mathbf{q}})\right) \approx 3.3\%$ (also marked in Fig. 3.6 bottom) towards $(l, b) = (247.5^\circ, -30^\circ)$[7] for the V data and $\min\left(P(\chi^2_{\mathbf{q}})\right) \approx 6.5\%$ towards $(l, b) = (281.5^\circ, -19^\circ)$ for the V5 data.

These two results agree well with the previously estimated (Eriksen et al., 2007) intrinsic modulation parameter value $A = 0.114$ at scales $\ell_{\rm max}\leq 40$, as they lie well within "one-sigma" region of the distribution of log-likelihoods obtained from 1000 simulations modulated with the modulation of $A_{40} = 0.114$. However we note that it will always be difficult to reject such modulation at high confidence level as it is also realized (to this or a greater extent) on average in $\sim 8\%$ of GRF simulations (Sec. 3.5.4).

Note that the distribution of the probabilities of the joint multi-region analysis (Fig. 3.13) has a very flat and extended maximum and, for example, the minimal joint-probability in

[7]This and the following result is accurate to within the tolerance of about $\sim \pm 7^\circ$ resulting from the low-resolution search involving only 96 directions over a hemisphere in the *HP* 4 pixelization scheme.

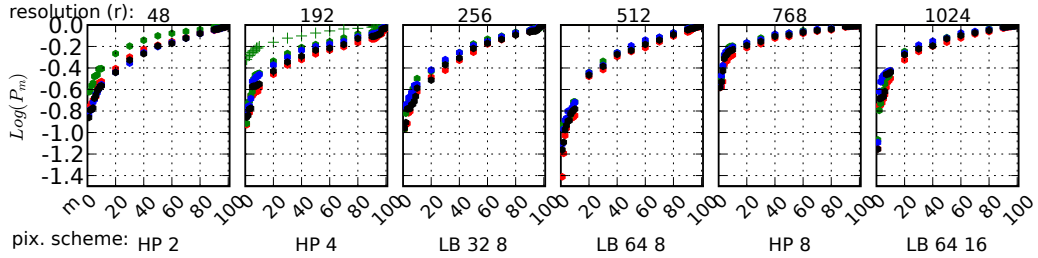


Fig 3.12: Results of the "multi-region" analysis with the dipole-corrected V (V5) data. Joint probabilities $P\left(\chi^2_{\mathbf{q}}\right)$ (Eq. A-7) for 600 *multi-masks*. The probabilities calculated with the spectral data maps from channels Q, V, W, INC are plotted with red, green, blue and black dots respectively. WMAP5 V5 data are plotted with green crosses (+) in *HP* 4 only. The V dataset original dipole has been replaced by the simulated dipole and removed in case of V5 data. Each dot (cross) corresponds to the joint probability of one *multi-mask*. From left to the right, the panels show the results with increasing resolution of pixelization schemes (see Table 3.1) with 100 different *multi-masks* in each panel. Only the "mean" data is shown since all other results remain almost unchanged. For better visualization in range $m \in [10, 90]$ we plot only every 10'th sorted probability value.

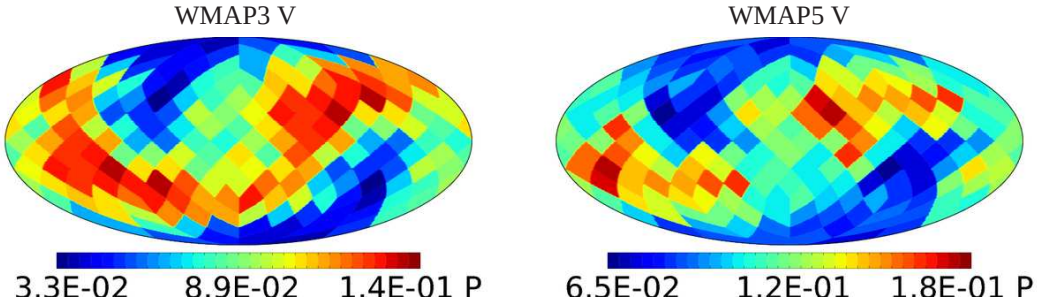


Fig 3.13: Results of the multi-region analysis, using the set of 96 *LB* 1 2 *multi-masks*, of the filtered in SH space up to $\ell_{\max}= 40$, low-resolution V data (left), and V5 data (right). The color in each pixel encode the multi-region, joint "probability of exceeding" derived with the *LB* 1 2 *multi-mask* rotated to the direction of the center of that pixel. Note that half of pixels in each map is redundant.

the V data in the reported direction is only 0.2% smaller from the probability corresponding to the direction close to the galactic pole, which is roughly 50° away from the minimal probability direction.

Analogous analysis, involving the *LB* 1 2 *multi-masks*, but performed on the full resolution unfiltered WMAP V5 data, results in larger minimal-probabilities: $\min\left(P(\chi^2_{\mathbf{q}})\right) \approx 9.6\%$ and the probability is minimized towards $(l, b) = (225^\circ, -78^\circ)$.

Residual dipole of the WMAP V channel.

In Fig. 3.10 we see a deviation in the distributions of the mean in the V dataset (green dots) and in V5 dataset (green crosses) for most of the *multi-masks*. The fact that it is visible in most *multi-masks* suggests that the anomaly is not particularly sensitive to the *multi-mask* orientation, and that it comes from large angular scales. Indeed, as measured by the n_σ values in individual regions of the *multi-mask* with the lowest joint probabilities, no region significantly deviates from simulations.

However, we find that the V data are fully consistent with the GRF simulations if we remove the dipole from the data, which is roughly ~ 2 times larger than the one in our simulations [8]. Actually the dipole values in the datasets as measured by the multipoles $l = 1$ on the Kp03 cut sky power spectrum are: $47(\mu\mathrm{K})^2$, $54(\mu\mathrm{K})^2$, $45(\mu\mathrm{K})^2$ in Q, V, W datasets respectively. The corresponding values in the WMAP5 data are: $64(\mu\mathrm{K})^2$, $54(\mu\mathrm{K})^2$ in V5 and W5 maps respectively. The measurements of dipoles on initially dipole-free maps, using a sky mask, introduce a bias due to power leakage from other coupled multipoles, leading to non-zero dipole amplitudes. When sky-cut-generated dipoles are statistically accounted for, the result would yield: $20^{+104}_{-26}(\mu\mathrm{K})^2$, $27^{+104}_{-26}(38^{+104}_{-26})(\mu\mathrm{K})^2$, $18^{+104}_{-26}(27^{+104}_{-26})(\mu\mathrm{K})^2$ in the WMAP3 (WMAP5) data at 95% CL, and hence is consistent with vanishing intrinsic dipole (except for the V band channel). We note that the noise component generates dipoles with amplitudes of order $C_1 \sim 0.01(\mu\mathrm{K})^2$ which is about three orders of magnitude less than the leakage effect. However the 95% CL effect is not sufficient to explain the strong anomalies detected in the regional tests.

The anomaly is more visible in the difference of dipole amplitudes between different channels. The difference of $9 \pm 3(\mu\mathrm{K})^2$ (at 95% CL) between channels V and W is excluded using simulations at $> 99.9\%$ CL assuming that it is generated only by the power leakage from the cut sky. The difference in the WMAP V5 data is even larger: $11 \pm 3(\mu\mathrm{K})^2$.

As for the amplitude of the V band dipole again, it becomes anomalous as one considers not only the magnitude, but also its orientation. The dipoles generated due to power leakage are strongly aligned within the galactic plane (Fig. 3.14) as a result of the shape of the Kp03 sky mask. While the Q and W dipoles measured on the cut sky are close from each other and close to the galactic plane pointing at $(l, b) = (13^\circ, -8^\circ)$ and $(l, b) = (7^\circ, 5^\circ)$ respectively, the dipole of the V band points at $(l, b) = (350^\circ, 30^\circ)$ which is itself anomalous at $> 97\%$ CL. WMAP5 data yield the dipole orientation $(l, b) = (203^\circ, 28^\circ)$ (Fig. 3.14).

We note that all dipoles with $b > 30^\circ$ have much smaller (roughly by an order of magnitude) amplitude than the one in the V band of the WMAP, which we believe is the reason for strong detections in the regional statistics carried out in the previous sections. When combining the alignment of the V band dipole with its magnitude, the hypothesis that it's generated only via the power leakage can be excluded at a very high CL since out of 1300 simulations, and within the subset of 37 that have generated dipole aligned at $|b| > 30^\circ$ the maximal generated power is of only $13(\mu\mathrm{K})^2$ which makes even the CL

[8] During the final stage of this work this sort of anomaly was independently reported in Eriksen et al. (2008)

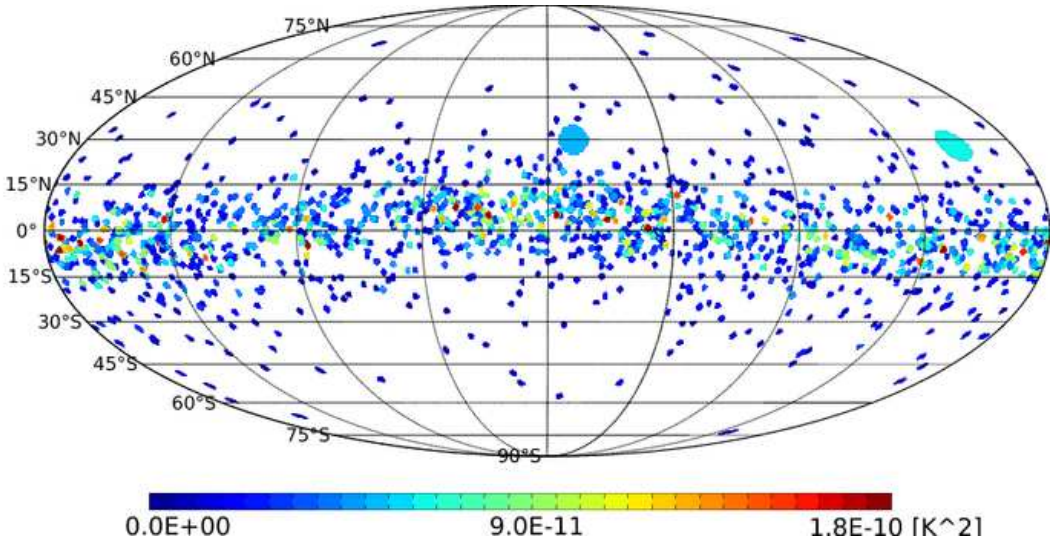


Fig 3.14: Distribution of dipole orientations generated due to Kp03 cut sky from 1467 full sky simulations with vanishing initial dipole. The large spot in the center (at the right-hand side) of the plot, indicates the orientation of the WMAP3 (WMAP5) V band dipole. The color scale reflects the amplitudes of the leakage generated dipoles.

estimation unfeasible, since this when compared to the $54(\mu\mathrm{K})^2$ of the V dataset ($65(\mu\mathrm{K})^2$ for V5), the simple χ^2 test implies a rejection basically without doubt to a very reasonable limits.

By reducing the dipole amplitude to the level consistent with simulations, or alternatively, by replacing it with our of our simulated dipoles, the data become consistent with our simulations at $< 2\sigma$ CL in the joint multi-region analysis (see Fig. 3.12) and at $< 1\sigma$ CL in all-*multi-masks* analysis at all resolutions (see Table 3.5 in the next section). Note that the presence of this dipole in the V band is of no cosmological consequences since the dipole is marginalized over for any cosmological analysis, but may be important for other low-ℓ analyses.

3.6.3 All-*multi-masks* analysis

We now discuss the result of the all-*multi-masks* analysis described in Sect. A-2.3. The corresponding results are presented in Table 3.5. We see that the data are consistent with the simulations at $\sim 68\%$ CL. The previously mentioned (in Sect. 3.6.2) tentative NG detection in the INC data in kurtosis has indeed the largest probability of rejecting across the scales (61%) but it turns out to be statistically completely insignificant.

The large scale variance distribution is found to be perfectly consistent with the GRF simulations.

We find a significant anomaly in the dipole component of the V band channel (see also Sec. 3.6.2) detected via distribution of means in both WMAP3 (99.8% CL) and WMAP5 (99.3% CL) data.

Table 3.5: Results of the all-*multi-masks* analysis for the signal dominated (co-added) and noise dominated (difference) maps. The columns content is as follows: (1) data set, (2) pixelization scheme, (3..5) "probability of rejecting" the consistency with GRF simulations (Sect. A-2.3) for each MOD. In case of V and V5 datasets, the probabilities for the data with corrected dipole component are given in brackets. We round to integer percentiles for probabilities $< 99\%$. The saturated values are marked with $\star$. We abbreviate the results consistent at given CL as: "no significant detections (CL)".

d	Reg.sk.	P^m_{rej}[%]	P^σ_{rej}[%]	P^S_{rej}[%]	P^K_{rej}[%]
(1)	(2)	(3)	(4)	(4)	(5)
INC	ALL	no significant detections (68%)			
Q	ALL	no significant detections (68%)			
V	*HP* 2	59 (20)			
	HP 4	99.8 (35)			
	HP 8	88 (16)	no significant detections (68%)		
	LB 32 8	99.5 (30)			
	LB 64 8	94 (37)			
	LB 64 16	86 (24)			
V5	HP 4	99.3 (15)	no significant detections (68%)		
W	ALL	no significant detections (68%)			
QV	ALL	$> 99.9^\star$			
VW	ALL	$> 99.9^\star$	no significant detections (95%)		
QW	ALL	$> 99.9^\star$			

All of the difference maps (see Sect. 3.6.5) show a very strong departures from foregrounds free, simulated, difference maps, most prominently detected in means distributions.

3.6.4 The "cold spot" context

A NG anomalous kurtosis excess of a wavelet convolution coefficients has been reported (e.g. Cruz et al. (2007)) in the southern hemisphere, and was found to be associated with the locally cold spot (CS) in the CMB fluctuations around galactic coordinates $(l, b) = (209^\circ, -57^\circ)$. In that work the wavelet convolution scales ranging from $\sim 6.6^\circ$ to $\sim 13.2^\circ$ in diameter were used, with anomaly being maximized at scales of $\sim 10^\circ$ with a rejection on grounds of Gaussianity assumption exceeding 99% CL. At the same time the authors note that the CS was not detected in the real space analyses.

The range of scales mentioned correspond roughly to the scales tested by the *HP* 4 ($\sim 14.6^\circ$) and *HP* 8 ($\sim 7.3^\circ$) pixelization schemes (see. Table 3.1).

In Fig. 3.15 we plot the scrambled n-sigma map of kurtosis from the single-region analysis, obtained from 100 *multi-masks* of the *HP* 8 pixelization scheme, which most closely corresponds to the scales, at which the CS was detected. The deviation of -3.6σ in the "the cold spot" direction, centered at $(l, b) = (209^\circ, -53^\circ)$ in Fig. 3.15 is clearly found along

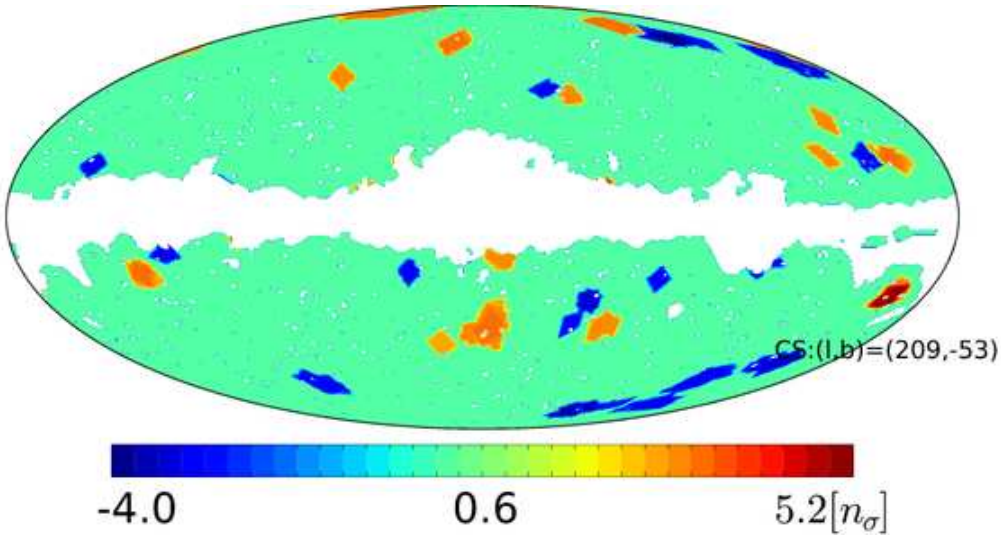


Fig 3.15: Kurtosis n-sigma map, thresholded at $n_{\sigma,th} = 3$, combined from 100 *multi-mask* of the *HP* 8 pixelization scheme. The "cold spot" is marked with "CS".

with many other, locally significant, deviations. This particular CS however is not found at the same location in e.g. *HP* 4 pixelization schemes or in skewness n-sigma maps, but rather it is shifted towards smaller galactic longitudes.

As an extension to the tested scales, in this section we run a high resolution test using pixelization scheme *HP* 16 corresponding roughly to scales of $\sim 3.7^\circ$ and the INC data. We use 10 additionally generated *multi-masks* in this resolution, and we perform the single region, joint multi-region, and all-*multi-masks* statistics. We also performed the same analysis using the filtered up to $\ell_{\max}= 40$ is SH space, low resolution ($n_s = 64$) maps in which the spot is clearly visible.

Although we find a locally negative KE and positive excursions from expected distributions by $\gtrsim 3\sigma$ around the CS direction in variance, we also find similar excursions at several other directions. The CS itself is well localized in the scrambled n-sigma maps of means with minimal value -2.9.

However, none of these detections (see also Fig. 3.7) hold under the scrutiny of the multi-region and all *multi-masks* analyses (Table 3.5), that find these local anomalies to be statistically insignificant.

3.6.5 Differential maps tests

We discuss results of the QV, VW and QW difference maps tests as a simple cross-check with the CMB signal dominated maps tests, and a rough estimation of the residual foregrounds amplitude. We limit the number of *multi-masks* to 10 and use $N_{\rm sim} = 10^4$ simulations in single region analysis and $N_{\mathrm{simPDF}(\chi^2_{\mathbf{p}})} = 10^3$ ($N_{\mathrm{sim}(\mathrm{C}_{\mathbf{p}})} = 9 \cdot 10^3$) simulations in joint multi-region analysis as before.

As shown in Table 3.5 the residual foregrounds are very strongly ($> 99.9\%$ CL) detected in all difference maps, due to anomalies in means distributions. In particular the n-sigma

and difference ($\Delta = m - \langle m \rangle$) maps of means distributions in QV, most prominently exhibit a dipole like structure oriented roughly at $(l, b) = (\ 260^\circ,\ 60^\circ)$ which is close to the kinetic dipole direction (Fig. 3.16a). The VW n_σ map has a similar dipole structure, but with opposite orientation, which however is absent in the QW n_σ map. We find this to be a consequence of the previously detected (Sect. 3.6.2) anomalous dipole component of the V band channel. Removing the dipole components from the difference maps, we have redone the three stages of the analysis, and although the dipole structure ceased to dominate, we still find a very high ($> 99.9\%$ CL) rejection probabilities for in means as quoted in Table 3.5.

In Table 3.6 we present the amplitudes of the residual foregrounds as measured by the variance of the Δ difference maps of means distribution for different scales as probed by our pixelization schemes. These remain in a good consistency with the limits given in Bennett et al. (2003a) for residual foregrounds contamination.

In Fig. 3.16(c-h) we show the n_σ maps with distribution of the regions in the difference maps outstanding from simulations at significance larger than 3σ (i.e. we use $n_{\sigma,th} = 3$). Clearly, the close galactic plane regions are strongly detected. We note that the the KQ75 sky mask partially removes the most affected regions around $(l, b) = (233^\circ, -10^\circ)$, $(l, b) = (259^\circ, 18^\circ)$ and the previously-mentioned $(l, b) = (199^\circ, -23^\circ)$.

It is interesting to note that the largest scale negative ($n_\sigma < 0$) anomalies seen in *HP* 2 (Fig. 3.16c-d) away from the galactic plane, can also be induced by the foregrounds dominating along the galactic plane (with $n_\sigma > 0$) due to a very strong linearity of foregrounds induced quadrupoles with strong maximums aligned along the galactic plane and consequently strong minimums allocated close to the poles (Fig. 3.16b). Such mechanism of foregrounds-generated linearity of the quadrupoles of the difference maps will not work in the foregrounds-free simulations, adding thereby to the observed large scale anomalies as probed via the largest regions. This effect is considerably smaller in the higher resolution pixelization schemes.

In order to test the consistency of our noise simulations with the noise of the WMAP data, and the approximation the uncorrelated, white noise and to constrain limits of the systematical uncertainties, we also performed analysis using Q12 and V12 difference maps in *HP* 2 pixelization scheme. The details of this analysis is given in Appendix A-3. Here, we briefly report the result that the systematical effects measured, as before, by the standard deviation of the difference Δ maps remains at level $< 1.7\mu$K at the scales corresponding the *HP* 2 pixelization scheme i.e. $\gtrsim 30^\circ$.

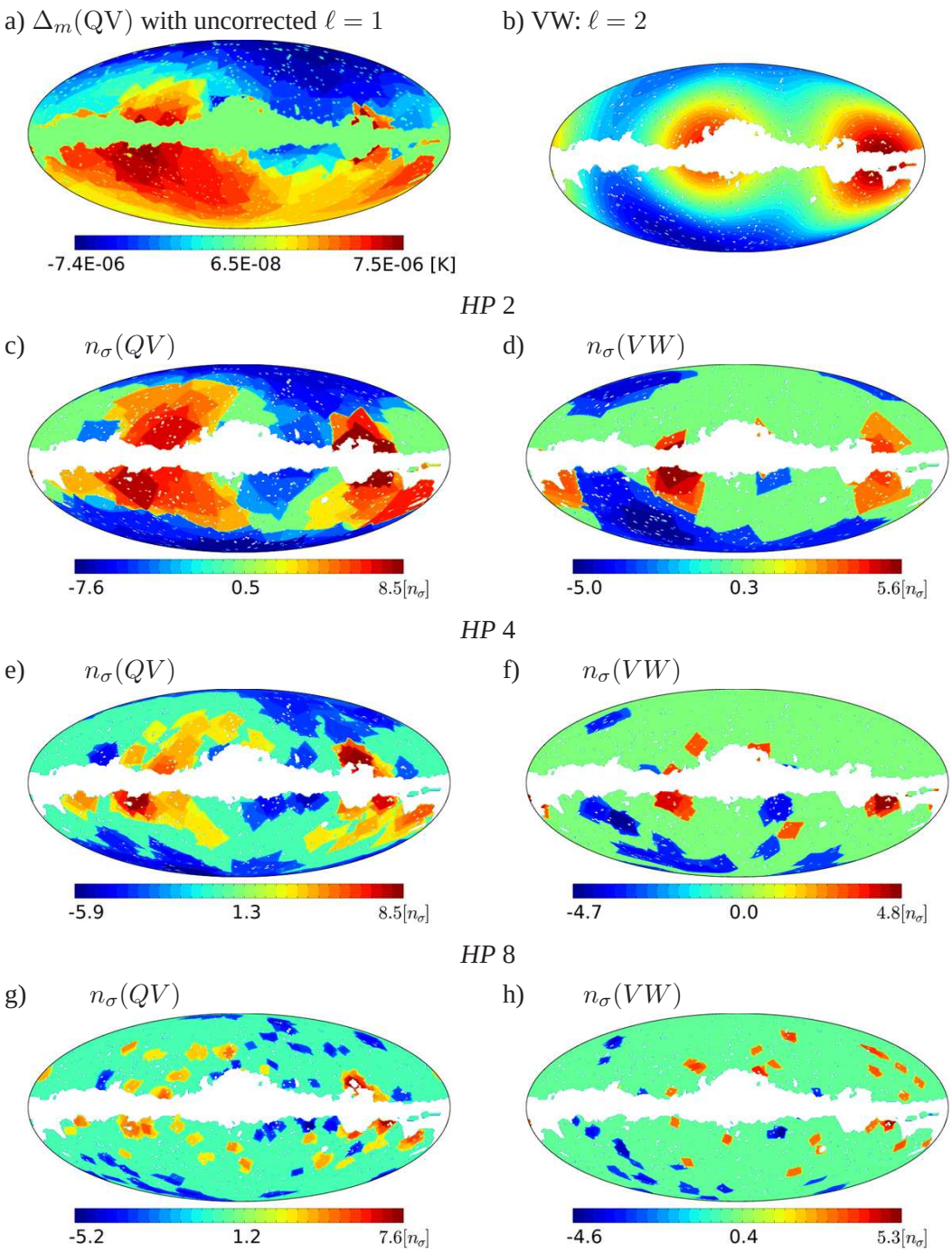


Fig 3.16: Results of the single-region difference maps analysis. In panel a) the residuals ($\Delta_m(\mathrm{QV}) = m - \langle m \rangle$) for the difference QV data, uncorrected for the anomalous V band dipole is plotted. In panel b) the quadrupole of the VW difference map and in panels c) through h) for the three *HP* pixelization schemes, we plot the thresholded at 3σ, n_σ maps of means distributions of the differential datasets QV (left) and VW (right) with the anomalous dipole component removed from the data. We make these maps (and maps of higher MODs) publicly available at `http://cosmo.torun.pl/~blew/SKregstat/`

Table 3.6: Residual foregrounds amplitudes in the cross-band difference maps. The columns contain: (1) pixelization scheme, (2) approximate angular size of the region as inferred from number of regions, (3) approximate corresponding multipole number $\ell = 180/\Omega_{\rm reg}$, (4..6) - standard deviation of the difference maps outside Kp03 sky mask.

Reg.sk.	$\Omega_{\rm reg}$ [deg]	ℓ	$\sigma(\mathrm{QV})$ [μK]	$\sigma(\mathrm{VW})$ [μK]	$\sigma(\mathrm{QW})$ [μK]
(1)	(2)	(3)	(4)	(5)	(6)
HP 2	29.3	6	2.7	1.9	3.8
HP 4	14.6	12	3.0	3.0	4.7
LB 32 8	12.7	14	3.9	3.5	5.1
LB 64 8	9.0	20	4.7	4.6	6.0
HP 8	7.3	25	4.9	4.9	6.2
LB 64 16	6.3	29	5.8	6.0	7.2

3.7 Discussion

3.7.1 Sensitivity, correlations and extensions

Although we have shown that the statistics is rather helpless to robustly detect the NG signals considered (defined in Sec. 3.5.2) via MODs higher than the variance, the statistics also proved to be a sensitive and precise tool for statistical isotropy measurements via variance and means. While we fail to detect the NG templates (inducing signals of rms $\sim 100\mu$K within spots of $\sim 10^\circ$ via skewness and kurtosis) in the multi-region NG analysis, the single region analysis detects these as locally significant ($n_\sigma \gtrsim 3$). Instead such template can be detected in the all-*multi-masks* analysis at $> 99.8\%$ CL via variance.

In Eriksen et al. (2004a) a regional statistical analysis was performed using a set of circular regions uniformly distributed across the sky. Our analysis is similar in spirit but uses different statistics and a richer sets of regions, varying both in size and shape. This approach has been validated by the fact that we have shown that the resulting statistical signal can be a sensitive function of the particular choice of regions. This is most prominently seen in case of the reported dipole anomaly in the V band of the WMAP data (Fig. 3.10), where it is easily seen that depending on the choice of the pixelization scheme (e.g. such as those associated with the results at the right-hand side of the top plots in Fig. 3.10) the obviously strong anomaly can be overlooked.

It is important to mention the correlations between various *multi-masks*. This correlation occurs since, although the *multi-masks* sample the data differently, in the end, the same data are being sampled. The degree of redundancy is directly connected to the number of *multi-masks* used in the analysis and the magnitude of the correlation is related to the size of the regions used in a pixelization scheme. To quantify this we carried out a simple statistical test only whose goal was to establish whether our test is statistically more sensitive to the change of the tested data itself or to the change of the *multi-masks* for various resolutions

parameter r and MOD.

If the results from different *multi-masks* were strongly correlated, then for a constant number of DOF one would expect the variance of the results measured over these *multi-masks* (e.g. χ^2 values) to be much smaller than the variance computed while fixing a *multi-mask* but varying dataset. On the other hand, if the variance of the results when changing dataset was much smaller than the variance when changing *multi-mask*, then the test would not be very sensitive to particular features in the data, or even unstable. To test this we calculated the R statistic defined as follows:

$$R(r, \mathrm{MOD}) = \frac{\left\langle \sigma_{\mathrm{sim}}(\chi^2(r, \mathrm{MOD})/\mathrm{DOF}_{\mathrm{eff}}) \right\rangle_m}{\left\langle \sigma_m(\chi^2(r, \mathrm{MOD})/\mathrm{DOF}_{\mathrm{eff}}) \right\rangle_{\mathrm{sim}}} \tag{3.4}$$

where $\left\langle \right\rangle_{\mathrm{sim}}$ denotes an average over simulations and $\left\langle \right\rangle_m$ denotes an average over *multi-mask* for a given simulation, and where $\mathrm{DOF}_{\mathrm{eff}} = DOF(r, m)$ is the effective number of degrees of freedom. Measuring this R statistic using our simulations, we find that the test is approximately equally sensitive at all resolutions and for all MODs, and that eventually it is little more sensitive to the change of the data under test than to the change of the *multi-mask* giving R values around 1.2.

The approach with arbitrary shape of the regions (*multi-mask*) and their orientation is quite flexible, and different shapes can possibly be used for different applications independently of the enforced sky cuts. This allows for a thorough test of the multivariate nature of this Gaussian field. One could also consider other statistics than the first MODs, as e.g. regional Minkowski functionals. Another possible extension is to apply a specific pre-filtering of the data in the SH space in order to expose for the test features dominating at particular scale. Such slicing of the data into subsets of maps according to some chosen ranges of multipoles could in principle significantly improve the sensitivity. The multi-region full-sky analysis though is restricted generally to the resolutions up to which the full covariance matrix analysis is feasible.

3.8 Conclusions

We introduce and perform a regional, real space test of statistical isotropy and Gaussianity of the WMAP CMB data, using a one-point statistics. We use a set of regions of varying size and shape (which we call *multi-mask*) allowing for an original sampling of the data. For each of the regions we analyze independently or simultaneously, the first four moments of distribution of pixel temperatures (i.e. mean, variance, skewness and kurtosis).

We assess the significance of our measurements in three different steps. First we look at each region independently. Then we consider a joint multi-region analysis to take into account the spatial correlations between different regions. Finally we consider an "all-*multi-masks*" analysis to assess the overall significance of the results obtained from different sky pixelizations.

We show that the results of such multi-region analyses strongly depend on the way in which the sky is partitioned into regions for the subsequent statistics and that our approach offers a richer sampling of the underlying data content. Consequently the all-*multi-masks* analysis provides a more robust results, avoiding possible biases resulting from an analysis constrained only to a single choice of pixelization scheme.

We find the three-year WMAP maps well consistent with the realistic, isotropic, Gaussian Monte-Carlo simulations as probed by regions of angular sizes ranging from 6° to 30° at 68% confidence level.

We report a strong, anomalous (99.8% CL) dipole "excess" in the V band of the three-year WMAP data and also in the V band of the WMAP five-year data (99.3% CL) (Sect. 3.6.2).

We test the sensitivity of the method to detect particular CMB modulation signals defined via the scale dependent modulation amplitude parameter ($A_{\ell_{\rm max}}$) for the case of modulation extending up to maximal multipole number of $\ell_{\rm max} = 40$ and $\ell_{\rm max} = 1024$. We are able to reject the modulation of amplitude of $A_{1024} = 0.114$ at $> 99.9\%$ CL and find that $A_{40} = 0.114$ can be statistically excluded only at $\sim 92\%$ CL (Sect. 3.5.4, 3.6.2). Given the WMAP V band data, we find that the large-scale hemispherical asymmetry is not highly statistically significant in the three-year data ($\sim 97\%$) nor in the five-year data ($\sim 93.5\%$) at scales $\ell \leq 40$. Including a higher-ℓ multipoles only decreases the significance of hemispherical variance distribution asymmetry.

We also test the sensitivity to detect a broad range of small (10° in radius) locally introduced NG signals, inducing non-vanishing kurtosis (and in general skewness) of rms amplitude $\sim 100\mu$K and find that the method is able to detect these as locally significant, but the overall impact in the joint multi-region analysis is unnoticed by mean, skewness and kurtosis, but is strongly detected ($\sim 99.8\%$ CL) by variance distributions. We conclude that the NG foreground-like signals will be easier to detect using local variance measurements rather than higher moments-of-distribution.

We also analyze our results in context of the significance of the "cold spot" (CS), reported as highly anomalous at scales corresponding to $\sim 10^\circ$ in diameter. While we notice the cold spot region as having locally anomalous, negative kurtosis-excess and locally increased variance (eg. Figs. 3.7, 3.15), we do not find these deviations to be globally statistically significant.

We easily detect the residual foregrounds in cross-band difference maps at average rms level $\lesssim 7\mu$K (at scales $\gtrsim 6^\circ$) and limit the systematical uncertainties to $\lesssim 1.7\mu$K (at scales $\gtrsim 30^\circ$) as a result of the analysis of same-frequency difference maps. These levels are consistent with the previously estimated limits.

Chapter 4

Hemispherical power asymmetry: parameter estimation from the Wilkinson Micowave Aanisotropy Probe cosmic microwave background radiation five-year data

The material presented in this chapter has been accepted for publication in Journal of Cosmology and Astroparticle Physics and is also available as Lew (2008a).

4.1 Abstract

We reexamine the evidence of the hemispherical power asymmetry, detected in the CMB WMAP data using a new method. We use a different data filtering, and preprocessing , and a different statistical approach and an independent parameter estimation from those previously used. At first, we analyze the hemispherical variance ratios and compare these with simulated distributions. Secondly, working within a previously-proposed CMB bipolar modulation model, we constrain model parameters: the amplitude and the orientation of the modulation field as a function of various multipole bins. Finally, we select three ranges of multipoles leading to the most anomalous signals, and we process corresponding 100 Gaussian, random field (GRF) simulations, treated as observational data, to further test the statistical significance and robustness of the hemispherical power asymmetry. For our analysis we use the Internally-Linearly-Coadded (ILC) full sky map, and KQ75 cut-sky V channel, foregrounds reduced map of the WMAP five year data (V5). We constrain the modulation parameters using a generic maximum a posteriori method.

In particular, we find differences in hemispherical power distribution, which when described in terms of a model with bipolar modulation field, exclude the field amplitude value

of the isotropic model $A = 0$ at confidence level of $\sim 99.5\%$ ($\sim 99.4\%$) in the multipole range $\ell \in [7, 19]$ ($\ell \in [7, 79]$) in the V5 data, and at the confidence level $\sim 99.9\%$ in the multipole range $\ell \in [7, 39]$ in the ILC5 data, with the best fit (modal PDF) values in these particular multipole ranges of $A = 0.21$ ($A = 0.21$) and $A = 0.15$ respectively.

However, we also point out that similar or larger significances (in terms of rejecting the isotropic model), and large best-fit modulation amplitudes are obtained in GRF simulations as well, which reduces the overall significance of the CMB power asymmetry down to only about 94% (95%) in the V5 data, in the range $\ell \in [7, 19]$ ($\ell \in [7, 79]$).

4.2 Introduction

The Gaussianity and the statistical isotropy of the fluctuations in the Cosmic Microwave Backgrounds Radiation (CMBR) are two generic features of current standard cosmological model and are compatible with the simplest inflationary scenarios. These predictions have been extensively studied in number of works. An incomplete list includes: Komatsu et al. (2003); Komatsu et al. (2008); McEwen et al. (2006a); Vielva et al. (2004); Wiaux et al. (2006); Mukherjee & Wang (2004); Savage et al. (2004); Naselsky et al. (2007); Cabella et al. (2004, 2006, 2005); Chen & Szapudi (2006a); Curto et al. (2008a, 2007); de Troia et al. (2007); Armendariz-Picon & Pekowsky (2008); Souradeep et al. (2006); Hajian & Souradeep (2006,?); Samal et al. (2008); Hansen et al. (2004a,b); Bernui et al. (2007b,a); Naselsky et al. (2005); Chiang et al. (2003); Chen & Szapudi (2006b); Gaztañaga et al. (2003); Copi et al. (2004, 2006a,b); Bielewicz et al. (2005); Abramo et al. (2006); de Oliveira-Costa & Tegmark (2006); Land & Magueijo (2005b, 2007); Jaffe et al. (2005); Efstathiou (2003b); Eriksen et al. (2004a, 2007, 2008); Land & Magueijo (2005b); Park et al. (2006); Shandarin (2002); Wu et al. (2001); Park (2004); Lew & Roukema (2008); Lew (2008b); Donoghue & Donoghue (2005); Yadav & Wandelt (2008) and references therein. Within the theory of inflation the primordial fluctuations are expected to form a Gaussian Random Field (GRF) at the leading order in perturbation theory. These statistical properties are imprinted in the CMB fluctuations, providing an interesting window onto the processes of the early Universe. Although the instrumental effects, like non-Gaussian, and non-isotropic noise, or eccentric beams, and astrophysical foregrounds effects, like Galactic, and extra-Galactic point sources, and extended sources of emission are either well controlled or corrected for, or masked out, a set of an unexpected anomalies of various magnitude and at various scales have been detected in the current CMB data (Cruz et al., 2007; Bernui et al., 2007a; Land & Magueijo, 2005a; Copi et al., 2004, 2006a; Abramo et al., 2006; Land & Magueijo, 2007; Eriksen et al., 2004a; Land & Magueijo, 2005b,c; Covi et al., 2006; Destri et al., 2008) (see also Huterer (2006); Martinez-Gonzalez (2008); Cabella & Marinucci (2008) for recent reviews and references therein). These anomalies call for plausible theoretical explanations since, if robustly detected, these can be used as valuable observables of the physics of the early Universe, or a new window on some of the late time effects as well (Inoue & Silk, 2007; Erickcek et al., 2008; Bernui & Hipólito-Ricaldi,

2008; Akofor et al., 2007; Destri et al., 2008; Brown & Crittenden, 2005).

We re-investigate the well-known, hemispherical power asymmetry observed in the CMB maps. We revisit the properties of the asymmetry, constrain parameters of the previously proposed bipolar modulation field model (Gordon et al., 2005) responsible for generation of the asymmetry, and we estimate the statistical significance using a generic maximum likelihood method, and realistic Monte-Carlo simulations. Given that we introduce and utilize a different method, from those previously used, and rely on different assumptions, while relax some other, our results can serve as a separate cross consistency and stability check. We provide a through tests of the method so as to validate the presented results.

For the first time we estimate the parameters of the hemispherical modulation for selected ranges of multipoles. Finally, we assess the significance of the hemispherical power asymmetry via direct comparison to the GRF CMB simulations.

The main differences and extensions to the previous analyses are: (i) we rely on the local real-space measurements of the variance as an estimator for the power asymmetry, (ii) we do not assume any priors on the probability distribution function (PDF) for any of the modulation parameters, and explore the likelihood function in the full (albeit sparse) parameter space and apply interpolation, (iii) we fully include the effects of the cut-sky, cross-multipoles power leakage and (iv) we analyze the power asymmetry in the various slices through the spherical harmonic space, pre-filtering the data prior to the analysis, rather than considering all scales scrambled together.

Since the full exploration of the parameter space is CPU expensive, our analysis is based on a few assumptions that greatly simplify, and speed-up the parameter estimation process. Using this different approach, while providing tests and justification for the assumptions made, we give a new estimates on the significance of the hemispherical power asymmetry anomaly.

We also discuss limitations of usage of the method with regard to the extent to which the assumptions of the method remain acceptably valid.

The organization of the chapter is as follows: In Section 4.3 we describe our datasets and CMB simulations. In Section 4.4 we present results of a statistics that measure the hemispherical power ratios. In Section 4.5 we focus on the properties of the power modulation model, our assumptions and tests of the assumptions, and then detail on our method for modulation parameter estimation. In Section 4.6 we present the results of various tests of the method. Results of the application of the method to the real CMB data are presented in section 4.7. Discussion and conclusions are given in sections 4.8 and 4.9 respectively.

4.3 Data and simulations

For the main analysis in the chapter we use the WMAP five-year foreground reduced CMB temperature maps (Hinshaw et al., 2008) from differential assemblies (DA) V1 and V2, because these spectral channels provide the best trade off between foregrounds of different spectral properties (i.e. the blue tilted galactic dust emission, and red-tilted galactic

synchrotron and free-free emissions). We co-add these observations (and corresponding simulations) using the inverse noise pixel weighting scheme. We will refer to these maps as V5. We generate simulations using the fiducial best fit ΛCDM model power spectrum of Dunkley et al. (2008) (constructed using the mean likelihood parameters) which we call $C_\ell^{\rm fid}$.

Also, for comparison purposes, we will use the five year release of the Internally-Linearly-Coadded (ILC5) map and also for additional tests the Harmonic-Internal-Linear-Combination (HILC5) map (Kim et al., 2008).

In Fig. 4.1 we plot the power spectra of the data sets that will be used in the power modulation parameters estimation analysis. For the purpose of the analysis (to be explained latter) we create a fitted power spectrum to the ILC5 data by concatenating the $C_\ell^{\rm fid}$ power spectrum in the limit of low multipoles ($\ell < 30$), with the cubic spline fit to the piece-wise averaged, full-sky power spectrum, reconstructed from the ILC5 data for multipoles $\ell \geq 30$. We will call this fitted spectrum $C_\ell^{\rm ILCfit}$. We cut off the residual strong foregrounds in the ILC5 map at the $\pm 350 \mu$K threshold; a level estimated from GRF foregrounds-free simulations.

Note that in the analysis of the modulation parameter estimation we will focus only on the large scale multipoles ($\ell \leq 80$), where the differences in the power spectrum due to the 350μK chop of the residual galactic contamination remaining the ILC5 data are completely unimportant, and where the signal-to-noise (SNR) ratio yields $\mathrm{SNR} \gtrsim 10^2$ (Fig. 4.1).

Throughout the chapter we will use the V5 data along with the KQ75 sky mask unless otherwise noted. For comparison purposes, the ILC5 and HILC5 datasets are used without any masks throughout of the chapter [1].

[1] Although we realize that the usage of the unmasked ILC maps can lead to somewhat overestimated constraints on any power asymmetry signals due to residual foregrounds contamination, we include the analysis of these maps in the full considered multipole range ($\ell \leq 80$) mostly for comparison purposes.

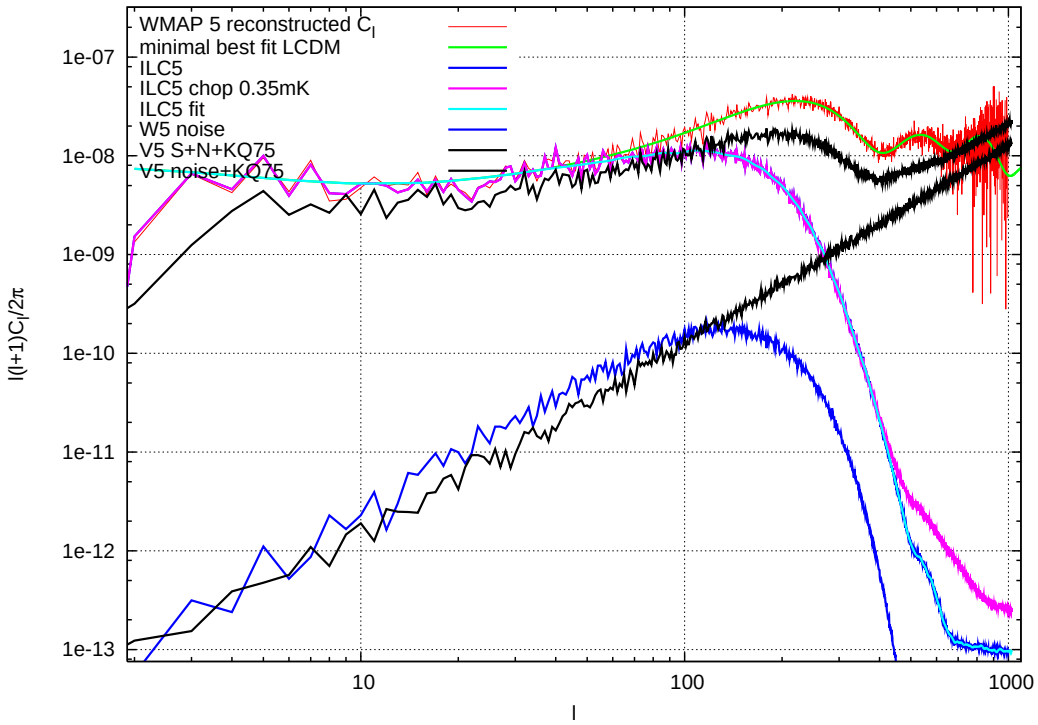


Figure 4.1: Power spectra of the data sets used in the analysis. We generate the GRF simulations of the V5 data using the best fit minimal ΛCDM model (green line) to the reconstructed WMAP5 data power spectrum (red line). The corresponding cut-sky pseudo power spectrum from one of the V5 simulations and V5 noise power spectrum are plotted in black (the top-black curve and the bottom-black curve respectively). Simulations of the ILC5 data are created using the fitted power spectrum $C_\ell^{\rm ILCfit}$ (light blue). The reconstructed from the ILC5 data power spectrum is plotted using the dark blue (top) line. The noise power of the most noisy channel of the WMAP data (W) smoothed with 1 degree (FWHM) Gaussian beam is also plotted (dark-blue bottom) and used here to place an upper limit constraint on the amount of noise in the ILC data. The power spectrum of the ILC5 data chopped at temperature threshold of 350μK is also shown (pink line). In the parameter estimation analysis we will use only the range of multipoles $\ell \leq 80$.

4.4 Hemispherical power ratio

We begin the analysis of the hemispherical power asymmetry by computing the following statistics:

$$\begin{aligned} R_{{\rm NS},\ell_{\rm max}} &= \max_{\hat{\mathbf{n}}_{\mathbf{S}}}\left(\frac{\sigma_{\rm N}(\hat{\mathbf{n}}_{\mathbf{S}},\ell_{\rm max})}{\sigma_{\rm S}(\hat{\mathbf{n}}_{\mathbf{S}},\ell_{\rm max})}\right) \\ r_{{\rm NS},\ell_{\rm max}} &= \min_{\hat{\mathbf{n}}_{\mathbf{S}}}\left(\frac{\sigma_{\rm N}(\hat{\mathbf{n}}_{\mathbf{S}},\ell_{\rm max})}{\sigma_{\rm S}(\hat{\mathbf{n}}_{\mathbf{S}},\ell_{\rm max})}\right) \end{aligned} \tag{4.1}$$

These are simply the maximal ($R_{{\rm NS},\ell_{\rm max}}$) and the minimal ($r_{{\rm NS},\ell_{\rm max}}$) ratios of the hemispherical standard deviations, found in the all sky search over directions $\hat{\mathbf{n}}_{\mathbf{S}}$, that define the

orientation of two hemispheres. The $\sigma_{\rm N}$ and $\sigma_{\rm S}$ values are the cut-sky (in case of V5 data) or full-sky (in case of ILC5/HILC5 data) hemispherical standard deviations of a map.

To define the grid of directions ($\hat{\mathbf{n}}_{\mathbf{s}}$), we choose to use the first 96, ring-ordered directions, defined by the pixel centres of the HEALPIX pixelization scheme (Górski et al., 2005) of resolution parameter $n_s = 4$. The measurements are performed using V5, ILC5 and HILC5 datasets within either, chosen ranges of multipoles, or as a function of a cumulative maximal multipole number $\ell_{\rm max}$.

The results of this survey are summarized in Fig. 4.2

The hemispherical asymmetry in the filtered bins of multipoles seems to be localized within the range of multipoles $\ell_{\rm assym} \in [8, 15]$ in all of the dataset: V5, ILC5 and HILC5.

Note that the first point in every plot: i.e. the quadrupole ($\ell = 2$), has the ratios always equal unity, because only the quadrupole map is used in variance measurements, and since the single-multipole maps have a point (anty)symmetry, due to the properties of the spherical harmonics, the variance is identical in the two hemispheres. In the range of multipoles $\ell \in (2, 30)$ we process the maps containing only two neighboring multipoles: eg. for the $\ell = 3$ we use combined maps of multipoles $\ell = 2$ and $\ell = 3$; for $\ell = 4$ we use combined maps of multipoles $\ell = 3$ and $\ell = 4$, and so on. For higher multipoles ($\ell > 30$) the bins are larger, and are defined by the two neighboring plotted points in Fig. 4.2. In the case of the cumulative plots (second and fourth row in Fig. 4.2) we use all multipoles from $\ell = 2$ up to $\ell_{\rm max}$, and due to the cumulative process, the curves on the right-hand side plots, (showing the statistics in the large-ℓ limit), do not exhibit much of details, since most of the map power (variance) comes from the low multipoles. In the limit of large multipoles, the multipole range-filtered maps show much more details as they are not overwhelmed by the power of the low multipoles. In particular, notice the strong, systematical deviation away from the simulation mean in case of the ILC5 and HILC5 data, starting with multipoles $\ell \gtrsim 150$. These are most likely caused by the extended foregrounds, and point sources remaining in the maps, since we do not apply any sky cuts with these data. We over-plot the results for the HILC5 data using the confidence level contours derived from the ILC5 simulations, just to compare them with the results obtained for the ILC5 data.

We also notice that in case of the multipole, range-filtered V5 data, (top row in Fig. 4.2) some asymmetry is also seen in the range $\ell \in [29, 40]$. On the same plot it appears that the "northern" hemisphere is anomalous due to the decrement in power as compared to the simulations in multipole range $\ell_{\rm assym} \in [8, 15]$.

For the case of the multipole, range-filtered results (first and third row in Fig. 4.2) we derive the joint "probability of rejecting", using a generic multivariate calculus, and for the V5 data we obtain result consistent with the simulations at $< 82\%$ confidence level for both the maximal hemispherical ratio ($R_{{\rm NS},\ell_{\rm max}}$) and the minimal hemispherical ratio ($r_{{\rm NS},\ell_{\rm max}}$) regardless of whether or not the off-diagonal terms of the covariance matrix are included. This result corresponds to the full range of multipoles up to $\ell = 300$.

As a final note to this analysis we point out, that the low significance of the power asymmetry, illustrated in Fig. 4.2, may result from the fact that we analyzed the data in

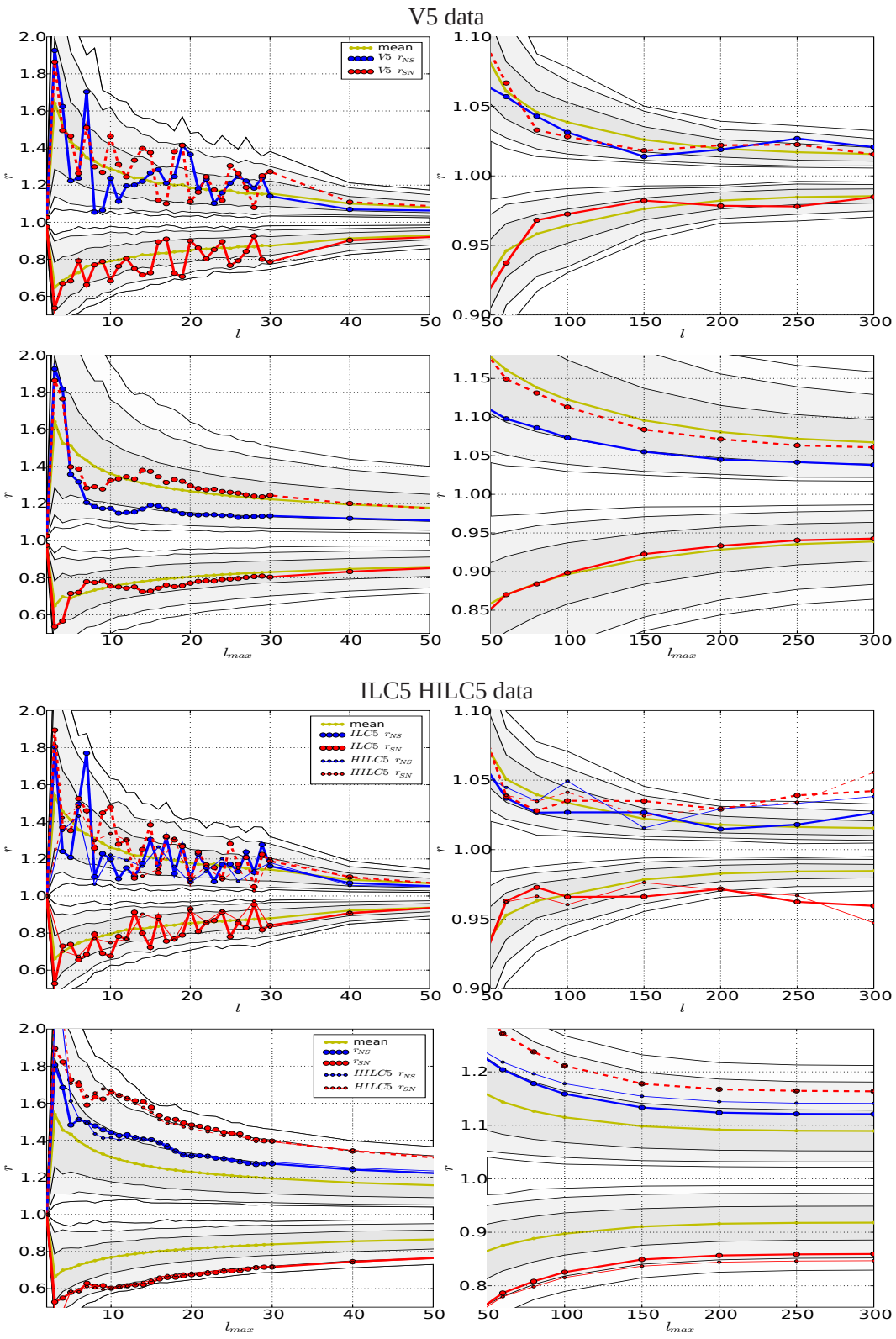


Figure 4.2: Hemispherical power asymmetry in filtered in the spherical harmonic space, V5 data (top panel) and in the ILC5 and HILC5 data (bottom panel). In the first row of each panel we plot the maximal ($R_{\mathrm{NS},\ell_{\mathrm{max}}}$) (blue)/minimal ($r_{\mathrm{NS},\ell_{\mathrm{max}}}$) (red) ratios of the standard deviations as a function of the filtered range of multipoles defined by the two ℓ values corresponding to any of the two neighboring plotted points (see text for details). In the second row of each panel we plot the ratios as a function of the cumulative maximal multipole number ℓ_{max} defining the considered range of filtered multipoles: $\ell \in [2, \ell_{\mathrm{max}}]$. The red (dashed) line is the inverted minimal ($1/r_{\mathrm{NS},\ell_{\mathrm{max}}} = r_{\mathrm{SN},\ell_{\mathrm{max}}}$) ratio of the hemispherical standard deviations, plotted to appreciate the power asymmetry by direct comparison with the blue lines. The gray bands represent the 68%, 95% and 99% confidence level contours. The simulations' mean is plotted with yellow solid line. The hemispherical asymmetry mostly appears to be confined to the multipoles range $\ell_{\mathrm{assym}} \in [8, 15]$ in V5, ILC5 and HILC5 dataset.

a thin slices through the spherical harmonic space (every two multipoles) up to $\ell = 30$. Although the full covariance matrix analysis should in principle, be stable to that, it's possible that other binning of the data may lead to somewhat different result. In order to check that we similarly estimated the joint "probability of rejecting" within the multipole range $\ell_{\rm assym} \in [8, 15]$ for the V5 data, but we found the data to be consistent with simulations at confidence level as low as 60%.

In the following we will constrain the properties of the hemispherical power asymmetry in greater detail.

4.5 Modulation parameters estimation

In Eriksen et al. (2007) an approach for estimating the bipolar modulation parameters has been concisely outlined, and implemented to obtain the constraints on the modulation parameters. In that work a Gaussian PDF proposal for all model parameters has been assumed, except for the modulation orientation axis, for which a flat PDF proposal has been used. As will be shown, the exact shape of the likelihood function may have and important effect on the overall significance of the power asymmetry anomaly, so in contrast to that work we directly reconstruct the likelihood function using a grid method.

Since the full parameter space operations are time consuming, we introduce few assumptions that greatly simplify the parameter estimation process. In the next section we discuss them one by one and provide appropriate justifications.

4.5.1 Bipolar modulation model parametrization

We generalize the parametrization of the CMB modulation, form the one defined in our previous work (Lew, 2008b), to account for a modulation that is effective only within a requested range of multipoles $(\ell_{\rm min}, \ell_{\rm max})$. A CMB observation $T(\hat{\mathbf{n}})$ of the GRF CMB realization ($T_{CMB}(\hat{\mathbf{n}})$) within a bipolar modulation model can be written as:

$$T(\hat{\mathbf{n}}) = \mathbf{B}(\hat{\mathbf{n}}, \hat{\mathbf{n}}') \star \Big(T_{CMB}(\hat{\mathbf{n}}')\big(1 + M(\hat{\mathbf{n}}')\big) + F(\hat{\mathbf{n}}')\Big) w(\hat{\mathbf{n}}) + N(\hat{\mathbf{n}}) w(\hat{\mathbf{n}}) \qquad (4.2)$$

where the modulation field $M(\hat{\mathbf{n}})$ is defined as:

$$M(\hat{\mathbf{n}}) = A^{\ell_{\rm max}}_{\ell_{\rm min}} \hat{\mathbf{m}} \cdot \hat{\mathbf{n}} \qquad (4.3)$$

where $\hat{\mathbf{n}}$ is a unit vector and M is a bipolar modulation field, oriented along direction $\hat{\mathbf{m}}$ with amplitude $A^{\ell_{\rm max}}_{\ell_{\rm min}}$, which modulates the CMB component only within the specified range of multipoles between $\ell_{\rm min}$ and $\ell_{\rm max}$. The $F(\hat{\mathbf{n}})$ and $N(\hat{\mathbf{n}})$ denote the residual foregrounds and the noise component respectively. The $\mathbf{B}(\hat{\mathbf{n}}, \hat{\mathbf{n}}')$ represents the real-space beam convolution kernel of the instrument, or any other effective convolution that has been applied to the data. The $w(\hat{\mathbf{n}})$ is the mask window function which can assume either 0 for masked pixels or 1 for unmasked pixels. In case of the ILC5 data $w = 1$. We will constrain the parameters $A^{\ell_{\rm max}}_{\ell_{\rm min}}$ and $\hat{\mathbf{m}}$ in different ranges of multipoles in order to investigate the modulation as a function of scale.

4.5.2 Assumptions

To facilitate the reconstruction of the multidimensional likelihood function in the following analysis we will rely on three assumptions. We assume that (i) the noise in our dataset in the range of multipoles under consideration can be neglected, (ii) that the dataset maps are foregrounds free, and (iii) that the residual systematical effects of the modulation-induced change to the underlying power spectrum of the CMB does not significantly influences the modulation parameter estimates. In the following sections we will discuss each of the assumptions in greater detail.

Signal to noise ratio

For the purpose of the analysis we assume that the noise in the CMB dataset (or simulations) has no significant impact on the CMB component modulation parameters estimates. The accuracy of this assumption is scale dependent. In Fig. 4.3 we plot the signal to noise ratio (SNR) as deduced from Fig. 4.1.

We choose to work with the filtered data at scales below $\ell \leq 80$, where the signal-to-noise ratio per single multipole is approximately $SNR \sim 10^2$ (Fig. 4.3 thick lines). Note that for the case of a real-space, variance measurements in a map composed of a range of multipoles, the SNR is larger due to the cumulative effect (Fig. 4.3 thin lines).

Foregrounds

We rely on the foregrounds reduced data of the cleanest WMAP5 channel - V and a conservative sky mask KQ75. As for the ILC5 data the residual galactic foregrounds are clearly seen in the map. We drastically reduce their amplitude by a sharp cut at the level of $\pm 350 \mu$K (limit deduced from foregrounds free simulations). Of course this doesn't remove the foregrounds but somewhat reduces their impact on the regional variance measurements. It should be noted that due to the residual foregrounds, a caution is recommended by the WMAP team when analysing this map at scales $\ell \gtrsim 10$ (Hinshaw et al., 2007). We present the results of the full sky ILC5 analysis for comparison purposes with the results obtained using an extensively masked V5 data. The V5 data should be therefore more reliable in the limit of large multipoles. However it will be interesting to compare the results between the two analyses both in the limits of low multipoles, where the ILC map should be reliable, and in the limit of large multipoles, where some residual foreground contaminations are present.

Modulation effects to the power spectrum

The modulation inevitably leads to a change in the underlying power spectrum at all scales, due to the multiplicative dipole component. Assuming a modulation orientation along the "z-axis" direction, the modulation field M expressed by a spherical harmonic of degree 1

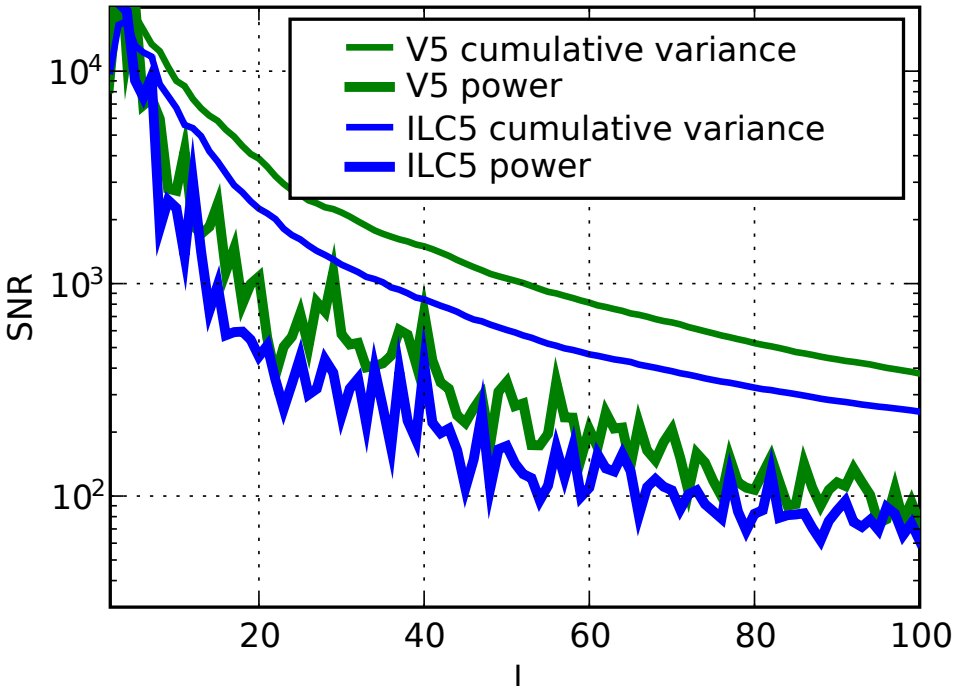


Figure 4.3: An estimate of the signal-to-noise ratio (SNR) per multipole in the pseudo-power spectrum ($C_\ell^{\rm signal}/C_\ell^{\rm noise}$) (thick lines) in the ILC5 (blue) and V5 (green) dataset and the signal-to-noise ratio of the cumulative variance measurements (thin lines) defined as: $\mathrm{SNR}(\ell) = \sum_{j=2}^{\ell}(2j+1)C_j^{\rm signal}/\sum_{j'=2}^{\ell}(2j'+1)C_{j'}^{\rm noise}$ in the ILC5 (blue) and V5 (green) dataset.

and order 0 is: $M(\hat{\mathbf{n}}) = A_{\ell_{\rm min}}^{\ell_{\rm max}} a_1^0 Y_1^0(\hat{\mathbf{n}}) = A_{\ell_{\rm min}}^{\ell_{\rm max}} a_1^0 \cos(\theta)$, [2] where $a_1^0 = 2\sqrt{\frac{\pi}{3}}$, and the spherical harmonic expansion of the modulated map reads:

$$T(\hat{\mathbf{n}}) = \sum_{\ell,m} a_\ell^m Y_\ell^m(\hat{\mathbf{n}}) + 2\sqrt{\frac{\pi}{3}} A_{\ell_{\rm min}}^{\ell_{\rm max}} \sum_{\ell,m} a_\ell^m \cos(\theta) Y_\ell^m(\hat{\mathbf{n}}) \tag{4.4}$$

where the first term corresponds to the initial CMB component and the second term corresponds to the modulated component.

Using the recurrence formula for the associated Legendre polynomials

$$(\ell - m + 1)P_{\ell+1}^m(x) = (2\ell+1)xP_\ell^m(x) - (\ell+m)P_{\ell-1}^m(x) \tag{4.5}$$

it is straightforward to see that

$$\cos(\theta)Y_\ell^m(\hat{\mathbf{n}}) = \frac{\ell-m+1}{2\ell+1}Y_{\ell+1}^m(\hat{\mathbf{n}}) + \frac{\ell+m}{2\ell+1}Y_{\ell-1}^m(\hat{\mathbf{n}}) \tag{4.6}$$

where $x \equiv \cos(\theta)$.

According to Eq. 4.6 the modulation leads to redistribution of power of a given multipole ℓ of the modulation component map on to the two neighboring multipoles $\ell - 1$ and

[2] We use the Abramovitz & Stegun (1972) notation convention, and their phase definition of the spherical harmonics throughout the chapter.

$\ell+1$ removing totally power from the multipole ℓ. Of course the power in ℓ'th multipole is restored by redistribution of power of the $\ell+1$ and $\ell-1$ multipoles. In general, the power is redistributed not only within the same m mode but also within the $m\pm1$ modes when the modulation orientation is allowed to assume an arbitrary orientation.

It can also be shown that the modulated map has statistically more power than the corresponding non-modulated map. This is clearly seen in the modulated map power spectrum (Fig. 4.4 top panel) as a systematic departure from the initial power spectrum, whose magnitude depends on the modulation amplitude $A_{\ell_{\rm min}}^{\ell_{\rm max}}$. In our analysis we will account for this systematical effect by a calibration of the map in the real space by the standard deviation (Fig. 4.4 middle panel). This approach does not eliminates the systematical effects, however it reduces them to a sub-percent levels (Fig. 4.4 bottom panel) for the modulations that are cosmologically relevant.

We study these systematical effects using a sample of 1 000 full sky ILC5 simulations, which we modulate to various extent. Then we reconstruct the average modulated power spectra, and compare it to the average non-modulated power spectra (Fig. 4.4). Note that e.g. for modulations $A\lesssim0.6$ the systematical effects of the modulation are smaller than 1% after a proper re-calibration of the map. In practice the larger modulations will be much stronger penalized due to the violation of the statistical isotropy, than due to the systematical deviations from the initial, fiducial power spectrum. It is also worth noting that the deviation from the initial power spectrum in the case of smaller, and much more relevant modulations is negligible, after variance recalibration (green curves in the bottom plot of Fig. 4.4).

Since we will be working with selected, filtered ranges of multipoles (in order to investigate the modulation hypothesis as a function of scale) the multipoles at the upper and lower ends of that range will have their power significantly suppressed, due to the modulation-induced power-leakage outside the chosen range. To improve the effectiveness of the calibration process, for a given multipole range of interest $[\ell_1,\ell_2]$ we will calibrate the modulation-altered simulations, and the data by standard deviations, calculated on the non-modulated, filtered maps, without the outer-most multipoles: i.e. calculated within range $[\ell_1+1,\ell_2-1]$. This improves the statistical consistency by several percent to the case when the calibration is done within the full considered multipole range $[\ell_1,\ell_2]$. This procedure has been actually used for the results given in Fig. 4.4.

The bottom panel of the Fig. 4.4 suggests that the fiducial power spectrum model used by Eriksen et al. (2007): $C_\ell^{\rm modulated}/C_\ell^{\rm fid}=a(l/l_0)^b$ with additional freedom allowed by the tilt parameter in the reference power spectrum could result in smaller residuals, however in principle, a possibly large running could also be needed in order to account for the residual, systematical discrepancies.

We also check that the same kind of relation (as depicted in Fig. 4.4) with very similar amplitude of the systematical effects is obtained for the case of data de-modulation (see. section 4.5.3).

We will test the accuracy of the method to reconstruct the modulation parameters in section 4.6.

4.5.3 Method

Our method is based on measurements of variances in two opposing hemispheres, in a sample of 3 000 GRF V5 and ILC5 simulations and in the corresponding data. We use 96 different directions in the northern hemisphere, that define the axial-symmetry axis of the two hemispherical regions. The directions are defined by the pixel centres of the HEALPIX pixelization scheme with the resolution parameter $n_s = 4$. In case of the analysis with V5 data the hemispherical regions are defined outside the KQ75 sky mask, and therefore the number of available pixels in these regions (N_k) may vary from one orientation to another.

In each region $k = [1, N_r]$ (here $N_r = 2$) we measure:

$$r_k = \frac{\sigma_k^2}{\sigma^2} \tag{4.7}$$

where σ_k^2 is the variance of the CMB within kth region and σ^2 is the variance of the whole map (outside the KQ75 sky mask in case of the V5 data). For the set of N_r regional measurements we define the corresponding χ^2 value as:

$$\chi^2 = \sum_{k=1}^{N_r} (r_k - \langle r_k^{\rm sim} \rangle)^2 / Var(r_k^{\rm sim}) \tag{4.8}$$

Note that given that we rely on the local variance measurements it is justified, to the limits to which CMB represents a GRF realization, and to the extent to which a cosmic covariance is unimportant, that we neglect the off-diagonal terms in the covariance matrix, and treat the regional variance measurements as independent variates. Although we realize that the r_k quantities defined in Eq. 4.7 should in principle follow a Fischer F-distribution, it is not clear which distribution a sum given in Eq. 4.8 should follow, and as such we reconstruct the likelihood function using a generic χ^2 distribution [3].

We seek for the best fit between the data and GRF simulations in terms of the hemispherical variance distributions (r_1, r_2) by de-modulating the data under assumptions given in Sect. 4.5.2. This approach is therefore a non-standard one, due to the fact that generically it's the simulations that are being fit to the data, rather than the data to simulations. However within the approximations given in Sect. 4.5.2, it is possible to reverse the process, by demodulating data, while retaining a formal correctness and allowing thereby to avoid a time-consuming processing of large number of simulations for each cell of the parameter space. Under no-noise and no-foregrounds assumptions we rewrite the the Eq. 4.2 as:

$$T(\hat{\mathbf{n}}) \;\; = \;\; \mathbf{B}(\hat{\mathbf{n}}, \hat{\mathbf{n}}') \star \Big(T_{CMB}(\hat{\mathbf{n}}') \big(1 + M(\hat{\mathbf{n}}')\big) \Big) w(\hat{\mathbf{n}}) \tag{4.9}$$

It is clear that apart from the beam smoothing effects, the de-modulation of the observed map $T(\hat{\mathbf{n}})$ is simply a division by the factor $\big(1 + M(\hat{\mathbf{n}}')\big)$. In order to account for the beams, using few spherical harmonics transformations (SHT), we pre-process the V5 data and simulations as follows:

[3] We will discuss and take into account possible consequences of this approximation in section 4.7.3.

1. downgrade the simulations/data to resolution $n_s = 128$

2. SHT analysis of the full sky V5 maps to obtain $a^m_{\ell,\mathrm{V5}}$ coefficients, and deconvolve them with the average V channel beam transfer function

3. SHT synthesis using $a^m_{\ell,\mathrm{V5,nobeam}}$ coefficients to obtain a map including the first 128 multipoles (in order to account for the power leakage from cut sky in point 4)

4. apply KQ75 sky mask to remove the foregrounds in the deconvolved map (there are no foregrounds in the simulations)

5. SHT analysis of the cut-sky deconvolved map to obtain $a^m_{\ell,\mathrm{V5,nobeam,cutsky}}$ coefficients up to $\ell_{\max} = 80$

We store the final set of $a^m_{\ell,\mathrm{V5,nobeam,cutsky}}$ coefficients to produce the beam free maps for any requested range of multipoles. Note that the two beams of the V1 and V2 WMAP channels are practically identical, however we still average between them to deconvolve the ILC maps. In the limit of the highest, considered multipoles the operation of deconvolution has an impact of few percent, as compared with the convolved power spectrum, as can be inferred from the shape of the beam transfer function. Also note, that since we operate in the signal dominated regime there is no danger to artificially blow up the high-$\ell(\approx 80)$ multipoles.

De-smoothing by the V band beam transfer function, leads to a power increase at $\ell = 80$ by about 10%. This can be easily estimated from the transfer function itself, since the transmittance for $\ell = 80$ is about 95%, which in the deconvolved power spectrum translates onto an increase by a factor of $1/0.95^2 =\sim 1.1$

We remind that the noise in case of V5 is, of course, present in the simulations. The approximation of "no noise" only means that we assume that the estimates of the modulation parameter, that modulates the pure CMB component, is not much altered by the fact that we are actually deconvolving noisy observations, rather than a pure CMB component, which in general does not make sense unless the signal strongly dominates the noise.

In case of the ILC5 data we create an effective "beam" transfer function by dividing the fit to the ILC5 power spectra (see. 4.3 for details) by the fiducial best fit ΛCDM model power spectrum (Dunkley et al. (2008) generated using the mean likelihood parameters): i.e. $b^{\mathrm{eff}}_\ell = \sqrt{C^{\mathrm{ILCfit}}_\ell / C^{\mathrm{fid}}_\ell}$. We divide the ILC5 a^m_ℓ coefficients ($\ell \leq 80$) by this function to match the ILC5 map power spectrum to the pure CMB component power spectrum. We will analyze this data with the 3000 GRF signal-only, full sky CMB simulations, generated with the best fit fiducial power spectrum C^{fid}_ℓ[4].

As a result, such preprocessed data sets and simulations (apart from the cut sky effects which are identical in the two) are consistent with the fiducial, best fit, theoretical power

[4] Actually for the data preparation process we only generate the GRF a^m_ℓ coefficients and then generate maps for any requested range of multipoles

spectrum, which we verify experimentally. We can rewrite the Eq. 4.9 as

$$T_{\rm nobeam}(\hat{\mathbf{n}}) \;\; = \;\; T_{CMB}(\hat{\mathbf{n}})\big(1+M(\hat{\mathbf{n}})\big)w(\hat{\mathbf{n}}) \tag{4.10}$$

Based on our preprocessed dataset we assured that the inferred modulation parameter will correspond to the modulation of the pure CMB component (to the extent where the assumptions given in section 4.5.2 are valid).

We seek for the best-fit modulation map $(1+M(\hat{\mathbf{n}}))$ such that if the observations $(T_{\rm nobeam}(\hat{\mathbf{n}}))$ are divided by it, the resulting map will yield the best consistency with the GRF simulations in terms of the statistics given in Eq. 4.8.

4.5.4 Parameter space

As was mentioned in the previous section, we use 96 different directions in the northern sky, that define a set of orientations of the N_r regions ($N_r = 2$ for hemispherical regions). The set of regions uniformly covers the whole sphere. The 96 directions define our search space, and the corresponding search parameter that we will call $\hat{\mathbf{n}}_{\mathbf{s}} = \{1..96\}$.

Additionally, we use 192 directions over the full sky, that define the orientation of the de-modulation axis $\hat{\mathbf{m}} = \{1..192\}$. The directions are defined by the pixel centres of the HEALPIX pixelization scheme of resolution parameter $n_s = 4$. Those that are localized in the northern hemisphere overlap with the directions defining the regions orientations. These directions define our modulation orientation space.

In the most general case we probe the likelihood function for the modulation amplitudes in range $A_{\ell_{\rm min}}^{\ell_{\rm max}} \in [0.0, 0.2]$ with step $\Delta = 0.01$, and in range $A_{\ell_{\rm min}}^{\ell_{\rm max}} \in [0.2, 0.3]$ with step $\Delta = 0.02$, and in range $A_{\ell_{\rm min}}^{\ell_{\rm max}} \in [0.3, 0.5]$ with step $\Delta = 0.05$, and in range $A_{\ell_{\rm min}}^{\ell_{\rm max}} \in [0.5, 0.7]$ with step $\Delta = 0.1$. These values define our modulation amplitude space.

As will be shown in Sect. 4.7 including the large modulations ($A \gtrsim 0.5$) mostly explores completely unimportant regions of the likelihood function, which is why our grid in this region is much sparser. In general however, the amount of the possible hemispherical variance asymmetry in the GRF simulations depends on both: the underlying power spectrum shape, and the selected range of multipoles.

Additionally we perform search in different bins of multipoles $(\ell_{\rm min}, \ell_{\rm max})$. The range of the multipoles tested is summarized in Table 4.5.4.

According to the CMB WMAP5 best fit ΛCDM model [5] our considered range of the multipoles: i.e. $\ell \leq 80$ make up for only about 32% of the total power in this model (of which cumulative variance we calibrate to unity at the maximal computed multipole number of $\ell = 2000$: $\sigma(\ell_{\rm min}, \ell_{\rm max}) = \sum_{\ell=\ell_{\rm min}=2}^{\ell_{\rm max}=80}(2\ell+1)C_\ell^{\rm fid} / \sum_{\ell=2}^{2000}(2\ell+1)C_\ell^{\rm fid}$). However it was shown in Lew (2008b) that the modulation of $A \approx 0.1$ extending all the way up to $\ell_{\rm max} = 1024$ (at which about 96% of the total CMB power is used) is excluded at a high confidence level ($> 99\%$ CL).

[5] http://lambda.gsfc.nasa.gov/data/map/dr3/dcp/params/c_l/wmap_lcdm_sz_lens_wmap5_cl_v3.dat

Throughout the analysis we work on maps of the HEALPIX resolution n_s, which depends on the considered range of multipoles so as to yield the condition: $n_s \geq \ell_{\max}/2$.

4.5.5 Parameter estimation

For each direction from our search space (see section 4.5.4) we reconstruct the likelihood function for each of the the modulation parameters values $\theta = (A, \hat{\mathbf{m}})$ and for each considered multipole range. As a first step we a perform minimization of the likelihood over the search parameter $\hat{\mathbf{n}}_\mathbf{s}$, in order to select only the measurements that maximize the possible variance distribution anomaly. We next derive the corresponding marginal posterior distributions using flat prior probabilities $\Pi(\theta|\mathcal{M}) = const.$ at each cell of our parameter space. Therefore, the maximum likelihood inference will lead to the same results as the maximal posterior results since according to the Bayes theorem:

$$\mathcal{P}(\theta|\mathcal{M}, T(\hat{\mathbf{n}})) \propto \mathcal{L}(T(\hat{\mathbf{n}})|\mathcal{M}, \theta)\Pi(\theta|\mathcal{M}) \tag{4.11}$$

where $\mathcal{P}(\theta|\mathcal{M}, T(\hat{\mathbf{n}}))$ denotes the posterior distribution, and $\mathcal{L}(\theta|\mathcal{M}, T(\hat{\mathbf{n}}))$ denotes the likelihood of the parameters θ within the hypothesized model $\mathcal{M}$ defined in Eq. 4.3,.

For the estimates on the modulation amplitude, the marginalized, one-dimensional probability distribution is interpolated using cubic spline interpolation, before computing the expectancy value, modal value, and confidence ranges. The marginalization over the modulation directions is not performed directly on the grid nodes, but rather on an interpolated, on the surface of the sphere, for each value of the modulation amplitude independently, posterior.

For the estimates on the modulation orientation, the marginalized, two-dimensional probability distribution is interpolated using two-dimensional tension splines on sphere (Renka, 1997). The marginalization over the modulation amplitude is not performed directly on the grid nodes, but rather on an interpolated, for each modulation orientation independently, posterior. The interpolation is done using a cubic splines and a dense equispaced grid. We also tested and compared the "interpolations" using spherical harmonics analysis, followed by $\mathrm{FWHM} = 14^\circ$ (the approximate size of the search step) Gaussian smoothing and synthesis up to $\ell_{\max} = 30$, and found that the results: i.e. the maximum likelihood value orientations and confidence contours are reasonably similar, with those where the fitted tension spline surface was used. The usage of cubic interpolations or spherical harmonic approximations in principle can lead to oscillations in the PDF that exceed below zero value especially in the tails of the distribution. In case of one dimensional interpolations we circumvent this problem by replacing a cubic interpolation with a piece-wise linear interpolation. Although this step will break the continuity of the first derivative of the PDF function, we mostly probe the likelihood function dense enough so that these effects are relatively small and unimportant. In case of two-dimensional interpolations we find that the potential oscillations, if exist, are small and happen far outside the considered confidence levels. In particular such oscillations would lead to artificial generation of multiple isolated contours for a given CL which we generally do not observe and consequently do

not consider this to be a problem. The most affected artifacts of the applied interpolation are observed for cases where the most preferred modulation amplitude is vanishing or is close to zero. In such cases, of course, there is no information on the modulation orientation.

While deriving our results we choose to rely on dense two-dimensional interpolations using tension splines on sphere rather than on, somewhat arbitrary, spherical harmonics analysis approach. Depending on the tension parameter the interpolating surface approaches the Delaunay triangulation (linear interpolation) solution for large values of the tension parameter, and cubic splines solution for zero-tension parameter. The interpolated surface on an equidistant (in galactic latitude and longitude) grid is dense enough so that is could easily be projected without holes onto an equal-pixel-area HEALPIX grid to ease the integration over the sphere in pixel space.

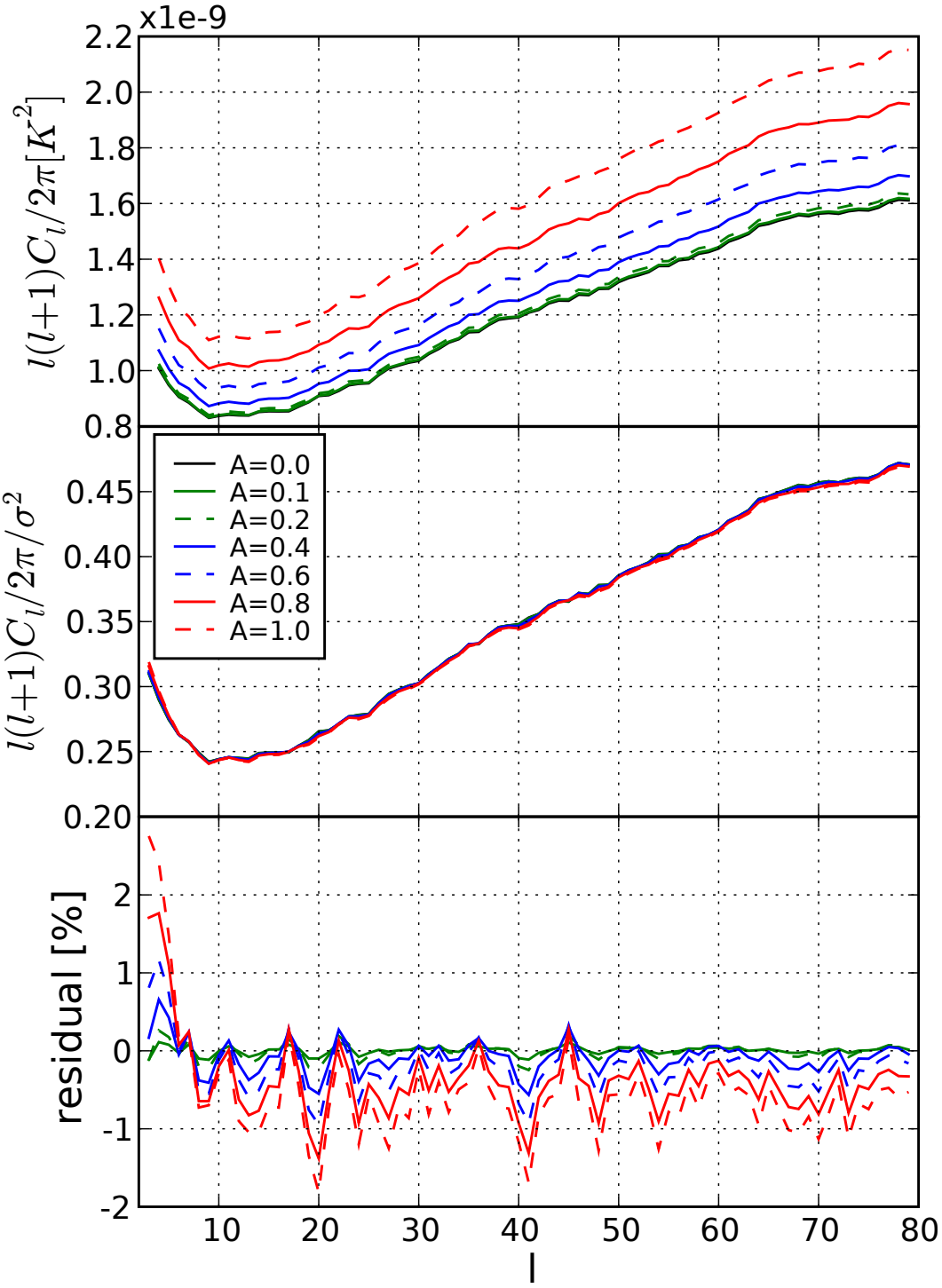


Figure 4.4: *Top panel:* Modulation induced systematical changes to the fiducial power spectrum as inferred from the average power spectrum from a sample of 1 000 simulations. The overall shift, approximately by a constant factor in all multipoles is clearly seen. *The middle panel:* Calibrated in real space, by the standard deviation, average power spectra of the modulated maps. *Bottom panel:* Residual, fractional systematical deviation from the original (non-modulated) power spectrum.

$\ell_{\rm min}\backslash\ell_{\rm max}$	7	20	30	40	60	80
2	$3 \sim 6$ (6.8) [21.2]	$3 \sim 19$ (13.8) [43.1]	$3 \sim 29$ (16.9) [52.6]	$3 \sim 39$ (19.4) [60.4]	$3 \sim 59$ (23.8) [74.2]	$3 \sim 79$ (27.9) [87.2]
6		$7 \sim 19$ (7.0) [21.9]	$7 \sim 29$ (10.1) [31.4]	$7 \sim 39$ (12.6) [39.2]	$7 \sim 59$ (17.0) [53.0]	$7 \sim 79$ (21.1) [66.0]
19			$20 \sim 29$ (3.0) [9.5]	$20 \sim 39$ (5.6) [17.3]	$20 \sim 59$ (10.0) [31.1]	$20 \sim 79$ (14.1) [44.1]
29				$30 \sim 39$ (2.5) [7.8]	$30 \sim 59$ (6.9) [21.6]	$30 \sim 79$ (11.1) [34.6]
39					$40 \sim 59$ (4.4) [13.8]	$40 \sim 79$ (8.6) [26.8]
59						$60 \sim 79$ (4.2) [13.0]

Table 4.1: Summary of the tested multipole bins. Note that as explained in section 4.5.2, in the actual analysis we discard the outermost multipoles of the considered multipole ranges. For clarity, in each cell, we explicitly write down the filtered multipole ranges used. The numbers in round brackets indicate the percentage of the variance within a considered range out of the total power in the best fit ΛCDM model, calculated as explained in the text. The numbers given in square brackets indicate the percentage of variance within a considered range of multipoles out of the total CMB signal variance within the first 80 multipoles of the fiducial power spectrum $C_\ell^{\rm fid}$.

4.6 Tests of the method

To test the correctness of the code, and the sensitivity of our approach we use a GRF white noise simulations. The choice of the white noise helps to generate a GRF simulation in which the cosmic variance effects, leading to accidental, unequal hemispherical power distribution, are suppressed, by giving as much power to high-ℓ modes, as to the low-ℓ modes. The magnitude of the low-ℓ modes dispersion is $\propto \sqrt{N_\ell/N_{\ell'}}$; i.e is larger than in the high-ℓ modes, where $N_{l\in\{\ell,\ell'\}} = 2l+1$ is the number of a_ℓ^m coefficients at multipoles ℓ, ℓ' respectively, where $\ell > \ell'$. Therefore equalizing power between different multipoles helps to better control the experiment: i.e. correctly interpret the results of the tests with the synthetic data given some input requested modulation parameters.

In case of realistic GRF simulations based on the CMB power spectrum, the hemispherical power asymmetry allowed within the cosmic variance uncertainty, is much larger due to the fact that the lowest multipoles (with the smallest number of modes) make up for the main part of the map's total variance, while the higher ℓ multipoles, even though more numerous, are strongly suppressed.

Full sky tests In the following tests there are no effects from any instrumental beams, nor cut skies. We generate a white noise realizations in resolution $n_s = 64$ and modulate them with modulation amplitude of $A = 0.1$ and with modulation axis $\hat{\mathbf{m}}$ oriented at $(l, b) = (225^\circ, -27^\circ)$.

At first we test the correctness of the code by using an analytical proposal for the PDF of the χ^2 values. Since we operate on white noise, zero mean and unit variance GRF simulations, its statistical properties are well known, and we therefore approximate the χ^2 value of the Eq. 4.8 as:

$$\chi^2 = \textstyle\sum_{k=1,2}(\sigma_k^2 - \langle\sigma_k^2\rangle)^2/Var(\sigma_k^2) \tag{4.12}$$

where

$$\langle\sigma_k^2\rangle = \frac{(N_k-1)}{N_k}\sigma^2, \qquad \text{and} \qquad Var(\sigma_k^2) = \frac{N_k^2}{2(N_k-1)}\sigma^4 \tag{4.13}$$

are the expectation value of the mean in the sample, and the expected variance of the variances in the sample of variates drawn from the GRF filed (Kenney & Keeping, 1951). The $\sigma^2 = 1$ is the variance of the Gaussian PDF from which the GRF is drawn, and N_k is the number of pixels in kth region.

We find that the statistics correctly reproduced the initial modulation amplitude and orientation well within the "one-sigma" confidence level (Fig. 4.5*a*) in all tested cases. Also, the tests show that using the white noise simulations, the method is able to reject the hypothesis of statistical isotropy at a very high confidence level (at $\gg 4\sigma$ CL in this case). The precision of the reconstructed, via interpolation, modulation direction orientation is surprisingly good ($\sim 2^\circ$) given a poor resolution of the search: $\sim 14^\circ$ (Fig. 4.5*b*)

We also test the reconstruction of the modulation parameters using a set of 1000 GRF simulations from which we derive the averages and variances of the regional variance realizations, and proceed according to equations 4.7 and 4.8 for $N_r = 2$. We notice an increase in the peakedness of the PDF when the Monte-Carlo ("MC") probed estimates are used, as compared to the theoretically ("TH") derived estimates. We speculate that the difference might come from the fact that we assumed the field variance value σ to be unity. In practice, we will always rely on a sample of 3000 MC simulations for estimates of local variance distributions.

Note that the simulation 3 in Fig. 4.5a, plotted as a peculiarity, found in one of our tests, traces the correct value of the injected modulation ($A = 0.1$) with an accuracy of about 5%.

Cut-sky tests and other subtleties We further test the stability of the method while varying the number of simulations used to probe the hemispherical (regional) mean and variance expectations. We check the dependence on the increase of the number of simulations from 1000 to 3000. We perform tests with the KQ75 sky cut, and test sensitivity of the method using different number of regions: $N_r = \{2, 10\}$. The regions for $N_r > 2$ are defined as an axial-symmetric patches, equally dividing galactic latitude into a symmetrical about the equator (before rotation) regions. In principle the increased number of regions could potentially have impact on the accuracy of the method. We use the effective number of degrees of freedom equal to the number of unmasked regions, in order to derive the likelihood value.

The results of the tests are presented in Figs. 4.5*c-f*. We find that the use of the increased number of simulations does not significantly influence the estimates of the reconstructed mode values, nor the shape of the marginalized PDF. Also it is clear that the increase of the number of regions, used in the statistics, broadens the marginalized PDFs, making thereby the statistic less sensitive.

Furthermore, we see that the accuracy of the method, in case of the cut-sky maps, is generally found at the level of few, up to several (in the worst case) percent of the level of the injected modulation ($A = 0.1$), which is of the same order as the unknown, initial (resulting from a random realization) modulation[6] of our white noise maps.

We note that the selected and presented "simulation 3" is one of the worst cases found in our tests, and as such, we give more attention to it in variety of tests summarized in Fig. 4.5*c* where the reconstructed distribution exhibits bi-modality. In general however, the simulations result in unimodal distributions, like those depicted in Fig. 4.5*e*.

We find that generally, in the presence of the cut sky, the modulation orientation is correctly reconstructed within 50% to 68% CL contours (Fig. 4.5*d*), regardless of the number of regions used in the statistics (2 or 10) however in the worst case simulation (as in case of the "simulation 3" with the two-region statistics) it is found as far as within the 95% CL

[6] We refer to the initial unknown, accidentally unequal power distribution in a GRF white noise simulation as a "modulation" since it's the modulation amplitude that we measure, but of course there's no reason to believe that any modulation effect, as defined in this chapter, exists in the GRF simulations.

contour.

We also checked the difference between different statistical approaches: i.e. between maximization[7] and marginalization over the modulation orientation. We find that both - the full sky and cut sky tests yield similar, or almost identical results. We will show that the situation will not be the same in case of the real CMB data or CMB simulations due to the effects we mentioned in the beginning of this section.

Finally we tested the statistical biases of the method under the cut sky conditions (Fig. 4.5*f (left)* red lines) and found that, within the obtained accuracy, no significant statistical bias is noticed.

Summary We find that with our chosen search resolution, the method traces the correct solution to within a few percent accuracy for the full sky measurements with respect to the injected modulation amplitude value, and from few up to several percent accuracy ($\lesssim 18\%$) for the cut sky case, with about 68% of the estimates yielding an accuracy better than $\sim 7\%$ (Fig. 4.5f). In terms of the absolute errors of the reconstructed value of the modulation amplitude parameter A, for the injected amplitude of $A = 0.1$ the errors are roughly an order of magnitude smaller: ~ 0.005 and ~ 0.007 for the full and cut sky cases respectively. In case of the larger number of regions the sensitivity of the method is worsened (eg. for the case of $N_r = 10$) and therefore in the following analysis we will only rely on the two-region statistics.

As for the reconstructed modulation direction, we find that mostly the correct direction is reconstructed within $\sim 50\%$ CL limits for the two regions statistics, and well within 50% CL limit for 10 regions case for the cut sky and full sky cases.

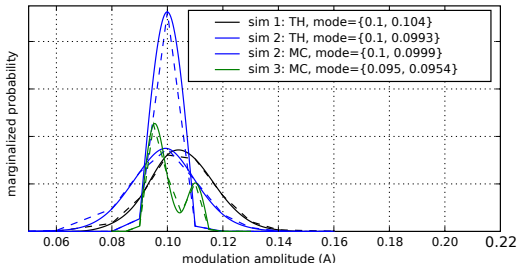
It is important to note that even with the white-noise simulations the initial, unknown modulation, resulting from random, and unequal distribution of power in the sky is at level of $A \lesssim 0.005$, which is of the same order of magnitude as the accuracy which we obtain in the full sky tests.

As for the reconstructed modulation orientation, mostly the correct direction is found to be within the 50% to 68% CL limits in case of the cut sky reconstructions, while the typical angular size of the 50% and 95% CL contours are $\sim 20^\circ$ and $\sim 35^\circ$ respectively (Fig. 4.5). The full-sky reconstructions CL contours are slightly smaller. Note that the modulation direction is reconstructed via interpolation to within a few degrees accuracy for the full sky case with the search resolution of about 14°, which is surprisingly good. For the cut sky case the accuracy is approximately at the level of several degrees.

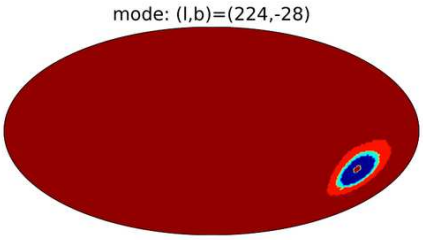
[7] In case of maximization over the modulation direction orientation, we have used the modal values, found in the fitted, two-dimensional maps of the likelihood function, for each modulation amplitude. We found this method to improve the smoothness of the resulting PDF, since our parameter search space is very sparse - only 192 directions over the entire sky.

Full sky white-noise maps tests with initial modulation: $A = 0.1$ $(l, b) = (225^\circ, -27^\circ)$

a) marginalized modulation amplitude PDF:

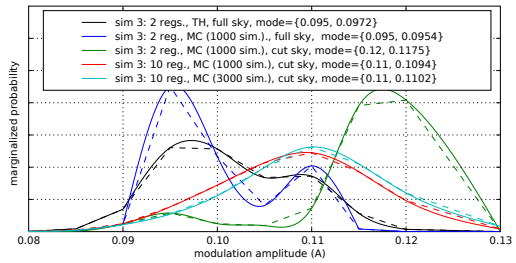


b) marginalized 50%, 68% and 95% CL modualtion orientation limits

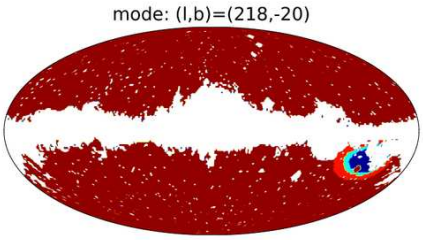


Cut sky (KQ75) white-noise maps tests with initial modulation: $A = 0.1$ $(l, b) = (225^\circ, -27^\circ)$

c) marginalized modulation amplitude PDF:

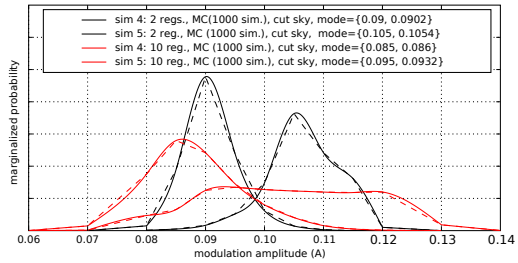


d) marginalized 50%, 68% and 95% CL modualtion orientation limits



Two-region statistic tests with cut sky (KQ75) white-noise maps and with initial modulation: $A = 0.1$ $(l, b) = (225^\circ, -27^\circ)$

e) marginalized modulation amplitude PDF:



f) statistics of the reconstructed ML modulation amplitude values from 100 GRF simulations

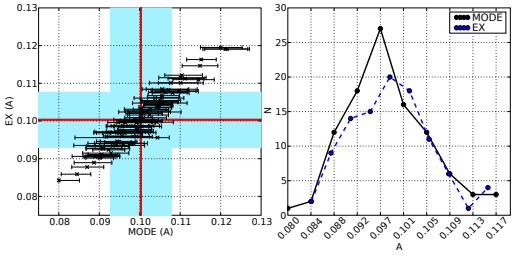


Figure 4.5: White noise maps tests of the accuracy of the method to reconstruct the injected modulation amplitude and its orientation. *(a)* Constraints on the modulation amplitude, from three, chosen, full sky, white-noise, modulated maps. The three colors code three selected simulations. The dashed lines for each simulation connect points at which the likelihood function was calculated, and the solid lines represent a cubic polynomial (spline) interpolation, combined with a linear interpolation in the tails of the distribution so as to avoid the oscillations into negative values. The abbreviations "TH" (for the black and the blue-spiky curves) and "MC" (for the green and the broad-blue lines) in the legend indicate respectively: the theoretically derived, and Monte-Carlo probed values of the expected means and variances of the regional (hemispherical) variances (Eq. 4.13). In the "MC" case a sample of 1 000 simulations is used. The mode values of the probed and interpolated PDFs are also given. *(b)* Limits of the modulation orientation, from one of the modulated, white noise, full sky, GRF realization. The 50% (dark blue) , 68% (light blue) and 95% (red) confidence level ranges are plotted. The small red dot indicates the location of the $\hat{\mathbf{m}}$ direction, and in the top of the plot the ML modulation direction is given. We obtain similar results in all tested cases. *(c)* As in panel *(a)* but for the KQ75 cut-sky, modulated, white-noise maps tests. We focus here only at the peculiar, "worst-case" – the "simulation 3", subject to different statistical approaches as indicated in the legend. For comparison, the "TH" PDF, and the PDF for the full sky case for this simulation, are also replotted. *(d)* Limits of the modulation orientation for the sky cut "simulation 3", using $N_r = 10$ region statistic. The details are as in panel *(b)*. *(e)* Constraints on the modulation amplitude, from two another (less-extreme than "simulation 3") simulations, derived using the two-region statistics (black curves), and the corresponding PDFs for 10 regions statistics (red curves). *(f) (left)* Reconstructed modulation PDF mode values versus expectation values from 100 cut-sky simulations, with 68% CL error bars derived individually from each PDF independently, plotted with the corresponding histograms (*(f) right*). The CL ranges were integrated from the mode value. The red lines indicate the mean value of the scatter for each direction, and the shaded area encompasses one standard deviation of the distribution.

4.7 Results

4.7.1 Modulation amplitude

The marginalized over the modulation direction, probability distributions (posteriors) of the modulation amplitude parameter as derived from the WMAP V5 and ILC5 data are plotted in figures 4.6 and 4.7 respectively for different ranges of filtered multipoles (see table 4.5.4). We remind that only the analysis involving the V5 data was performed using the KQ75 sky mask.

Table 4.2 summarizes the modal values of the distributions as parameter estimates and the 68% and 95% confidence level limits.

In Table 4.3 we concisely summarize the results of the modulation significance analysis: ie. the analysis in which we derive the minimal confidence levels, at which the modulation value of $A = 0$ cannot be excluded. We specify the expectancy values, mode values of the distributions, and the corresponding significance. We choose to calculate the confidence intervals - or rather, since we're working on posterior probability distributions, in the nomenclature of the Bayesian language, the credibility intervals, by integrating from the modal value, rather than from the expectancy value.

To visualize these results, we plot the estimated modal values of the posterior distributions as a function of considered $\ell_{\min}$ and $\ell_{\max}$ values (Fig. 4.8). For each multipole bin, in Fig. 4.8 we also indicate the minimal confidence level (see Table 4.3) it takes to keep the modulation $A = 0$ i.e. the non-modulated, isotropic model, as an viable option.

It is clear that the modulation amplitude depends on the considered multipole range and hence on scale. Is is generally seen that large values of the best-fit modulations mostly come from the large scales, while for high multipole bins, the best-fit modulations are much smaller. As explained in section 4.6, within GRF realizations this is somewhat expected due to the nature of cosmic variance effects.

Looking at distributions in Figs. 4.6 and 4.7, it is apparent that the modulation for $\ell \in [7, 19]$ (and also $\ell \in [3, 19]$), in the V5 data, (blue curves (crosses) in Fig. 4.6) is strongest and most significant, as it takes the confidence level as high as $\sim 99.5\%$ to include the $A = 0$ value (see.Table 4.3 and Fig. 4.8). The appearance of some asymmetry in this range also seems consistent with the results presented in Section 4.4. The range of multipoles $\ell \in [20, 29]$ of the V5 data does not seem to prefer any modulation value as its modal value is almost zero, and consequently while increasing values of $\ell_{\max}$, for $\ell_{\min} = 7$, the overall significance falls, as this multipoles bin is included, but then systematically increases as larger multipoles are added, which is consistent with the shape of the PDF functions preferring some non-zero modulation for higher multipole bins like $\ell \in [30, 39]$.

The best-fit modulations for $\ell \in [20, 39]$ and $\ell \in [30, 39]$ range from $A = 0.07$ to $A = 0.10$ and exclude the isotropic model, depending on the data, at confidence level of about 96% to 99%, while the best-fit modulation for $\ell \in [7, 39]$ in the ILC5 data, with the modal amplitude of $A = 0.13$, exclude the isotropic model, at confidence level as high as

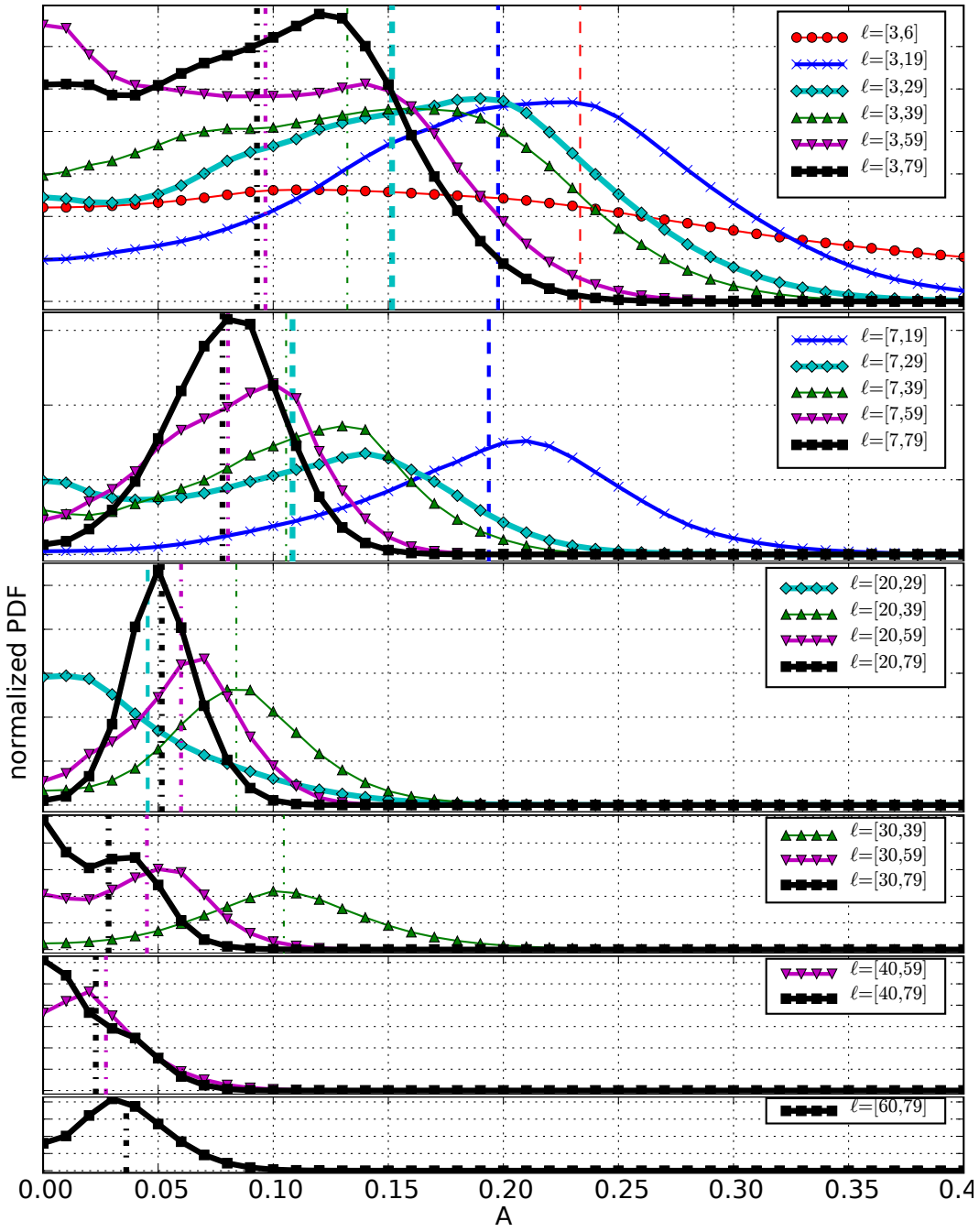


Figure 4.6: Constraints on modulation-amplitude parameters as a function of considered multipole bin from the V5 data. From the bottom to the top we plot data with less filtering of the large scale multipoles. For each PDF the corresponding expectancy value is marked by vertical dashed lines for bins $\ell \in [3, 6]$, $\ell \in [3, 19]$, $\ell \in [3, 29]$, and dash-dotted line for bins $\ell \in [3, 39]$, $\ell \in [3, 59]$ and $\ell \in [3, 79]$. Within each group the increasing line width corresponds to increasing value of $\ell_{\max}$. Only every 100'th point of the interpolated, marginalized PDF was plotted, so the data points do not correspond to the actual grid nodes. We have truncated the plot at A=0.4 to maximally expose the most interesting regions, while keeping the same scale throughout all panels.

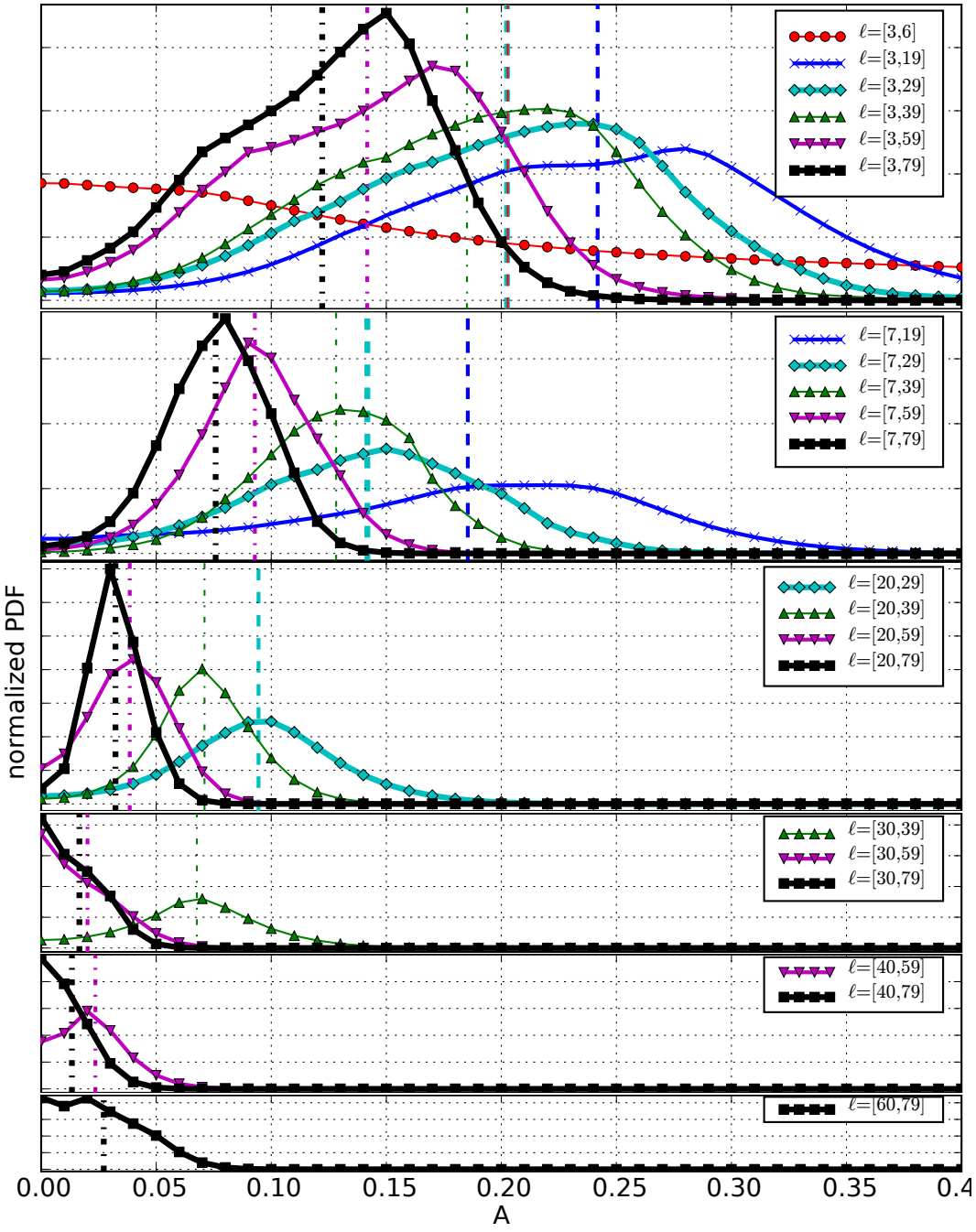


Figure 4.7: Constraints on modulation-amplitude parameter as a function of filtered multipole bin from the ILC5 data. Details as in Fig. 4.6.

99.9%. We will further test the significance of these results in section 4.7.3.

Note also that, some of the marginalized PDFs exhibit bi-modality (eg. the PDF corresponding to the multipole range $\ell \in [3, 59]$ – magenta-line ($\triangledown$) in Fig. 4.6). This bi-modality results from the marginalization itself, and is not observed in the full non-marginalized distribution. Since the likelihood function does not depend on the orientation of the modulation axis for the modulation amplitude $A = 0$, while it does depend on the modulation orienta-

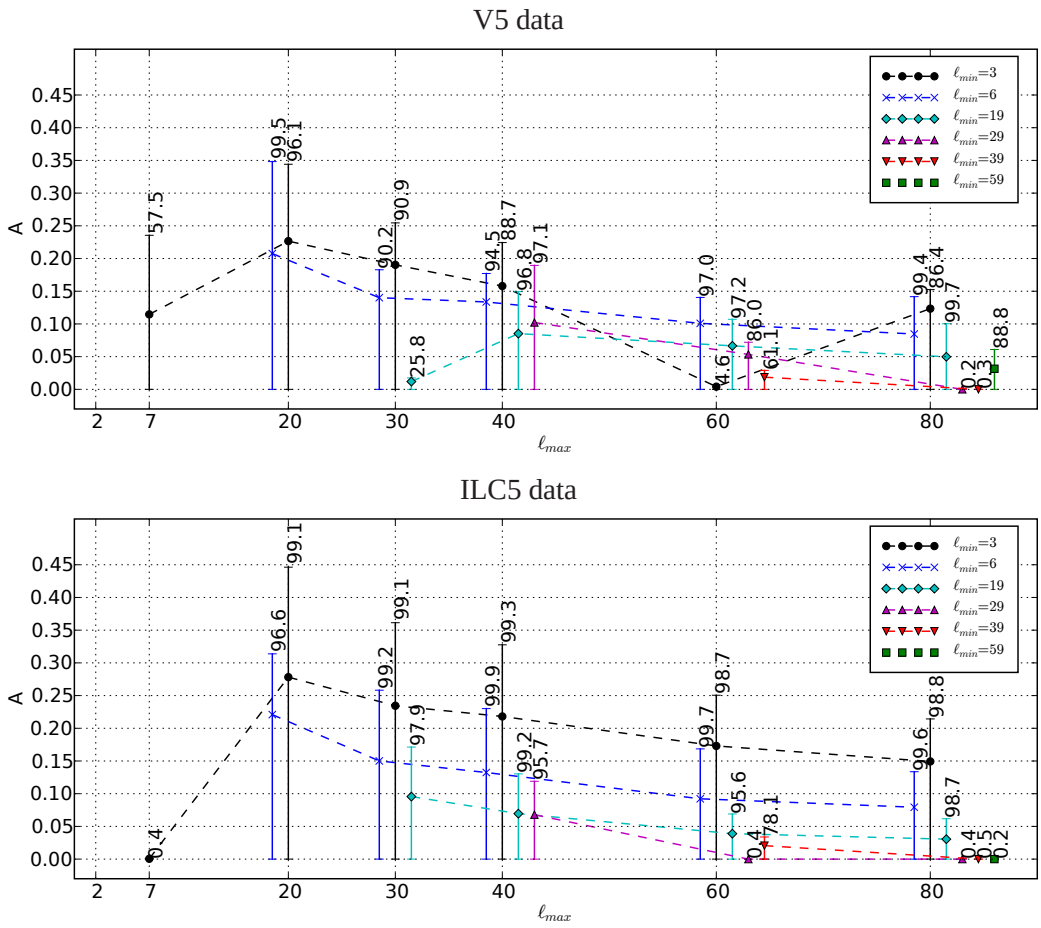


Figure 4.8: Results of the modulation amplitude parameter estimation from the V5 data (top panel) and ILC5 data (bottom panel). Modal values of the posterior distributions are plotted for different bins of multipoles filtered out from the dataset. The error bars represent the confidence limits that include the $A = 0$ value. The corresponding confidence levels (given in percents) are given as annotations by the data points.

tion very strongly for modulations $A \gg 0$, the likelihood surface shall tend to be peaky for large values of A, and flat for $A \approx 0$. As such, depending on how strong the preference of some direction happens to be, it is possible to accumulate in the marginalization process a second peak (the second mode) out of somewhat less-preferred, but constant at certain level, likelihood values along $A \approx 0$ direction. As a result the aforementioned range $\ell \in [3, 59]$ of V5 data yields a small significance in Table 4.3 (in terms of rejecting an isotropic model). We have also processed the results by using maximization over the modulation orientations instead of marginalizations, and as expected, the maximized PDFs are unimodal and more strongly exclude the non-modulated, isotropic models, however we choose the more conservative, and more correct method of marginalizing over the non plotted dimensions.

From Fig. 4.8 it is easy to see that the modulation amplitude estimates are mostly similar between the two datasets, and that generally the amplitude of the modulation decreases

with increasing multipole number ℓ in the two datasets. A common feature between the two datasets is that for high multipole bins $\ell \in [40, 59]$, $\ell \in [40, 79]$ and $\ell \in [60, 79]$ the best-fit modulation amplitudes are small or zero, despite that the amount of variance carried by bin eg. $\ell \in [40, 79]$ is as high as $\sim 27\%$ of the total variance carried by the full range of considered multipoles (i.e. from $\ell = 3$ to $\ell = 79$, see.table 4.5.4). The multipoles range $\ell \in [29, 40]$ instead participate to the total variance only by $\sim 8\%$, and consequently, the the best-fit modulation of this range, estimated to be $A = 0.10$ ($A = 0.07$) for the V5 (ILC5) data, is effectively destroyed, as the higher multipole bins are included, most likely due to simply dominating power in the added multipole bins.

There are few significant differences between the datasets as well. Firstly, we notice, that the ILC5 estimates are generally slightly, but systematically larger from the V5 estimates. Also, in particular, the full sky ILC5 posteriors, for multipole bins $\ell \in [7, 29]$ and $\ell \in [20, 29]$ strongly prefer some non-vanishing modulation amplitude, in contrast to the V5 data (compare cyan-diamonds in Fig. 4.6 and 4.7). This results in an almost constant significance of excluding $A = 0$ as higher multipole bins are being included (see first three rows of the Table 4.3 in section for ILC5 data).

It is interesting to note a small difference in range $\ell \in [3, 6]$ (see table 4.3) in which the ILC5 slightly prefer a vanishing best-fit modulation. Note that among our considered multipole bins, the ILC5 should be reliable basically only within this lowest range. In contrast, some non-zero preferred modulation is obtained with the V5 data, however the value is still largely consistent with the vanishing modulation at confidence level as low as $\sim 58\%$.

Some differences between the datasets are of course expected due to the cut-sky effects, which preclude filtering of exactly the same range of multipoles due to the power leakage effects in case of V5 data. Also caution is needed in the interpretation of the ILC5 data for higher multipole bins, as residual foregrounds in the regions around the Galactic center, may have some impact on the results. In particular, these residual foregrounds might be responsible for the significant alteration of the shape of the PDF function in the multipole bin $\ell \in [20, 29]$ (cyan-diamonds in figures 4.6 and 4.7) towards an increased significance in favor of non-isotropic models.

4.7.2 Modulation orientation

We now focus on the modulation orientation as a function of our chosen multipole bins, as specified in table 4.5.4. The maximum likelihood modulation orientations are summarized in table 4.4

The best-fit orientations found in the analysis depend in general on the considered range of the multipoles. In particular, we see that the hemispherical power asymmetry, as measured here by the modulation orientations, generally tend to shift from larger galactic latitudes to the smaller galactic latitudes, as data of higher multipole bins are processed or cumulatively added up as $\ell_{\max}$ value increases. This was previously also noticed in Lew

(2008b) but using a different method. This effect seems to be seen in the first and second rows (as moving from the left to the right) for each of the data in table 4.4, or by comparing vertically in columns the first three rows as the low multipoles are removed from the analysis. This is also directly seen in the distributions plotted in Fig. 4.9, where it can be easily deduced in which bins the effects of small or vanishing modulations can be ignored. However we note that the effect is not present in every bin, and at most of order of few tens of degrees, and therefore we refrain from making any far-reaching speculations based on it.

While analyzing the distributions in Fig. 4.9 note that, those of them that correspond to very small or vanishing modal values of modulation amplitudes, have a very extended confidence level contours that cover large fraction of the sphere. These naturally result from a very flat likelihood function, and therefore any inference based on these cases is largely speculative and irrelevant. This is however expected, because for a vanishing modulation amplitude, there is no information on its orientation either.

We note that it is possible there is some degree of correlation between the plots for a given row, resulting from an cumulative effect of adding higher multipole bins. The possible changes to the resulting distribution will jointly depend on the modulation amplitude and orientation in the added bin, but also on the amount of variance carried by the that bin as specified in table 4.5.4.

In table 4.4 we plot the directions, in galactic coordinates, of the maximum posterior values found in the modulation orientation analysis. Out of curiosity we also provide the angular separation of these directions from the ecliptic south pole to check for any possible extra alignments.

4.7.3 Modulation significance

In the previous section we have shown that, for some multipole ranges, the reconstructed, marginalized probability distribution function of the modulation amplitude, excludes the vanishing modulation value ($A = 0$) at a very high confidence level. It is important to ask whether this result is really robust, and whether or not we should reject the standard isotropic model of the Universe, at least, at some of the scales: i.e. those corresponding to the distributions with the strongest modulation detections, and the highest non-zero modulation significances. In particular, at least three ranges are of most concern: $\ell \in [7, 19]$, and $\ell \in [7, 79]$ for which the $A = 0$ can be excluded at 99.5% and 99.4% CL respectively, using the KQ75 sky-cut V5 data, and where the modulation parameters are constrained to be within ranges $(0.07)0.14 < 0.21 < 0.26(0.31)$ at 68% (95%) CL and $(0.02)0.05 < 0.08 < 0.11(0.13)$ at 68% (95%) CL respectively. Also the aforementioned range $\ell \in [7, 39]$, for which the $A = 0$ can be excluded at 99.9% CL using the full-sky ILC5 data, and where the modulation parameter is constrained to be within range $(0.06)0.10 < 0.13 < 0.17(0.20)$ at 68% (95%) CL.

What we have done in the previous sections, is that we have estimated the best-fit modulation parameters (amplitude and orientation) with respect to the average from large amount of GRF simulations. Using average from large number of simulations ensures that we com-

pare the data to really isotropic distributions, as any deviations from the statistical isotropy, even those resulting from the cosmic variance, will be averaged out. Although the measurements also quantified the allowed magnitude of deviation from the ideal isotropy, allowed within the cosmic variance, via the standard deviation in the χ^2 tests, there were number of explicit, or implicit assumptions or simplifications made on the way, like for example the usage of the χ^2 distributions, or neglecting the cosmic covariance effects, or residual foregrounds to name few.

In Lew (2008b) we have performed the full covariance matrix analysis in two hemispherical regions with the same search parameter space as detailed in Section 4.5.4. We used 1000 GRF simulations (500 for covariance matrix estimation and 500 for probing the PDF of the underlying χ^2 distribution) and another 1000 simulations, modulated with amplitude of $A = 0.114$. The simulations were filtered up to $\ell_{\rm max} = 40$. We found that on average about 8% of the GRF simulations exhibited a more unusual power distributions, than those found in the modulated simulations. Here, although we use the same density in the search space, we have improved somewhat the method by using smooth interpolations.

In order to further test the robustness, and the significance of the power asymmetry anomalies, and circumvent all possible imperfections of the method, in the following we will pursue a similar test. We process 100 GRF simulations of the V5 data through our parameter estimation pipeline, and compare the results with the real data. Such approach should always be an ultimate test of the robustness, as it must give correct results independently from the assumptions taken in the method.

We will focus on the aforementioned multipole bins: $\ell \in [7, 19]$ and $\ell \in [7, 79]$ of the V5 data, and $\ell \in [7, 39]$ of the ILC5 data. Within these ranges the power asymmetry seems to be very strong and very significant (see. table 4.3 and figures 4.6 and 4.7).

In Fig. 4.10 we plot the results of the modulation parameter estimation for all tested simulations along with the WMAP data. While it is clear that most of the simulations do not prefer any significant, if any, modulation amplitude values, at least few simulations, in our sample, yield modulations that are stronger in the considered range of multipoles, than those found in the data. Also, from the shape of the PDF it is easy to infer that the significance of rejecting $A = 0$ in these few cases will be even larger than in the case of the selected, most anomalous results from the WMAP data.

We find that 7 out of 100 V5 simulations yield stronger best-fit (modal) modulation amplitudes, and 6 of them also yield a more significant rejection of the $A = 0$ parameter value, than the V5 data in the range $\ell \in [7, 19]$. Similarly, for the range $\ell \in [7, 79]$ 5 simulations yield stronger and more significant best-fit modulation values.

Consequently, we conclude, that the significance, as inferred simply from integrating the PDF (as given in table 4.3) is not quite robust. In light of these results we estimate the overall significance of possible modulation signals in the analyzed WMAP CMB maps at the level of about $\sim 94\%$ to $\sim 95\%$ depending on the particular range of multipoles. This remains greatly consistent with our previous results reported in Lew (2008b) for the same data.

As for the ILC5 data we find that three simulations out of 100 yield a more significant rejection of $A = 0$ hypothesis, and curiously, only one in those three also yields a stronger modulation within the multipole range $\ell \in [7, 39]$. Therefore, the corresponding overall significance of the power asymmetry, in this particular multipole range, is still as high as about $\sim 99\%$. We note however, that since we did not use any sky masks in this case, this result can probably be safely considered as somewhat overestimated, as any residual galactic foregrounds are likely only to increase the level of the hemispherical power asymmetry, rather decrease it.

Table 4.2: Constraints on the modulation amplitude parameter for the V5 and ILC5 data. The table contain the modal parameter values and the corresponding 68% and 95% (in brackets) confidence level limits.

V5 data

$\ell_{\rm min}\backslash\ell_{\rm max}$	7	20	30	40	60	80
2	(0.00) $0.00 < 0.11 < 0.29$ (0.55)	(0.01) $0.12 < 0.23 < 0.29$ (0.34)	(0.00) $0.07 < 0.19 < 0.24$ (0.28)	(0.00) $0.05 < 0.16 < 0.21$ (0.25)	(0.00) $0.00 < 0.00 < 0.15$ (0.20)	(0.00) $0.00 < 0.12 < 0.15$ (0.18)
6		(0.07) $0.14 < 0.21 < 0.26$ (0.31)	(0.00) $0.00 < 0.14 < 0.18$ (0.20)	(0.00) $0.06 < 0.13 < 0.16$ (0.18)	(0.01) $0.05 < 0.10 < 0.12$ (0.14)	(0.02) $0.05 < 0.08 < 0.11$ (0.13)
19			(0.00) $0.00 < 0.01 < 0.06$ (0.12)	(0.01) $0.05 < 0.09 < 0.12$ (0.15)	(0.01) $0.04 < 0.07 < 0.09$ (0.11)	(0.02) $0.04 < 0.05 < 0.07$ (0.09)
29				(0.01) $0.06 < 0.10 < 0.14$ (0.19)	(0.00) $0.00 < 0.05 < 0.07$ (0.09)	(0.00) $0.00 < 0.00 < 0.04$ (0.06)
39					(0.00) $0.00 < 0.02 < 0.03$ (0.07)	(0.00) $0.00 < 0.00 < 0.03$ (0.06)
59						(0.00) $0.01 < 0.03 < 0.05$ (0.07)

ILC5 data

$\ell_{\rm min}\backslash\ell_{\rm max}$	7	20	30	40	60	80
2	(0.00) $0.00 < 0.00 < 0.26$ (0.52)	(0.07) $0.16 < 0.28 < 0.33$ (0.41)	(0.06) $0.14 < 0.23 < 0.28$ (0.33)	(0.06) $0.13 < 0.22 < 0.26$ (0.30)	(0.03) $0.09 < 0.17 < 0.20$ (0.24)	(0.03) $0.08 < 0.15 < 0.18$ (0.20)
6		(0.01) $0.12 < 0.22 < 0.28$ (0.31)	(0.04) $0.09 < 0.15 < 0.20$ (0.24)	(0.06) $0.10 < 0.13 < 0.17$ (0.20)	(0.04) $0.07 < 0.09 < 0.12$ (0.15)	(0.03) $0.05 < 0.08 < 0.10$ (0.12)
19			(0.02) $0.06 < 0.10 < 0.13$ (0.17)	(0.02) $0.05 < 0.07 < 0.09$ (0.12)	(0.00) $0.02 < 0.04 < 0.06$ (0.07)	(0.01) $0.02 < 0.03 < 0.04$ (0.06)
29				(0.00) $0.04 < 0.07 < 0.10$ (0.12)	(0.00) $0.00 < 0.00 < 0.03$ (0.05)	(0.00) $0.00 < 0.00 < 0.02$ (0.04)
39					(0.00) $0.01 < 0.02 < 0.03$ (0.05)	(0.00) $0.00 < 0.00 < 0.02$ (0.03)
59						(0.00) $0.00 < 0.00 < 0.03$ (0.06)

Table 4.3: Results of the modulation amplitude parameter estimation for the V5 and ILC5 dataset. The table contain the minimal confidence levels (in percents) at which the parameter value of $A = 0$ cannot be excluded (bold face numbers) and the expectancy (in round brackets) and the modal (in square brackets) values of the corresponding distributions. See also Fig. 4.8.

V5 data

$\ell_{\min}\backslash\ell_{\max}$	7	20	30	40	60	80
2	**57.5** (0.23) [0.11]	**96.1** (0.20) [0.23]	**90.9** (0.15) [0.19]	**88.7** (0.13) [0.16]	**4.6** (0.10) [0.00]	**86.4** (0.09) [0.12]
6		**99.5** (0.19) [0.21]	**90.2** (0.11) [0.14]	**94.5** (0.11) [0.13]	**97.0** (0.08) [0.10]	**99.4** (0.08) [0.08]
19			**25.8** (0.05) [0.01]	**96.8** (0.08) [0.09]	**97.2** (0.06) [0.07]	**99.7** (0.05) [0.05]
29				**97.1** (0.10) [0.10]	**86.0** (0.05) [0.05]	**0.2** (0.03) [0.00]
39					**61.1** (0.03) [0.02]	**0.3** (0.02) [0.00]
59						**88.8** (0.04) [0.03]

ILC5 data

$\ell_{\min}\backslash\ell_{\max}$	7	20	30	40	60	80
2	**0.4** (0.20) [0.00]	**99.1** (0.24) [0.28]	**99.1** (0.20) [0.23]	**99.3** (0.19) [0.22]	**98.7** (0.14) [0.17]	**98.8** (0.12) [0.15]
6		**96.6** (0.19) [0.22]	**99.2** (0.14) [0.15]	**99.9** (0.13) [0.13]	**99.7** (0.09) [0.09]	**99.6** (0.08) [0.08]
19			**97.9** (0.09) [0.10]	**99.2** (0.07) [0.07]	**95.6** (0.04) [0.04]	**98.7** (0.03) [0.03]
29				**95.7** (0.07) [0.07]	**0.4** (0.02) [0.00]	**0.4** (0.02) [0.00]
39					**78.1** (0.02) [0.02]	**0.5** (0.01) [0.00]
59						**0.2** (0.03) [0.00]

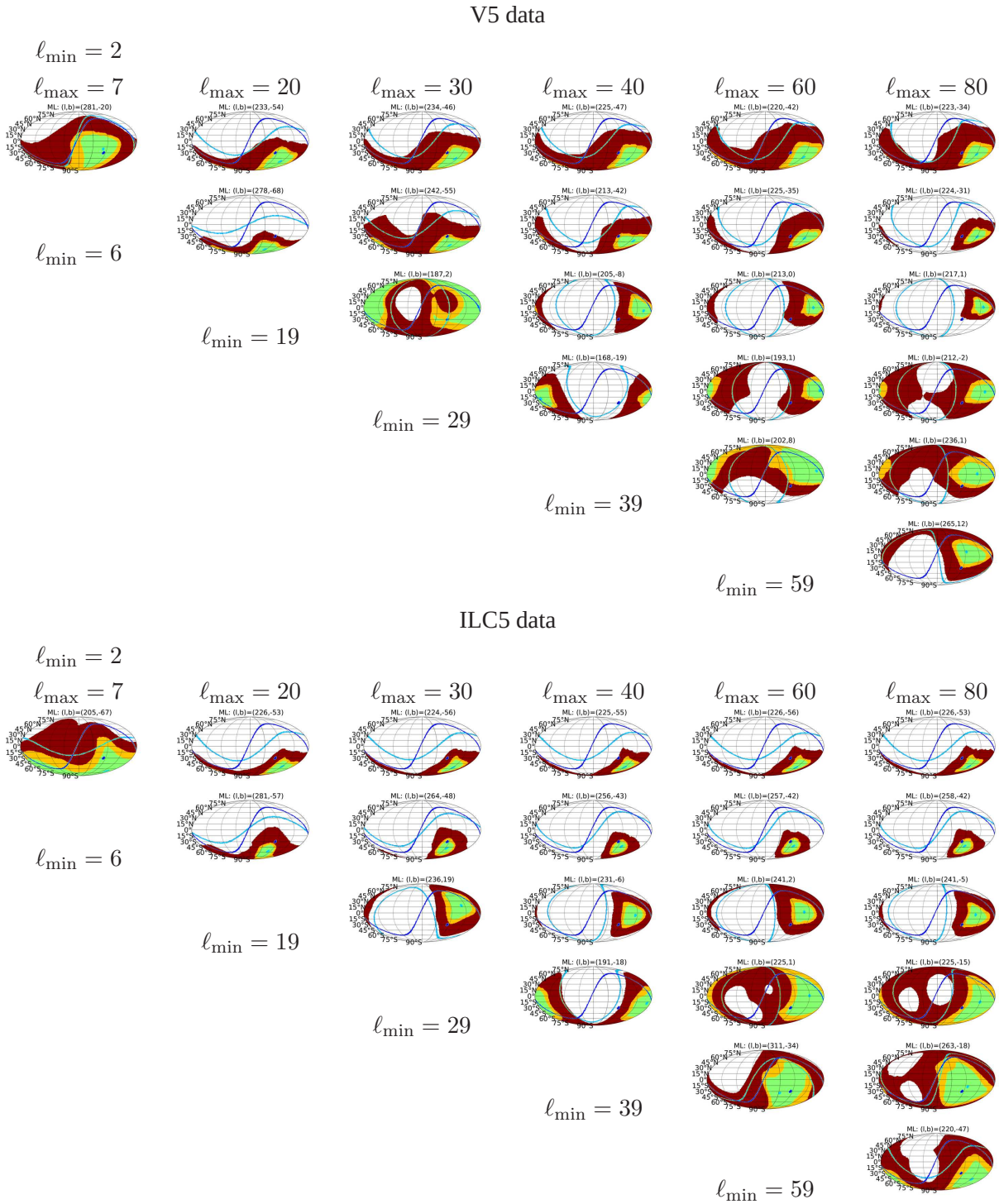


Figure 4.9: Constraints on the modulation orientation from the V5 and ILC5 data. For each considered multipole bin we plot the confidence regions corresponding to the 50% (green), 68% (yellow), and 95% (red) confidence levels, based on the interpolated maps of the posterior distributions. In each map the maximum likelihood orientation and the corresponding dipole plane are indicated using a light blue dot and line respectively. Additionally, for comparison, the ecliptic south pole and ecliptic plane are plotted in dark blue. In the top of each panel we give the galactic coordinates of the maximum likelihood solution. The arrangement of the panels is consistent with the cells in tables 4.5.4, 4.2, and 4.3.

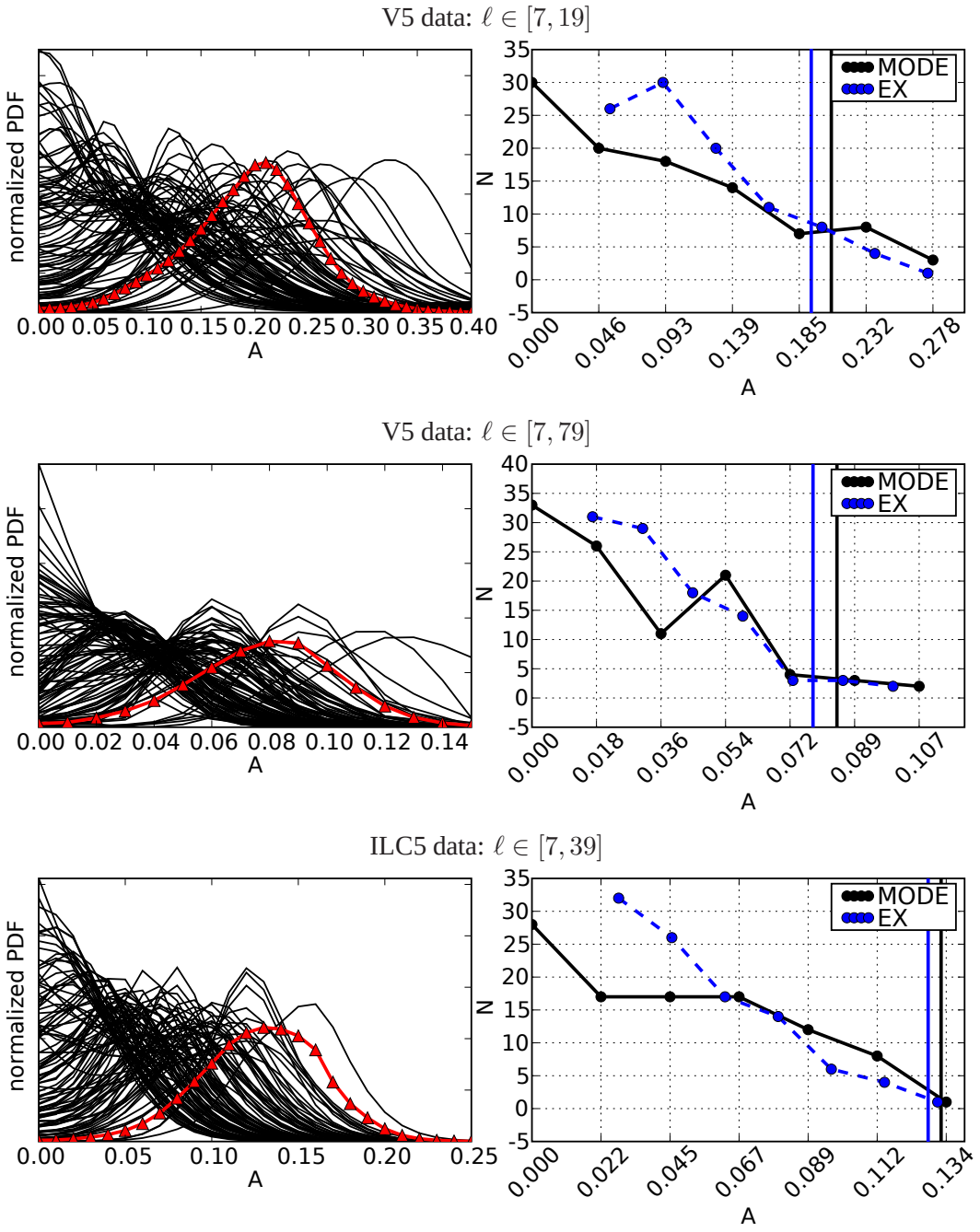


Figure 4.10: On the left-hand side panels we plot a comparison of the reconstructed, marginalized distributions of the modulation amplitudes from 100 simulations of the WMAP V5 data (top and middle panels) and WMAP ILC5 data (bottom panel). The WMAP data are plotted using red, thick lines (triangles). Only every 100th point of the interpolated, marginalized PDF is plotted. On the right-hand side panels we plot the corresponding histograms of the expectancy values and modal values derived from these distributions. The WMAP data values are marked with vertical lines. The plotted ranges yield the strongest, and most significant, plausible hemispherical anomalies in the data as inferred from the analysis in section 4.7.1.

Table 4.4: Results of the modulation orientation parameter estimation for the V5 and ILC5 data. The table contains the galactic coordinates of the maximum posterior modulation orientation and (in square brackets) the relative angular distance to the south ecliptic pole at $(l, b) = (276.4°, -29.8°)$.

V5 data

$\ell_{\min}\backslash\ell_{\max}$	7	20	30	40	60	80
2	$(281°, -20°)$, $[11°]$	$(233°, -54°)$, $[39°]$	$(234°, -46°)$, $[37°]$	$(225°, -47°)$, $[43°]$	$(220°, -42°)$, $[46°]$	$(223°, -34°)$, $[45°]$
6		$(278°, -68°)$, $[38°]$	$(242°, -55°)$, $[35°]$	$(213°, -42°)$, $[52°]$	$(225°, -35°)$, $[43°]$	$(224°, -31°)$, $[45°]$
19			$(187°, 2°)$, $[90°]$	$(205°, -8°)$, $[70°]$	$(213°, 0°)$, $[67°]$	$(217°, 1°)$, $[64°]$
29				$(168°, -19°)$, $[96°]$	$(193°, 1°)$, $[85°]$	$(212°, -2°)$, $[67°]$
39					$(202°, 8°)$, $[81°]$	$(236°, 1°)$, $[49°]$
59						$(265°, 12°)$, $[43°]$

ILC5 data

$\ell_{\min}\backslash\ell_{\max}$	7	20	30	40	60	80
2	$(205°, -67°)$, $[56°]$	$(226°, -53°)$, $[43°]$	$(224°, -56°)$, $[45°]$	$(225°, -55°)$, $[44°]$	$(226°, -56°)$, $[44°]$	$(226°, -53°)$, $[43°]$
6		$(281°, -57°)$, $[27°]$	$(264°, -48°)$, $[21°]$	$(256°, -43°)$, $[21°]$	$(257°, -42°)$, $[20°]$	$(258°, -42°)$, $[19°]$
19			$(236°, 19°)$, $[62°]$	$(231°, -6°)$, $[49°]$	$(241°, 2°)$, $[46°]$	$(241°, -5°)$, $[42°]$
29				$(191°, -18°)$, $[77°]$	$(225°, 1°)$, $[58°]$	$(225°, -15°)$, $[49°]$
39					$(311°, -34°)$, $[30°]$	$(263°, -18°)$, $[17°]$
59						$(220°, -47°)$, $[46°]$

4.8 Discussion

The results outlined in sections 4.7 indicate that different multipole ranges yield a different best fit modulation value, and that the modulation orientation also slightly varies from one multipole range to another.

In particular, the best-fit modulation orientation dependence when higher multipole bins are included cumulatively is not as strong, as when the added multipole bins are considered individually. Generally, we notice that within the best-fit orientations, that also yield a large modulation values ($A \gtrsim 0.1$) the high-ℓ multipole bins prefer a close galactic plane orientation, while the low-ℓ multipole bins rather prefer orientations with larger galactic latitudes.

The analysis of the modulation amplitude within few multipole bins yielded a large, best-fit modulation amplitudes, that seem to significantly reject the isotropic Universe model (with $A = 0$). However as much as few in one hundred GRF simulations, processed as data, also yielded a similar or larger modulation values, and also excluded the $A = 0$ hypothesis at yet even higher confidence levels, than in the case of the V5 data. This effectively reduces the overall significance at which the isotropic model of the Universe can be rejected, down to only about 94% or $\sim 95\%$ using the V5 data in the range $\ell \in [7, 19]$, and $\ell \in [7, 79]$ respectively.

We therefore pursued the analysis of the modulation signals in a two partially complementary ways. While the first approach addresses the question of "*how large and how significant is the best-fit modulation of the data?*", the second approach quantifies "*how consistent is the best fit-modulation as compared with the GRF simulation expectations?*". The second approach should be more robust since it is free of any, possibly inaccurate, assumptions that could result in underestimation of the size of the errors in the statistic, and in the result lead to spurious detections. These problems are effectively eliminated in a direct comparison with the GRF simulations.

Curiously the ILC5 data in the multipole range $\ell \in [7, 39]$ still seem to be anomalous at a high CL of about 99%; level almost as high as quoted in (Eriksen et al., 2007). However contrary to that work, we have not applied any sky masks to this data, and therefore these results, given here only for comparison purposes with the V5 data, should still be treated with caution.

It would be interesting to perform similar analysis using the ILC5 data but with included sky cut, and to check the dependence of the analysis while varying the sky cuts from less to more aggressive. Also, it could be interesting to check the robustness and the significance in other multipole ranges than those two, tested in section 4.7.3. In principle, it would also be interesting to include other available renditions of the ILC maps, to see the stability of the modulation to different foregrounds cleaning pipelines. We defer these issues for possible future work.

4.9 Conclusions

We performed tests of the hemispherical power asymmetry found in the CMB WMAP data for different bins of multipoles in two ways.

At first, we introduced a statistics that searches for the orientation of two opposing, hemispherical regions that maximize or minimize the hemispherical variance ratio, and compared these with the expectations from the GRF simulations. We found that the maximal asymmetry revealed this way is found within a multipole range $\ell \in [8, 15]$, with the southern hemisphere having larger variance than the northern hemisphere. When these results are compared to the GRF simulations, the northern hemisphere appears to be suppressed below the average expectation.

Secondly, we have introduced and tested a new method for measuring the power asymmetry in the CMB data, as quantified within a bipolar modulation model (Gordon et al., 2005). For the first time we constrained the modulation parameters as a function of various multipole bins. For each multipole range, we obtained the constraints on the the modulation amplitude and orientation. Based on the analyzed, up to the maximal multipole $\ell_{\rm max} = 80$, datasets i.e. the WMAP five-year inverse noise co-added, KQ75 sky cut map from the V channel (V5), and the five-year, full-sky, foregrounds cut ILC map (ILC5) we found that:
(i) generally the modulation amplitude decays as higher multipole bins are cumulatively added or independently analyzed,
(ii) the best fit modulation amplitude is small $A < 0.03$ and insignificant for multipoles beyond $\ell \approx 40$
(iii) the most anomalous signals in terms of the modulation amplitude and its significance come from multipole range $\ell \in [7, 19]$, and $\ell \in [7, 39]$ in the V5 and ILC5 data respectively. For these ranges the significances of rejecting the isotropic cosmological model are 99.5% and 99.9% respectively and the constraints on the best fit, (PDF modal) modulation values are: $(0.07)0.14 < 0.21 < 0.26(0.031)$ and $(0.06)0.10 < 0.13 < 0.17(0.20)$ at 68% (95%) CL respectively.

Focusing on the two selected multipole ranges we performed an additional tests of the significance using GRF simulations processed as data, and found that similar or stronger and more significant (in terms of rejecting the isotropic model) modulation values are obtained in 6 (1) cases in 100 simulations, which decreases the overall significance of the power asymmetry in the CMB down to 94% (99%) in V5 (ILC5) data respectively. To complement the results in the limit of high multipoles as well, we additionally tested the range $\ell \in [7, 79]$ of the V5 data that also yields a strong and significant (99.4%) best-fit modulation value - $(0.02)0.05 < 0.08 < 0.11(0.13)$ at 68% (95%) CL - but when this result was compared with the GRF simulations the effective significance is again decreased down to about 95%.

Although the significance in case of the ILC5 data is still rather high, we warn that the results in this case were obtained without any sky cut, and therefore the asymmetry significance can be overestimated due to residual foregrounds.

Finally we note that a further analysis of the significance in terms of comparison with

GRF simulations of other multipole ranges would be interesting, as well as analysis of the power asymmetry in the ILC data as a function of different sky cuts.

Chapter 5

Gaussianity tests using Minkowski functionals: a high significance, large scale non-Gaussianity detection in the WMAP CMB data.

In this chapter we focus on Gaussianity tests, optimized for searching the residual foreground contamination in the CMBR WMAP data. The presented results are part of the publication currently being in preparation.

5.1 Abstract

We perform a cut-sky Gaussianity analysis of the WMAP five-year, foreground reduced CMBR data using Minkowski functionals. By applying a band-pass filters in the spherical harmonic space, we analyze the maps exclusively in selected bins of multipoles. This effectively helps to test particular scales, which otherwise, given an unequal power distribution across the multipole scale, are dominated by strong signals from outside the bin. This way we complement the previous works, which utilized averaging of different scales by either: degrading the map resolution, or by applying "low-ℓ"-pass Gaussian filters.

We find the data inconsistent with Gaussian random field simulations (GRF) at confidence level at least 99.5%, due anomalous, negatively skewed temperature distribution covering large areas of the sky, in the multipole range $\ell \in (32, 128]$. We further test the robustness of these results in a few additional consistency checks. In particular, we find the anomaly to be associated with a small, but realized by a large number of pixels, excess of positive temperature pixels over the negative ones, and we find that this effect is decreasing for larger galactic latitudes, and is vanishing when more aggressive galactic plane masks are applied, which hints on small, but extended over a large area, residual foregrounds. Finally, we note that the anomaly would have been missed if the analysis involved only smoothed

or downgraded in resolution maps.

5.2 Introduction

Measuring levels of the primordial non-Gaussianity (NG) is one of the primary goals of the future cosmological surveys such as the PLANCK surveyor[1] designed for measurements of the cosmic microwave background radiation (CMBR) anisotropies, or LOFAR[2] designed for measurements of the of the primordial hydrogen distribution. These missions will provide a high sensitivity and resolution cosmological observations of high cosmological significance. The amplitude and shape of the primordial non-Gaussianity will become a very useful observable allowing to distinguish between different viable inflationary models (eg.Gangui et al. (2002); Rigopoulos et al. (2007); Barnaby & Cline (2008); Chen et al. (2007b); Holman & Tolley (2008); Enqvist & Takahashi (2008); Chen et al. (2007a); Sasaki (2008); Acquaviva et al. (2003); Bernardeau et al. (2006); Bartolo et al. (2004)). The CMBR will serve as one of the best windows on these processes, spanning through orders of scales, from close-horizon scales $\sim 10^4$ Mpc, all the way down to ~ 10Mpc scales, where non-linear gravitational collapse effects strongly dominate the primordial density fluctuations. Any primordial NG in the gravitational potential field is be imprinted on the CMB long before it decouples from interactions with baryon fluid at the time close to the recombination, and thereby it becomes accessible for direct measurements.

While the consistency of the (CMBR) data with a Gaussian random field (GRF) hypothesis has been extensively tested in the previous several years, using a whole battery of various statistical tests, (Park et al. (2006); Copi et al. (2008); Lew (2008a,b); Efstathiou (2004); Gordon (2007); Gordon & Trotta (2007); de Oliveira-Costa & Tegmark (2006); Copi et al. (2006b, 2004); Gaztañaga & Wagg (2003); Hansen et al. (2004b); Souradeep et al. (2006); Cabella et al. (2004)) the measurements of levels of the primordial NG only in the recent years started to be directly testable, (de Troia et al. (2007); Curto et al. (2008b,a); Chen & Szapudi (2006a); Cabella et al. (2005); Cayón et al. (2003); Yadav & Wandelt (2008); Hikage et al. (2006)), still, however with a precision that will have to be improved over the next years, as more precise observations of the CMB anisotropies will become available; in order to yield robust tests of the physics of the early Universe.

In the meantime it is important to study both the NG, and a degree of the consistency of the data with the GRF hypothesis via eg. realistic Monte-Carlo (MC) simulations. In the latter case, if a robust localization of a deviation is possible, such analyses would help to identify and eliminate the NG signals, that most likely will be of non-cosmological origin, but would also possibly confuse the estimates of the primordial NG.

In this paper we study the CMB WMAP maps using Minkowski functionals (MF), and also extend the and analysis in a variety of ways to cross-check the results. Several authors have already pursued a similar statistics (Colley & Gott (2003); Wu et al. (2001); Park

[1] http://www.rssd.esa.int/PLANCK

[2] http://www.lofar.org/

(2004); Eriksen et al. (2004b); Komatsu et al. (2008)) but most of such analyses operated on either: scrambled through many orders of resolutions maps, or smoothed maps in spherical harmonics space, in order to test the Gaussianity hypotheses at different scales.

However since the power (the variance per multipole) in the map is unequally distributed across the multipole scale (due to certain shape of the underlying power spectrum), it might be useful, when testing different angular scales, to eliminate the impact of the largest, and most uncertain (due to the cosmic variance effects) scales, in order to focus on the smaller scales exclusively, and perform an analysis, that will not be dominated, or obscured by the effects large scale multipoles. In this paper we intend to explore this idea, and we perform a statistics in different ranges of multipoles, using multipole band-pass filters in the spherical harmonics space prior to the tests.

5.3 Data and Analysis

5.3.1 Maps pre-processing

We utilize the WMAP five-year foreground reduced CMB temperature maps (Hinshaw et al., 2008) from differential assemblies (DA) V1 and V2, because these spectral channels provide the best trade off between foregrounds of different spectral properties (i.e. the blue tilted galactic dust emission and red-tilted galactic synchrotron and free-free emissions). We co-add these observations (and corresponding simulations) using the inverse noise pixel weighting scheme. We will refer to these maps as V5. We generate 3000 Monte-Carlo (MC) simulations, using the fiducial, best-fit ΛCDM model power spectrum of Dunkley et al. (2008) (constructed using the mean likelihood parameters) which we call $C_\ell^{\rm fid}$. We will also analogically utilize the Q and W channel data to trace spectral properties, and we will call them Q5 and W5 respectively. We will use the KQ75 sky mask throughout the analysis. All maps are calibrated by their variance prior the analysis.

5.3.2 Band pass filters

We choose to test the data in all possible ranges of multipoles $\ell \in (\ell_{\rm min}, \ell_{\rm max}]$, defined by: $\ell_{\rm min} \in \{5, 32, 64, 128\}$ and $\ell_{\rm max} \in \{32, 64, 128, 256\}$. We apply the top-hat filters in spherical harmonic space. Note that due to sky cut, these ranges should be treated only as approximations. We chose to remove the large scales ($\ell \leq 5$) from the analysis, because these are well known to have a strong non-Gaussian features, widely discussed in the literature (see. section 5.2).

5.3.3 Statistics

We perform a single region, cut-sky statistics using Minkowski functionals in the data and in the simulations at $N_\nu = 15$ thresholds, uniformly spanning within the $\pm\{3, 3.5, 4, 4\}\sigma$ temperature range depending on the resolution parameter of the analyzed maps n_s, defined as in the Healpix sphere pixelization scheme (Górski et al., 2005); where σ denotes the

standard deviation of a map or a simulation. Sine we will derive the Minkowski functionals (MFs) in the pixel space, in order to increase the reliability of the derivation of the covariant derivatives, the map resolution parameter n_s is always chosen to be equal to the maximal number of the multipole that is to be synthesized in the map. In case of our bins this will correspond to $n_s = \{32, 64, 128, 256\}$ respectively. The MFs are derived following the prescriptions outlined in Schmalzing & Gorski (1998); Hikage et al. (2006). These are also briefly summarized in the appendix A-5.

Since the values of the MFs are correlated between different temperature thresholds we derive the joint probability of the realization of MF in the data via direct comparison with the simulations, accounting for the covariance terms via the standard multivariate calculus as:

$$\chi^2_k = \big(v_{k,\nu_i} - \langle v_{k,\nu_i}\rangle\big)\hat{\mathbf{C}}^{-1}_{k,\nu_i\nu_j}\big(v_{k,\nu_j} - \langle v_{k,\nu_j}\rangle\big) \tag{5.1}$$

where k denotes the MF type. The covariance matrix is calculated from $N_{\rm cov} = 1500$ simulations, while the corresponding distribution of the derived χ^2 values is estimated from the remaining $N_{\chi^2{\rm PDF}} = 1500$ simulations. Therefore, for the $N_\nu = 15$ different temperature thresholds, the number of simulations used for the covariance matrix estimation is 100 times larger than its size, which should grant a sufficient convergence.

We derive the p-values for the data for each multipole bin, and for each MF type, using the linear interpolations in between the points of the probed χ^2 PDF as described in appendix of Lew (2008b).

5.4 Results

The results of the tests performed on the maps in the analyzed multipole bins are given in Fig. 5.1 and the results of the joint statistics for the V5 data are summarized in table 5.1.

In order to test the stability of the presented results in the same table we also attach the results under different combinations of the above mentioned numbers: i.e. $N_{\rm cov}$, $N_{\chi^2{\rm PDF}}$ and N_ν. It is apparent that the results are not particularly sensitive to the change of $N_{\rm cov}$, which suggests that the covariance matrix is sufficiently converged. However, a significant anomaly is seen in the case of circumference statistics in the multipole range $\ell \in (5, 64]$ for the joint statistic where as the corresponding plot in figure 5.1 does not seem to be anomalous. We find that the reason for the alarmingly low p-value in this multipole bin comes from the lowest-threshold data-point, which significantly stands out from the simulations, and builds up the main part of the χ^2 value. For the joint statistics, we will therefore inconsider the outer 2 (4) thresholds: i.e. we considered only the inner 13 (11) threshold data-points. As it is seen in the table 5.1 the low p-value for the range $\ell \in (5, 64]$ is caused by these tail outliers. However, number of pixels inside the threshold centered at $\sim 3.2\sigma$ away from the mean of the distribution for the number of pixels in the map of resolution of $n_s = 64$ and $\ell_{\rm max} = 64$ is very small, and therefore we do not put much attention to this, strong as it is, but possibly insignificant, due to some numerical uncertainties in the covariance matrix, detection. The inverse covariance matrix in this case is dominated just by few

outermost thresholds, and in the following we will not further investigate this deviation, and as far as the joint statistics is concerned, we will exclude the outermost thresholds from the analysis when calculating the joint probabilities.

At first, in the figure 5.1 it is easy to see a strong deviation in the fist Minkowski functional: i.e. describing the total excursion set above a given temperature threshold, in particular for the cases when the lowest order multipoles are filtered out. Most notably, the multipole ranges $\ell \in (32, 128]$, $\ell \in (32, 256]$ and $\ell \in (64, 256]$ and $\ell \in (128, 256]$ are particularly anomalous. This is main result of this paper, which to our knowledge is a new detection unnoticed in the previous analyses.

Generally, it is also easy to see that while adding more and more small scale data, the deviation from the simulation average tends to increase, which is somewhat expected, however it is also clear that the large scales, which here are already removed beforehand up to $\ell 5$, efficiently obscure, an otherwise very strong NG behaviour.

Secondly, it is important to note that the strongest deviations from the simulations appear in the thresholds close to the average temperature, where lie the bulk of the map pixels. This means that the anomaly is not some statistical fluke, but a very strong discrepancy. Furthermore, we notice that such strong deviation appears only in the first Minkowski functional, which measures simply the normalized number of pixels above a given threshold.

Note that the p-values that are smaller than $1/N_{\chi^2\mathrm{PDF}}$, such as the p-value for the range $\ell \in (32, 128]$ for the case of $N_{\chi^2\mathrm{PDF}} = 500$ and $N_\nu = 11$, are obtained using Gaussian extrapolation (as described in Lew (2008b)) and as saturated, should be treated with caution. Generally though the significance of the anomaly which is maximized in this particular range is larger than 99.5% CL.

5.4.1 Anomaly tests

The anomaly revealed in the previous section, in general may be caused in two ways, since the Minkowski functionals are not only sensitive measures of Gaussianity, but they are also a very sensitive to the shape of the underlying power spectrum. In order to exclude the possible biases due to the latter case, in figure 5.2 we plot the pseudo power spectrum of the WMAP V5 data with the cosmic variance 68%,95% and 99% limits from simulations and the simulation mean (top panel) and also the χ^2 value per degree of freedom (bottom panel). The joint probability as measured by the p-value for this consistency check is $\sim 35\%$, which yields a good consistency, however we notice a few strong outliers related to the glitch-like features in the pseudo-power spectrum. This was previously seen in eg. Hinshaw et al. (2006).

Since the anomaly is detected via the area functional it is interesting to plot the temperature distribution within the affected multipole range $\ell \in (32, 128]$. In figure 5.3 we plot the temperature histogram for the map in the detected multipole range outside the KQ75 sky mask. While the negative skewness of the distribution $S = -0.0243$ is not very significant with 98.5% CL to be rejected based on comparison with the 3000 simulations, the mean temperature value of the WMAP V5, in this multipole range, is inconsistent at very high

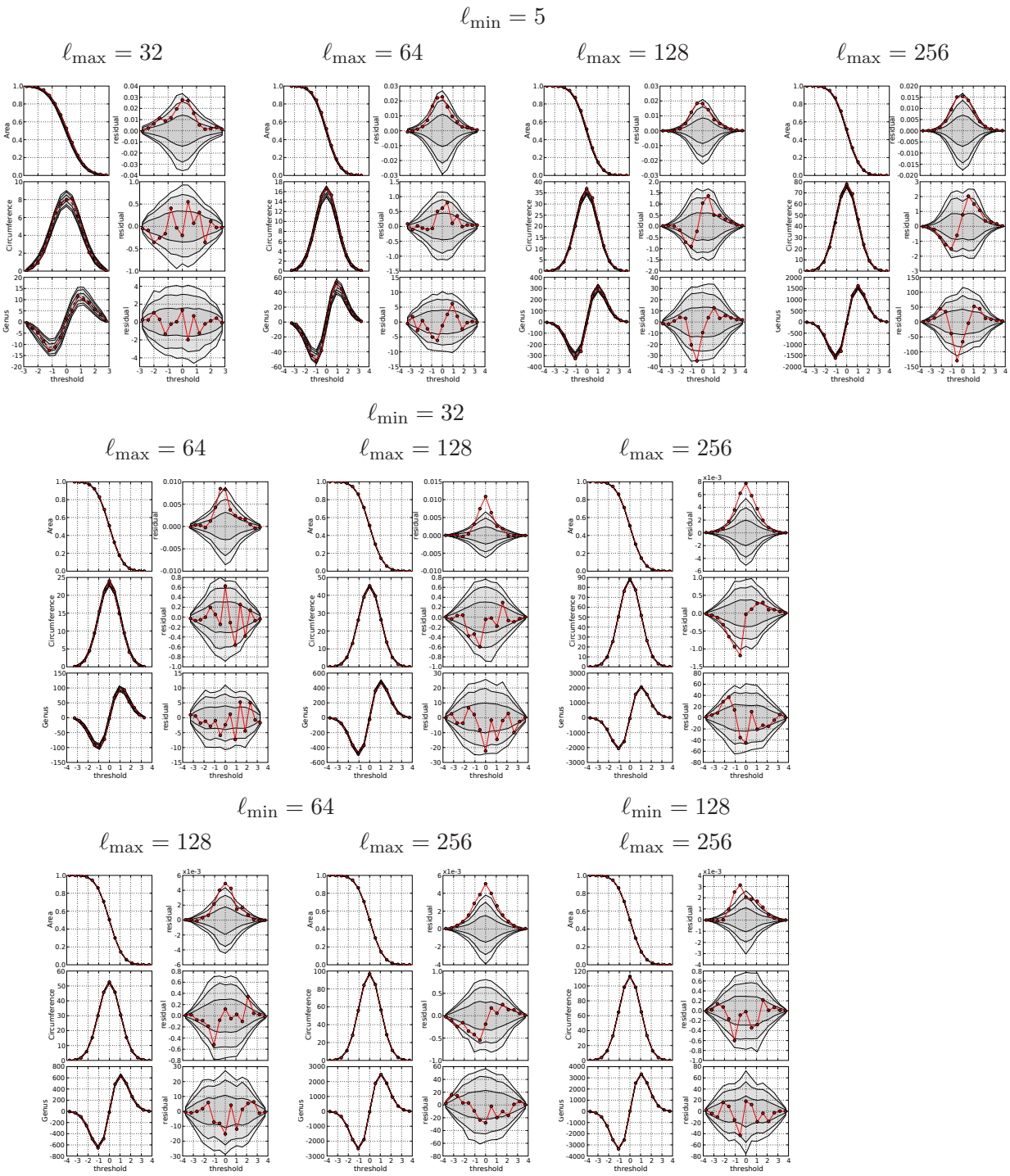


Figure 5.1: Results of the Minkowski functional analysis of the WMAP V5 data. The area, circumference and genus statistics are plotted for different tested multipole bins. For each bin, the left-hand side plots show the functionals themselves, along with the 68%, 95% and 99% significance limits derived from the simulations, and the right hand side plots show the residuals between the data and the simulations average and the same confidence limits.

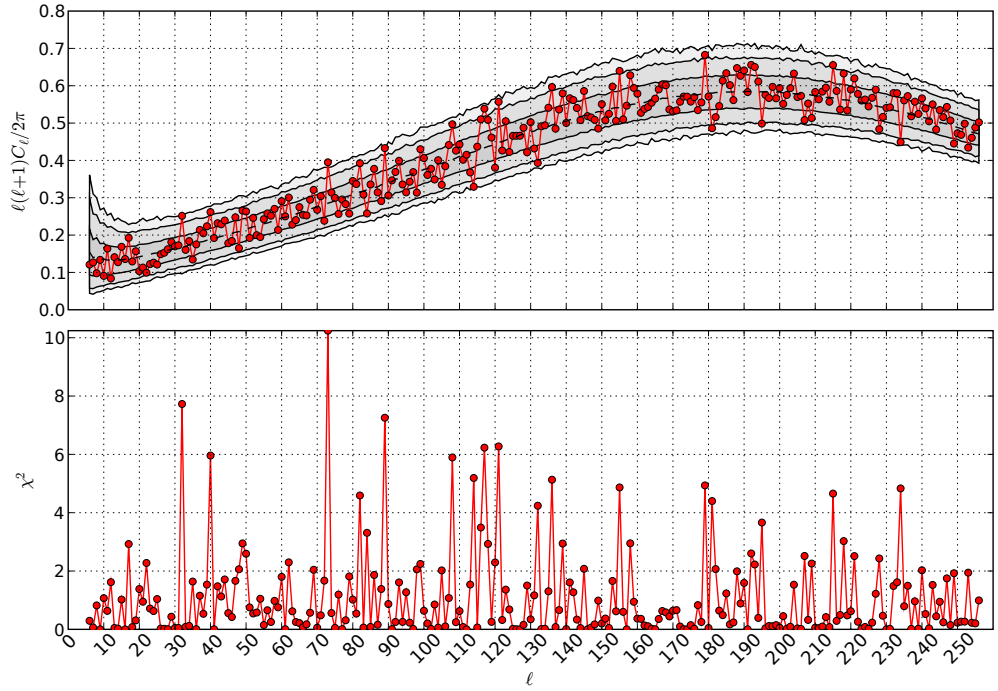


Figure 5.2: Consistency check: the pseudo power spectrum of the WMAP V5 data with the cosmic variance 68%,95% and 99% limits from simulations and the simulation mean (top panel), and the χ^2 value per degree of freedom (bottom panel).

confidence level of $\sim 99.99\%$. While this value is based on the extrapolation it might be important to complement the oddity of the mean of this map by noting that none out of the 3000 simulations, doesn't have as large mean value outside the KQ75 sky mask, and that within the simulated set of the maps the extremal temperatures are $\{-8.45, 7.44\} \times 10^{-07}$ with the standard deviation of the distribution: 2.15×10^{-07}, whereas the data value is 8.07×10^{-07}, and where the variance of the map is consistent with the simulations at 64% CL (as measured by the double-tail distribution). Even if we inverted the distribution by multiplying each mean value by -1, to check whether or not the shot noise is responsible for such high significance anomaly, the result would yield only 3 simulations in 3000 with means larger than the data value, and as such would still lead to an anomaly at $> 3\sigma$ CL.

5.4.2 Anomaly localization

We test whether the extended sky cuts would restore consistency between the data and simulations within the selected range of multipoles $\ell \in (32, 128]$. We gradually extend the KQ75 sky mask by masking out regions along the galactic disk of latitudes

$$|b| \leq \{10, 20, 30, 45\} \text{ [deg]}$$

and measure the response of the mean temperature outside of the extended mask. Under these extended masks the tail probability of finding the mean outside of the $\pm\langle T_{\rm data}\rangle$ is

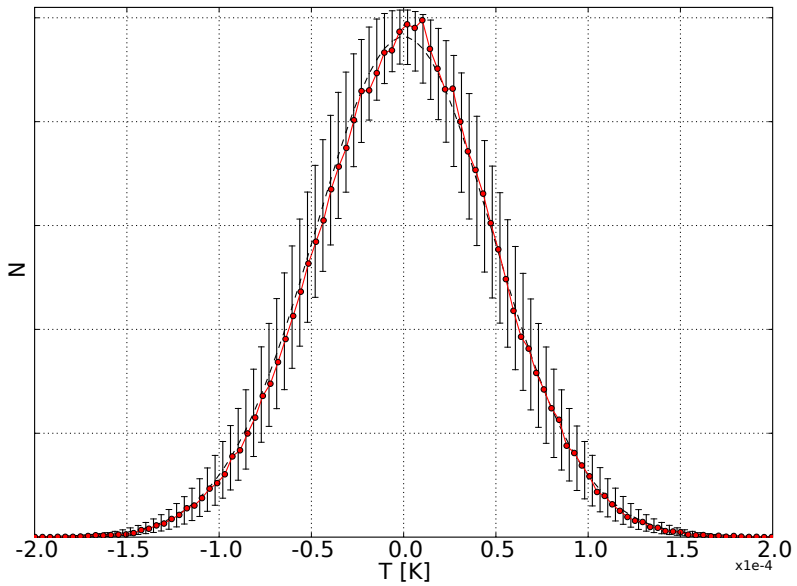


Figure 5.3: Temperature histogram for the filtered map in the $\ell \in (32, 128]$ multipole range, along with the simulations mean, and one standard deviation dispersion from simulations. It is easy to see a slight negative skewness of the distribution, and an overall excess of the positive temperatures close to the zero average temperature.

respectively

$$p \in \{0.016, 2.3, 54, 65\}\,[\%].$$

Clearly the enhanced masks resolve the problem of excessive number of positive temperature pixels, which helps in their localization.

In order to further localize the anomaly we performed a regional, one-point statistics as explained in Lew (2008b) for the filtered map in the considered multipole range, outside of the KQ75 sky mask. The analysis relies on regional measurements of the first four moments of temperature distribution in a number patches covering the sphere in many different ways so as to enable different sampling of the underlying fluctuations. For each region a frequentist statistic is performed, using a distribution of measurements obtained from simulations in the corresponding regions. We performed a "single region analysis" for the region sizes and shapes defined by the Healpix pixelization scheme of resolution $n_s = 4$ as explained in Lew (2008b). In figure 5.4 we plot the scrambled through all realizations of sphere pixelizations n_σ map for the distribution means also as defined as in Lew (2008b). The map shows only the strongest $n_\sigma > 3$, detected outliers found in the data.

While the strongest, outstanding regions revealed by this map could suggest that the temperature distribution anomaly is due to these regions, we compared the normalized temperature distributions of the tested map with the map with these two most outstanding regions masked out, and found basically no significant difference. This in fact should be expected, because the negatively skewed deviation in the distribution plotted in figure 5.3 comes from a large number of pixels which only slightly deviate from the zero level, while

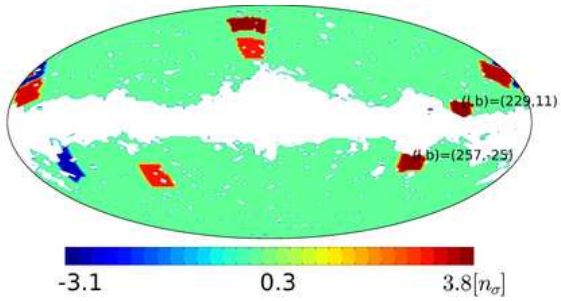


Figure 5.4: Thresholded at 3σ, n_σ map of deviations in distribution of means, as probed by the single region analysis, defined in Lew (2008b). The two strongest deviations are marked by their location.

the regional statistics plotted in figure 5.4 searches for the most outstanding regions. Consequently we do not find that the anomaly of the means, or as detected in the MFs analysis (figure 5.1) is caused directly by these regions. Of course, these as the hottest spots, and directly adjacent to the galactic plane could make some small part of this deviation but their impact is *definitely sub-dominant*.

In order to further validate these results we use additional extended mask that consists of the KQ75 sky mask and a mask covering the polar caps: $|b| > 45$, and we measure the distribution of means in data and simulations. As expected no improvement has been noticed, and the deviation was still at the level of 3σ.

5.4.3 Extended sky-mask cross-check

We redo the MF test for the V5 data map for the investigated multipole range: $\ell \in (32, 128]$, but with the KQ75 sky mask extended to also mask regions that yield $|b| < 30^\circ$ or $|b| < 45^\circ$. The result is plotted in figure 5.5 for the two cases. The joint p-values for the two cases, for each of the MFs are: $\{1.1, 18, 9\}$ [%] for the $|b| < 30^\circ$ case, and $\{26, 10, 23\}$ [%] for the $|b| < 45^\circ$ case, for $N_\nu = 15$ and $N_{\rm cov} = 1500$. For the case of $N_\nu = 13$ and $N_{\rm cov} = 1500$ the joint p-values yield: $\{0.99, 11, 5\}$ [%] and $\{20, 6, 19\}$ [%] respectively.

Clearly the anomaly is strongly sensitive to the size of the sky-cut and is weakening with more aggressive galactic plane region cuts, which could indicate some under/over-subtracted residual foregrounds.

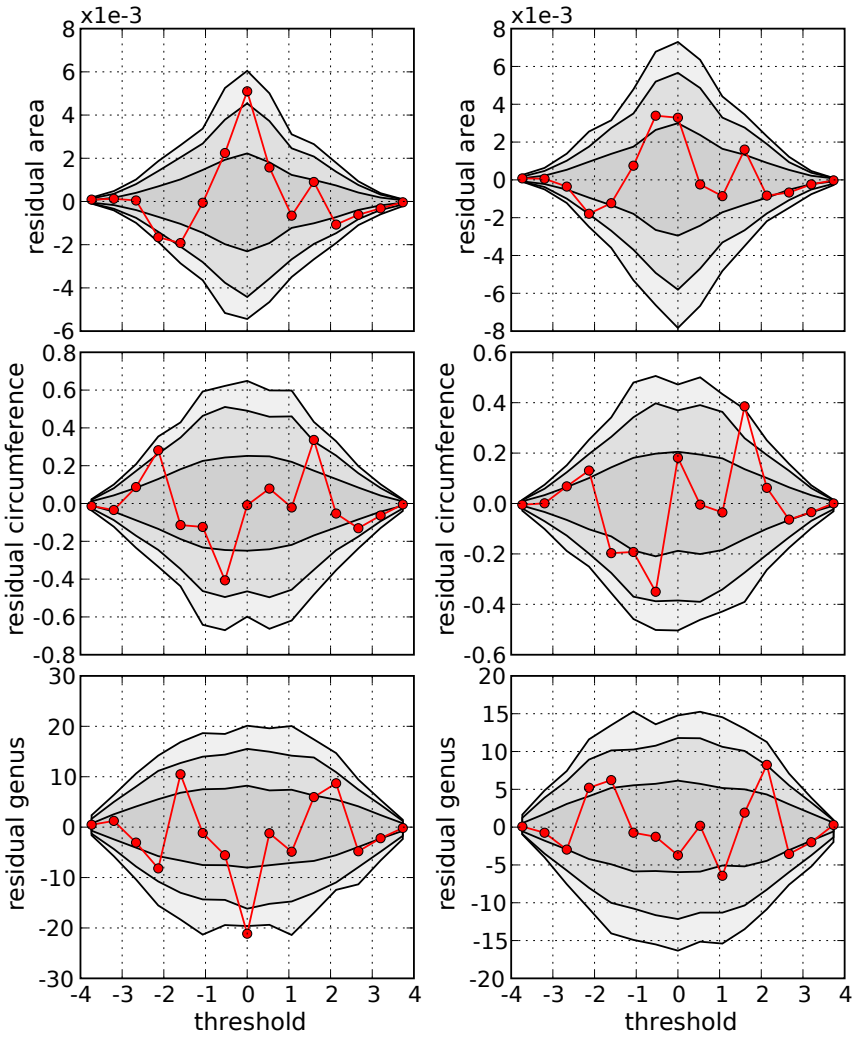


Figure 5.5: Residual plots for the results of the Minkowski functional cross-check analysis for the WMAP V5 data, filtered for the multipole range $\ell \in (32, 128]$ that previously yielded the strongest joint NG detection. The extended KQ75 sky-mask is used to additionally mask regions that yield $|b| < 30°$ (left panel) and $|b| < 45°$ (right panel).

Table 5.1: Joint statistic results of the MFs analysis of the WMAP V5 data, and tests of their stability to different numerical processing approaches. For each of the multipole bins, each triplet of numbers correspond to one of the MF type (area,circumference and genus respectively) and represent the p-values of the data in per cents. The p-values greater than 1% were approximated to integer percentile.

	$\ell_{\rm min}=5$				$\ell_{\rm min}=32$			$\ell_{\rm min}=64$		$\ell_{\rm min}=128$
MF type	$\ell_{\rm max}=32$	$\ell_{\rm max}=64$	$\ell_{\rm max}=128$	$\ell_{\rm max}=256$	$\ell_{\rm max}=64$	$\ell_{\rm max}=128$	$\ell_{\rm max}=256$	$\ell_{\rm max}=128$	$\ell_{\rm max}=256$	$\ell_{\rm max}=256$
	$N_{\rm cov}=1500$, $N_{\chi^2{\rm PDF}}=1500$, $N_\nu=15$									
area	10	15	15	14	9	0.50	10	3	1	2
circ.	7	0.03	6	7	14	40	4	56	14	79
genus	91	29	13	6	5	13	20	69	53	60
	$N_{\rm cov}=2500$, $N_{\chi^2{\rm PDF}}=500$, $N_\nu=15$									
area	11	12	15	14	10	0.38	10	0.75	0.80	2
circ.	7	0.08	6	6	16	38	3	56	14	74
genus	91	28	14	5	7	16	25	71	52	62
	$N_{\rm cov}=1500$, $N_{\chi^2{\rm PDF}}=1500$, $N_\nu=13$									
area	14	46	10	14	8	0.24	7	0.87	3	1
circ.	5	43	22	3	14	32	3	54	10	75
genus	85	35	61	7	12	11	15	54	37	62
	$N_{\rm cov}=2500$, $N_{\chi^2{\rm PDF}}=500$, $N_\nu=13$									
area	14	45	10	13	7	0.20	6	0.54	2	0.95
circ.	5	45	24	3	14	27	3	52	10	71
genus	86	36	62	6	13	13	18	56	36	62
	$N_{\rm cov}=1500$, $N_{\chi^2{\rm PDF}}=1500$, $N_\nu=11$									
area	10	32	17	15	12	0.13	4	0.47	2	0.55
circ.	2	66	14	5	12	31	1	37	8	63
genus	76	60	54	4	7	10	14	41	69	47
	$N_{\rm cov}=2500$, $N_{\chi^2{\rm PDF}}=500$, $N_\nu=11$									
area	12	31	18	14	12	0.01	3	0.36	0.76	0.43
circ.	3	68	15	5	11	28	2	37	8	62
genus	77	61	58	4	8	12	17	44	69	49

5.5 Discussion

It is important to consider whether or not the statistics performed in the previous section might be somewhat mislead or "biased". It is clear from figure 5.1 that in all cases the detection comes from the overall shift towards positive temperatures of the data with respect to the simulations, because only in this case, the difference between the data and the simulation average is positive everywhere for the first MF: i.e. the area functional. Such a shift then results in excessively large residuals in the Minkowski area functional, and suppression followed by excess in the Minkowski circumference functional. This is indeed observed in the plots. However, the data, as well as the simulations were divided by their standard deviations prior the analysis. Therefore, even if there was some e.g. point source, that would significantly stand-out in the data from the rest of the pixels, it would not significantly alter the overall variance, given the large number of pixels in the map, and consequently the thresholds at which the MFs would be calculated, would not be significantly shifted with respect to those used in the simulations. Even if this was the case though, then it would be difficult to explain the test involving the means. The histogram in the figure 5.3 shows that indeed large number of pixels in the filtered map, within the selected multipole range that yields the strongest detection, are shifted towards positive temperatures. This would indicate that the anomaly is not caused by the mismatch of the thresholds between the data and simulations. The filtered maps have no monopole value, and so the only effect that could generate such statistical up-shift of the temperatures must be related to some specific phase arrangements that build up regions of extra positive temperature excess. We recall that this effect is very small, but it affects a significant number of pixels in the map, and therefore is not related to some few outliers, as indicated by the performed regional analysis (figure 5.4). Additional argument in favour of this interpretation is the negative, mildly anomalous, skewness of the temperature distribution.

It could also be argued that it's not the data that is shifted towards positive values, but rather simulations are oddly shifted towards negative ones, as one could infer from the reported, in the previous section, the extremal temperature values obtained in the set of the simulated maps for the multipole range $\ell \in (32, 128]$. As it was stated, the negative extremal value was larger (in absolute value) than the positive extremal value. In fact the average mean temperature, although negative: -4.26×10^{-09}, is about 50 times smaller than the standard deviation of the mean temperatures. Therefore even if one were to correct for that small negative offset, the mean value of the data would still remain odd at the 3σ confidence level.

If there was a discrepancy in the total variance between the data and simulations, as measured outside of the applied mask, then it could be possible that the thresholds, at which the MFs were calculated in the data, would not correspond to those used in the simulation, which could naturally explain the anomaly. In order to check that we calculated the variances for the maps filtered in the range $\ell \in (32, 128]$. We found that the corresponding tail probabilities for the data (i.e. probabilities-of-exceeding the data value) are: 36% for

the KQ75 sky mask and 30% for the KQ75 combined with the mask covering regions with $|b| < 45°$.

We derived the Minkowski functionals directly in the real space. We based on the closest pixel match, when deriving the covariant derivatives, but we note that the approach could be easily improved by implementing appropriate interpolations and efficient, dedicated derivative algorithms, or alternatively by obtaining derivatives in an more conventional way: via the recurrence formulas for the Legendre polynomials in the spherical harmonics space. However the particular implementation of these issues cannot be responsible for the anomalies, because the strongest detections appear to be in the first Minkowski functional that simply measures the total normalized excursion set above given temperature threshold, and furthermore the maps we used are generated by a well-behaved functions, which should be easy differentiable, given sufficient number of points.

Finally we would like to note an interesting coincidence of negatively skewed temperature distribution, revealed in figure 5.3 and the recent report by Yadav & Wandelt (2008) about the high significance primordial NG detection with strong support for the positive value of the non-Gaussianity quadratic term coupling coefficient $f_{\rm NL}$. Positive value of $f_{\rm NL}$ (in the "WMAP" terminology egKomatsu et al. (2003)) corresponds to the negative skewness of the temperature distribution. However, Yadav & Wandelt (2008) tested the NG detection under different sky cuts and found that these are more or less stable under the extended galactic cuts. Moreover it is possible that the bi-spectrum based NG estimator of (Yadav et al., 2008, 2007) is immune to the kind of anomalies revealed in this work, therefore the connection is not clear.

It will be interesting to further test this signal by analyses involving different spectral channels of the WMAP, and the various ILC maps, which are generated using a different foregrounds cleaning methods than those used in the foregrounds reduced maps used in this paper.

5.6 Conclusions

We performed a Gaussianity analysis of the WMAP five-year, foreground reduced, cut-sky, CMBR data, using Minkowski functionals.

Since the power of the map per multipole, is unequally distributed across the multipole scale due to certain shape of the underlying power spectrum, it might be useful, when testing different angular scales, to eliminate the impact of the largest, and most uncertain (due to the cosmic variance effects) scales, in order to focus on the smaller scales exclusively, and perform an analysis, that will not be dominated, or obscured by the effects large scale multipoles. In this paper we intend to explore this idea, and we perform a statistics in different ranges of multipoles, using multipole band-pass filters in the spherical harmonics space prior to the tests. This way we complement the previous similar works which utilized averaging of different scales by either: degrading the map resolution, or by applying "low-

ℓ"-pass Gaussian filters.

We find the data inconsistent with a realistic Gaussian random field simulations at confidence level at least 99.5%, due anomalous, negatively skewed temperature distribution, covering large areas of the sky, in the multipole range $\ell \in (32, 128]$. We test the robustness of these results using few additional consistency checks. We find the anomaly appears to be associated with a small, but realized by a large number of pixels, excess of positive temperature pixels over the negative ones, and we check that this effect is mitigated in the larger galactic latitudes, and it is vanishing when more aggressive galactic plane masks are applied, which we interpret as hints of the residual, rather small, but extended over a large area, foregrounds. Finally, we note that the anomaly would not be detected, if the analyses involved only smoothed or downgraded in resolution maps.

Chapter 6

Tests of the statistical isotropy via point symmetries.

6.1 Introduction

As detailed in section 2, the most outstanding, and well-known anomalies of the CMB data with respect to the ΛCDM model, and the assumption of the Gaussian, random field (GRF) initial, primordial perturbations include: the apparent low quadrupole component of the CMB angular power spectrum, the phase alignments between the quadrupole and the octupole, and between the octupole and the multipole $\ell = 5$, the suppression of power in the octupole along an axis dubbed an Axis of Evil (AOE), a selective preference of particular spherical harmonic modes over other modes, glitch-like features of the reconstructed CMB power spectrum, or other power spectrum features, inducing a specific symmetries in the CMB sky, and the hemispherical anomalies in the power distribution across the sky which are discussed in more details in chapter 4, and localized NG features that we have sought for in chapter 3.

These features have been extensively studied via different statistical estimators, and various theoretical models have been devised to help explain these peculiarities (eg. Inoue & Silk (2007); Erickcek et al. (2008); Bernui & Hipólito-Ricaldi (2008); Akofor et al. (2007); Destri et al. (2008); Brown & Crittenden (2005); Gordon et al. (2005)). Most of these features, if really exist in the data, would violate the statistical isotropy or Gaussianity (or both), which in turn would call for an explanations that elude the predictions of the ΛCDM model.

In the following we analyze the shape of the reconstructed power spectrum in the lowest multipole regime ($\ell < 20$). One of the most apparent features is the jagged shape of the C_ℓ persistently preferring an odd ℓ modes over the even modes, which seem to by systematically suppressed (Fig. 6.1). In fact, most of the first 20, reconstructed even multipoles, have relatively less power than the neighbouring odd multipoles. This fact is independent from the analysed cleaned CMB map renditions and the power spectrum reconstruction method,

as will be discussed in Sec. 6.3. Such feature, if proved to be "statistically significant" could have a far reaching theoretical consequences as such features necessarily introduce a point-symmetries in the CMB fluctuations due to symmetry properties of the spherical harmonics.

In this chapter we reanalyze this problem statistically and we will investigate other aforementioned problems in the next chapter. Our analysis, although similar to previous works Land & Magueijo (2005c), implements a different statistics in order to independently cross-check significance, stability and the robustness of the previous results.

6.2 Data and simulations

We rely on the three year WMAP ILC map data Hinshaw et al. (2007) and its simulations. We assume it to be reliable for the full sky multipole amplitude and phase analyses in multipole range of $\ell \in \{2, .., 20\}$. From the minimal best fit ΛCDM model we have generated the full sky realistic GRF WMAP simulations including effects of beams and anisotropic noise in resolutions $n_s = 512$ in all WMAP channels. These maps were pre-smoothed with one degree FWHM Gaussian beam and used to generate the ILC simulated maps with the same weights, and region definitions applied as those originally used for the WMAP data by Bennett et al. (2003a). Whereas the white, anisotropic noise contribution to the lowest multipole moments is negligible we still have added it prior the smoothing, to simulate the ILC data as closely as possible. We have downgraded the full resolution maps the resolution parameter $n_s = 128$, although we realize that this resolution is still too generous for the range of the scales in the question. We have generated 2222 GRF ILC simulations which we use for the significance analysis.

6.3 Point symmetries

If some astrophysical, cosmological or instrumental process would induce a point-symmetries or asymmetries on the CMB maps[1]

$$T(\hat{\mathbf{n}}) = \pm T(-\hat{\mathbf{n}}) \tag{6.1}$$

(where $+$ corresponds to symmetry and $-$ correspond to antisymmetry), then it is quite straightforward to show that the angular power spectrum C_ℓ would have even or odd, for symmetry and antisymmetry respectively, multipoles suppressed or completely removed, depending on the degree of the symmetry.

Conversely, our motivation for pursuing the symmetry test of the data, is directly related to the observational fact that the reconstructed full-sky angular power spectrum seem to exhibit a suppression of even multipoles with respect to odd multipoles in the considered multipole range (Fig. 6.1). The feature particularly does seem to depend on the details of

[1]Given the specifications of the WMAP spacecraft this is in fact difficult to realize, however the feature under consideration is apparent enough to question its significance.

the power spectrum reconstruction method, nor does it depend on the CMB data considered (the first or the three year WMAP CMB data).

To see this we reconstruct the CMB power spectrum for the range $\ell \in \{2, .., 20\}$ using the MASTER method (Hivon et al., 2002). Below we briefly recall the basic reconstruction steps.

6.3.1 MASTER power spectrum reconstruction

The method of reconstructing the statistically unbiased, in general, cross-power spectrum estimator $\widehat{C}_\ell^{\rm AB}$ of the underlying, true CMB power spectrum C_ℓ, from the cut sky, noisy CMB observations from two different frequency channel (A,B) maps $\widetilde{T}^A(\hat{\mathbf{n}})$ and $\widetilde{T}^B(\hat{\mathbf{n}})$, is based on the fact that the ensemble averages of the so-called pseudo cross-power spectrum $\widetilde{C}_\ell^{\rm AB}$ and the true power spectra C_ℓ can be related as:

$$\langle \widetilde{C}_\ell^{\rm AB} \rangle = \mathbf{M}_{\ell\ell'} b_{\ell'}^A b_{\ell'}^B p_{\ell'}^2 \langle C_{\ell'} \rangle + \langle \widetilde{N}_\ell^{\rm AB} \rangle \tag{6.2}$$

where the $\mathbf{M}_{\ell\ell'}$ denotes the multipole-to-multipole coupling matrix, arising from partial sky observations and the resulting breaking of orthogonality of the spherical harmonics; the $b_{\ell'}^{A/B}$ and $p_{\ell'}$ denote the instrumental beam transfer function for a given frequency channel, and the pixel transfer function respectively; the $\langle \widetilde{N}_\ell^{\rm AB} \rangle$ denotes the ensemble average noise power spectrum estimated from Monte-Carlo simulations. The pseudo-power spectrum $\widetilde{C}_\ell^{\rm AB}$ is a power spectrum derived directly on the cut sky, using noisy observations as:

$$\widetilde{C}_\ell^{\rm AB} = \frac{1}{2\ell+1} \sum_m \widetilde{a}_{\ell m}^A (\widetilde{a}_{\ell m}^B)^\star \tag{6.3}$$

where

$$\widetilde{a}_{\ell m}^x = \int_{\Omega_{\hat{\mathbf{n}}}} {\rm d}\hat{\mathbf{n}} \widetilde{T}^x(\hat{\mathbf{n}}) w(\hat{\mathbf{n}}) Y_{\ell m}^\star(\hat{\mathbf{n}}) \tag{6.4}$$

are the pseudo spherical harmonics analysis coefficients of the $T^x(\hat{\mathbf{n}})$ observation ($x \in \{A, B\}$) and the $w(\hat{\mathbf{n}})$ is the sky mask window function with the corresponding spherical harmonics analysis coefficients

$$\widetilde{w}_{\ell m} = \int_{\Omega_{\hat{\mathbf{n}}}} {\rm d}\hat{\mathbf{n}} \widetilde{w}(\hat{\mathbf{n}}) Y_{\ell m}^\star(\hat{\mathbf{n}}) \tag{6.5}$$

and the angular power spectrum

$$\widetilde{w}_\ell = \frac{1}{2\ell+1} \sum_m |\widetilde{w}_{\ell m}|^2. \tag{6.6}$$

Then the unbiased estimator of the angular CMB power spectrum $\widehat{C}_\ell^{\rm AB}$ is given by:

$$\widehat{C}_\ell^{\rm AB} = \left(\mathbf{K}^{\rm AB}\right)_{\ell\ell'}^{-1} \left(\widetilde{C}_\ell^{\rm AB} - \langle \widetilde{N}_\ell^{\rm AB} \rangle \right) \tag{6.7}$$

where the $\mathbf{K}_{\ell\ell'}^{\rm AB} = \mathbf{M}_{\ell\ell'} b_{\ell'}^A b_{\ell'}^B p_{\ell'}^2$, and where the coupling kernel $\mathbf{M}_{\ell\ell'}$ is solely defined by the sky cut $w(\hat{\mathbf{n}})$ by:

$$\mathbf{M}_{\ell\ell'} = \frac{2\ell'+1}{4\pi} \sum_{\ell''} (2\ell''+1) w_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 \tag{6.8}$$

where the last term denotes the 3J Wigner coefficients. The corresponding covariance matrix of the estimated angular power spectrum is then given by:

$$\widehat{\mathbf{C}}^{\mathrm{AB}}_{\ell\ell'} = \left\langle \left(\widehat{C}^{\mathrm{AB}}_{\ell} - \langle \widehat{C}^{\mathrm{AB}}_{\ell} \rangle_{\mathrm{MC}} \right) \left(\widehat{C}^{\mathrm{AB}}_{\ell'} - \langle \widehat{C}^{\mathrm{AB}}_{\ell'} \rangle_{\mathrm{MC}} \right) \right\rangle_{\mathrm{MC}} \tag{6.9}$$

where $\langle \rangle_{\mathrm{MC}}$ denotes an average over the MC simulations.

In Fig. 6.1 we plot the reconstructed power spectrum of the CMB fluctuations from different cross-channel data sets as well as the spectrum reconstructed directly from the full sky foregrounds cleaned ILC map, with over-plotted best fit fiducial ΛCDM model power spectrum. Indeed, except for the multipole $\ell = 14$ every other even multipole is suppressed with respect to the neighbouring odd multipole. For clarity we have only plotted the cross-power spectra between channel Q1 and all other channels, however the remaining cross-power spectra are widely consistent with those plotted in Fig. 6.1

6.3.2 Statistic

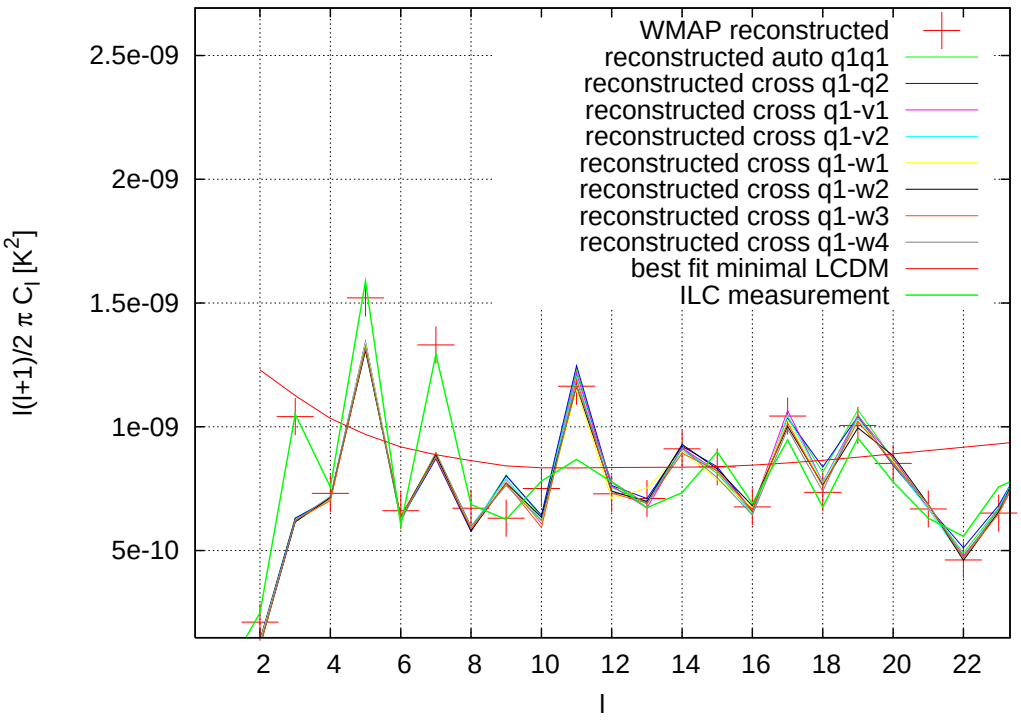


Fig 6.1: Reconstructed low-multipoles power spectrum of the CMB fluctuations as a function of used data set. Cross-power spectra and auto-power spectra were calculated with MASTER method. The green line represents the power spectrum obtained from the direct full sky SHT of the ILC map, whereas the big red crosses indicate the WMAP-team reconstructed power spectrum using a maximum-likelihood method in pixel space.

We propose to test the statistical significance of this asymmetrical signal via the follow-

ing correlation statistics:

$$\mathcal{C}_{\ell_{\rm min}}(\ell_{\rm max}) = -\frac{\langle T(\hat{\mathbf{n}})T(-\hat{\mathbf{n}})\rangle}{\langle T^2(\hat{\mathbf{n}})\rangle + \langle T^2(-\hat{\mathbf{n}})\rangle} \tag{6.10}$$

where

$$T(\hat{\mathbf{n}}) = \sum_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} T_\ell(\hat{\mathbf{n}}). \tag{6.11}$$

The ideally correlated sky - as in case of single multipole CMB maps - would yield $\mathcal{C} = \pm 1$ with "$-$" for the symmetrical case (i.e. for even multipoles) and with "$+$" for the anti-symmetrical case (i.e. for odd multipoles). We note that other symmetries like: $T(-\theta,\phi) = \pm T(\theta,\phi)$ or $T(\theta,-\phi) = \pm T(\theta,\phi)$ will not lead to a systematical suppression of either even or odd multipoles. Rather, such symmetries suppress power in both: even and odd multipoles proportionally. In case of the earlier (plane) symmetries the power is zeroed for those coefficients that yield odd $\ell+m$ value (see Appendix A-4 for an exact derivation), while the latter symmetries imply reducing the phase information in all $a_{\ell m}$ coefficients down to either $\phi = 0$ or $\phi = \pi$, and so the $a_{\ell m}$ coeffiecients become real, and the reality condition simplifis to: $a_{\ell -m} = a_{\ell m}$.

If the CMB fluctuations are a realization of a GRF, then no point-like correlations of this type are expected, beyond those that are enforced by an artificial cutoff of range of multipoles of interest, as well as those resulting from the ISW tail in the low-ℓ-end of the power spectrum (6.1). We test the correlations in the CMB multipole, range filtered maps $\ell \in [\ell_{\rm min}, \ell_{\rm max}]$ for varying $\ell_{\rm max}$ values (Eq. 6.11). We increase the $\ell_{\rm max}$ value by two multipoles at one step, in order to reduce the effects of otherwise induced point symmetries only from the fact of considering unequal number of even and odd multipoles.

Since the $\ell = 2$ of the best fit ΛCDM model has more power than $\ell = 3$, and also due to the fact that the nature of the low quadrupole value might be unrelated to the features in question, we therefore consider two cases. In the first case we synthesize maps starting from the quadrupole: $\ell_{\rm min} = 2$ and vary the $\ell_{\rm max}$ as $\ell_{\rm max} \in \{3, 5, 7, ...19, 21\}$, and in second case we synthesize maps starting with the octupole: $\ell_{\rm min} = 3$ and vary the $\ell_{\rm max}$ as $\ell_{\rm max} \in \{4, 6, 8, ...20\}$.

Also, due to the Integrated Sachs-Wolfe tail of the ΛCDM CMB power spectrum yielding the $C_2 > C_3$ the distinction of the two cases help to assess the asymmetry significance problem in two opposite limits: when the synthesized maps will be dominated by the positive and negative correlations.

It is important to account for correlation effects between different multipole ranges, and as such, our statistics is useful, because it operates on the synthesized real-space maps that combine selected multipoles together, and thereby the joint statistics is automatically assessed.

6.4 Results

The result of the tests involving 2222 MC simulations, used for evaluation of the confidence intervals, is plotted in Fig. 6.2. We estimate the magnitude of the uncertainties of derivation of $C_{\ell_{\min}}(\ell_{\max})$ values, jointly due to noise and residual foregrounds of the ILC map, to be $\approx 0.1\%$ for $\ell_{\max} = 21$ and less for smaller values of $\ell_{\max}$. As expected, it is apparent

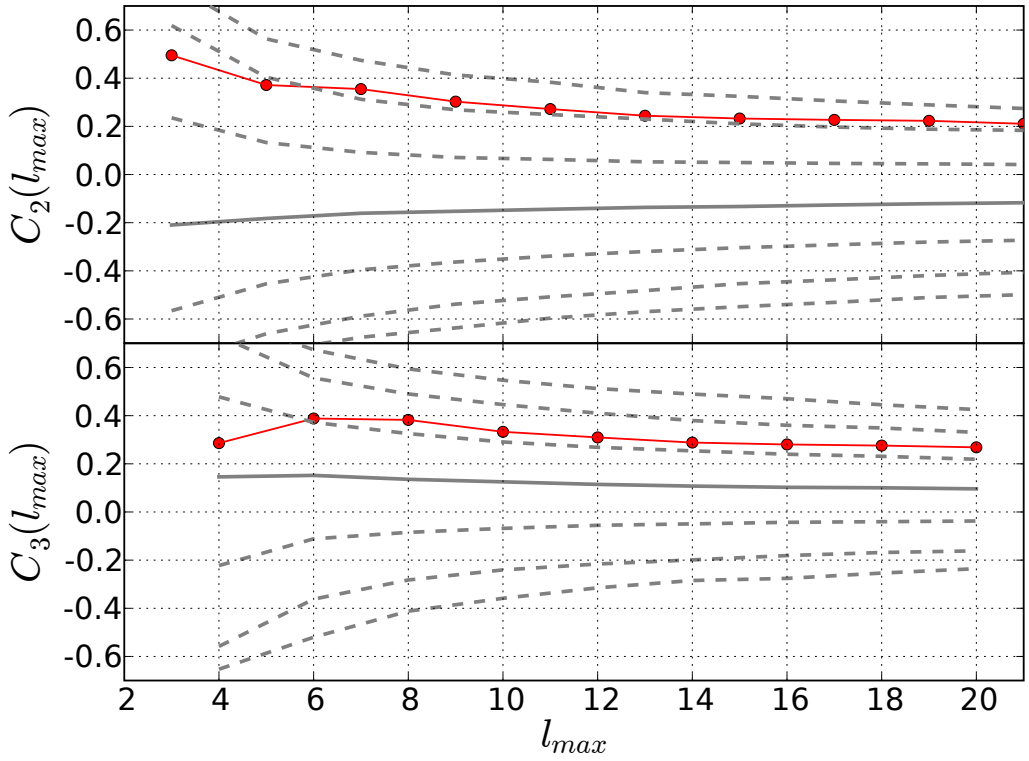


Fig 6.2: Point symmetry correlations in the three year WMAP ILC map data as a function of the maximal multipole number $\ell_{\max}$ of the multipole filtered map. Apparently the data are consistent with simulations at $\sim 1\sigma$ CL in case of the maps neglecting the quadrupole contribution, and are consistent at $\sim 2\sigma$ CL if the quadrupole is included. The maximal total joint anomaly is reached for $\ell_{\max} = 19$. The error bars for the data points are smaller than 0.1% and hence fit within the size of the dot.

that the mean correlation from the simulations assumes a negative values in case when the quadrupole is included, providing statistically stronger negative correlations due to symmetry properties, than could be cancelled out by the octupole component due local shape of the underlying power spectrum. In case when quadrupole is not included, the anti-symmetrical contribution from the octupole, which statistically is enhanced via ISW tail in the power spectrum, is not entirely cancelled out by $\ell = 4, ...$, thus leaving the simulation average positive.

The observed suppression of the even multipoles is maximized at $\ell = 19$ in case when the quadrupole is included with $P_{reject} \approx 98.4\%$, and it is clear that its significance is

almost constant as a function of $\ell_{\max}$ value for $\ell \geq 7$. This result is consistent with the results of Land & Magueijo (2005c), where a different statistical approach have been used.

Independently from the low quadrupole value (case when $\ell_{\min} = 3$), the suppression of the even multipoles in the reconstructed power spectrum of the CMB remains anomalous at confidence level $\approx 92\%$, and is maximized at the maximal considered multipole value of $\ell_{\max} = 20$.

Table 6.1: Results of the point symmetry correlations in the ILC WMAP3 year data.

$\ell_{\min} = 2$		$\ell_{\min} = 3$	
$\ell_{\max}$	P_{reject}[%]	$\ell_{\max}$	P_{reject} [%]
3	94.6	4	64.4
5	96.6	6	85.6
7	98.3	8	90.7
9	97.9	10	89.1
11	98.0	12	89.3
13	98.0	14	89.1
15	98.1	16	90.2
17	98.3	18	91.2
19	98.4	20	92.2
21	98.2		

6.5 Discussion

6.5.1 Point symmetries in the sky

The results from Sec. 6.4 show that the significance of the suppression of the even multipoles in the CMB power spectrum is mildly significant: $\approx 92\%$ CL in case when quadrupole is excluded from the analysis and is increased to $\approx 98.4\%$ CL if the quadrupole is included. While it is not easy to develop a viable theoretical model that could produce the considered point asymmetries in the sky, it is fairly easy to show that axis-symmetrical contaminations at fixed galactic longitude, or mirror symmetries about the galactic plane, which could be easily associated with the residual foreground emissions, would not lead to suppression of either even or odd multipoles of the CMB power spectrum. This statement seems to go against to what was discussed in Land & Magueijo (2005c).

Chapter 7

Search for preferred reference frame in the CMB data.

In this section we present some of the results from the ongoing project dedicated to generalized statistical tests of the significance of the alignments in the observational CMB data sets. The final results enforced by the significance analysis based on Monte-Carlo simulations will be submitted to the Journal of Cosmology and Astroparticle Physics.

7.1 Introduction

In this chapter we focus on a set of tests, related to the well known, low multipole anomalies of the CMB data, and compare our results with those previously reported, as well as with those obtained from more generalized tests.

While looking thorough results of different statistical tests, (e.g. Komatsu et al. (2003); Komatsu et al. (2008); McEwen et al. (2006a); Vielva et al. (2004); Wiaux et al. (2006); Mukherjee & Wang (2004); Savage et al. (2004); Naselsky et al. (2007); Cabella et al. (2004, 2006, 2005); Chen & Szapudi (2006a); Curto et al. (2008a, 2007); de Troia et al. (2007); Armendariz-Picon & Pekowsky (2008); Souradeep et al. (2006); Hajian & Souradeep (2006,?); Samal et al. (2008); Hansen et al. (2004a,b); Bernui et al. (2007b,a); Naselsky et al. (2005); Chiang et al. (2003); Chen & Szapudi (2006b); Gaztañaga et al. (2003); Copi et al. (2004, 2006a,b); Bielewicz et al. (2005); Abramo et al. (2006); de Oliveira-Costa & Tegmark (2006); Land & Magueijo (2005b, 2007); Jaffe et al. (2005); Efstathiou (2003b); Eriksen et al. (2004a, 2007, 2008); Land & Magueijo (2005b); Park et al. (2006); Shandarin (2002); Wu et al. (2001); Park (2004); Lew & Roukema (2008); Lew (2008b); Donoghue & Donoghue (2005); Yadav & Wandelt (2008) and references therein), many different numbers denoting confidence levels of rejecting tested hypotheses, on grounds of inconsistencies due to Gaussianity or statistical isotropy, or related - non-randomness, are generally found. This is because there are many different aspects of the CMB map that one can possibly test. Given no prior knowledge about whether particular features are correlated or not, re-

sults of statistical tests may over or underestimate the true "oddity" of the tested features. Therefore, in case of correlated anomalies (each of which may be a function of some unknown background physical process) a simple multiplicative combining of probabilities of assumed-as-independent incidents, could lead to a significantly overstated result. Also, different statistical tests and estimators may exhibit different sensitivities to particular features in the data. These tests could in principle also be dependent on one another, to some extent.

It is therefore important to test not only the unlikeliness of a selected features of data, but also test the processes that correlate presumably independent phenomena, or devise tests that jointly account for as many different aspects of the data as possible (Rakic & Schwarz, 2007). Such tests however could be either computationally intensive of inconclusive - i.e. giving consistent results within an assumed model work frame, but contradicting the results of more selectively sensitive tests. In particular one could easily develop some specific test that would seek for some particular properties arrangement found in the data, and find it very difficult to reproduce then in simulated data-sets, and conclude a strong inconsistency of the data with the simulations, only because of ill defined test. Otherwise in case of a general test and lack of any anomalies detection, one could always argue that the test is not sensitive enough to detect a signal found by other, more selective tests.

In our analysis we rely on the Gaussian Random Field (GRF) realizations, and assume these as a working reference null hypothesis, although, given that the process behind the formation of the plausible anomalies of the low-ℓ multipoles are unknown, we will not assume that individual multipoles of the data are actually independent. In our calculations therefore, we rely on the full covariance matrix estimates, as derived from the GRF simulations, to constrain limits on the mode-to-mode correlations, which appear to some extent simply due to the cosmic covariance effects. While such an approach can at most quantify the degree of consistency of the data with the GRF simulations, and will not provide any further information about the nature of the potential anomalies, or their source, it is a good start before a specific studies of a more general models, either non-Gaussian, or non-Isotropic are pursued.

In the following, we revisit the issue of the previously mentioned "Axis of Evil" (AOE) which has originally been reported by de Oliveira-Costa et al. (2004); Tegmark et al. (2003); de Oliveira-Costa & Tegmark (2006). In these works a statistic based on an angular-momentum estimator $M_\ell(\hat{\mathbf{n}})$ was used to search alignments between multipoles, by comparing the directions $\hat{\mathbf{n}}^\star$ in which the estimator

$$M_\ell(\hat{\mathbf{n}}^\star) = \max_{\hat{\mathbf{n}}} \frac{\sum_{m=0}^{\ell} m^2 |a_{\ell m}(\hat{\mathbf{n}})|^2}{\sum_{m=0}^{\ell} |a_{\ell m}(\hat{\mathbf{n}})|^2} \tag{7.1}$$

is maximized. The $a_{\ell m}$ denote the coefficients of the spherical harmonics analysis. Throughout the analysis we assume a convention of using $\hat{\mathbf{n}}^\star$ for the directions that maximize (alternatively minimize) a particular estimator in a search over all directions in the sky. By construction, this estimator gives more preference to higher m modes, and consequently prefers the sectoral harmonic modes over the zonal modes. It was found that the alignment between the quadrupole and octupole is significant; corresponding to a chance roughly 1

in 60, as compared with an isotropic field realizations. The search is obviously motivated by the visually sectoral (but somewhat rotated) appearance of the quadrupole and octupole components. Hence the general idea behind this test was to devise a method of assigning a vector to each multipole and using Monte-Carlo (MC) simulations to quantify the degree of inter-multipole correlations. These in the in the GRF should be statistically consistent with zero. This idea was then extended by Land & Magueijo (2005b) and generalized to assign vectors to all $a_{\ell m}$ coefficients and analyze the inter-ℓ alignments using only those vectors, that correspond to $a_{\ell m}$ modes, which accumulate most of the power of the ℓth multipole in a most preferred direction $\hat{\mathbf{n}}^{\star}$:

$$R_{\ell}(\hat{\mathbf{n}}^{\star}) = \max_{m}\big(r_{lm}(\hat{\mathbf{n}}^{\star})\big) \tag{7.2}$$

$$r_{\ell m}(\hat{\mathbf{n}}^{\star}) = \max_{\hat{\mathbf{n}}}\Big(c\frac{|a_{\ell m}(\hat{\mathbf{n}})|^2}{(2\ell+1)C_{\ell}}\Big), \quad \text{for } m \geq 0 \tag{7.3}$$

where $c = 1$ for $m = 0$, and $c = 2$ for $m \neq 0$.

There are however several subtleties to be considered in both of the tests. Firstly, the zonal harmonic mode (a_{20}) of the three year WMAP ILC data quadrupole appears to be yet even stronger than the sectoral mode (a_{22}) as it assumes $(93.55 \pm 0.03)\%$ of quadrupole power in a reference frame with polar axis oriented at $(l, b) = (333^{\circ}, 3^{\circ})$ in galactic coordinates, whereas the a_{22} mode absorbs $(93.08 \pm 0.03)\%$ of the quadrupole power in a reference frame of $(l, b) = (236^{\circ}, 68^{\circ})$ [1]. These values depend slightly on the data-set used for calculation, however the general picture of strong competition between different $a_{\ell m}$ modes remains the same.

Likewise, the power absorbed by the octupole coefficients a_{31} and a_{33} is $(91.29 \pm 0.03)\%$ and $(91.73 \pm 0.03)\%$ in a reference frames $(l, b) = (116^{\circ}, 14^{\circ})$ and $(l, b) = (237^{\circ}, 63^{\circ})$ respectively. Note that the angular separation between the preferred axes of a_{20} and a_{31} is only $\sim 40^{\circ}$. We also note that at the same time the preferred direction for a_{21} is $(l, b) = (111^{\circ}, 13^{\circ})$, towards which the mode absorbs $\sim 93.11\%$ of the quadrupole power, and which is very close (within $\sim 5^{\circ}$) to the most preferred orientation by the mode a_{31}.

Given that the difference between the amount of absorbed power between these modes is only $\lesssim 0.5\%$ in all cases, while the consequences to the question of alignments are obvious, it is important to also take into the account the second-best solutions, and in general all other significant solutions as well.

Moreover, there exists a plausible alignment (within an angular separation of 14°) between the most zonal harmonics in reference frames that maximize the power accumulated in modes a_{20} and a_{30}. Although the a_{30} does not absorb most of the power in a preferred frame, as opposed to a_{33}, the a_{20} does. To explore quantitatively these possibilities, we

[1] The estimate of the given uncertainties is based only on the estimate of the numerical precision with which the r_{lm} values were derived, but it doesn't take into account all other possible sources of uncertainties which in general will tend to increase these estimates. We will return to the issue of uncertainties latter in this chapter.

perform a real-space based statistics, corresponding to the maximal-momentum search. We reproduce the results of de Oliveira-Costa et al. (2004) and extend it for search for zonal-like modes that instead of maximizing the momentum estimator, minimize it.

Secondly, the momentum information - i.e. it's magnitude itself is a valuable one as well. In our test we utilize the momentum magnitude as a weight for measuring the inter-ℓ correlations. This way, a strongly correlated multipoles will be better extracted, compared to multipoles which accidentally happen to have a similar maximum momentum orientation, but aren't really either strongly sectoral nor strongly zonal.

Thirdly, given that the observations have a finite signal-to-noise ratio, there is a certain degree of freedom in interpretation of the preferred orientations $\hat{\mathbf{n}}$, that result in a maximal accumulation of multipole power in a single $a_{\ell m}(\hat{\mathbf{n}})$ mode. This is most easily seen in the low-ℓ range; especially in quadrupole and octupole with a dipole as an extreme case. For a given orientation of a dipole, there exist exactly one additional preferred, perpendicular orientation. Each of these orientations will result in an 100% of the dipole power being assigned to a either a_{10} or a_{11} in a selected coordinate system.
The situation with the quadrupole component is also ambiguous. Assuming that the a_{20} has no power at all, there exist exactly two reference frames and two corresponding preferred, and perpendicular to each other axes, in which the modes a_{21} and a_{22} will assume exactly 100% of the quadrupole power. This is easy to see via a proper rotation of e.g. real part [2] of the Y_{21} [3], which will become Y_{22} in a new reference frame. This degeneracy generally does not extend beyond $\ell = 2$, however we are not going to give a strict mathematical proof of that.

In Land & Magueijo (2005b) it was found that the inter-ℓ alignments not only involve the quadrupole and octupole, but actually extend over a multipole range $\ell = 2 \ldots 5$, and it was pointed out that the a_{53} absorbs $> 90\%$ of the $\ell = 5$ power in a reference frame consistent with the orientation of AOE ($(l, b) = (250^\circ, 60^\circ)$) within only several degrees. While it is natural that if a two multipole modes of different multipole number are aligned along an axis, they induce, to some extent, the alignments between other modes of their own multipoles, due to geometrical properties of the spherical harmonics, the answer to the question of which particular "preferred" orientation is intrinsically preferred may be ambiguous, given the finite uncertainty in power distribution between the competing modes. Also, if the joint uncertainties due to astrophysical residual foregrounds, instrumental noise, map cleaning and processing methods, as well as numerical errors yielded effects larger than $\approx 1\%$, leading thereby to changes in the most preferred $a_{\ell m}(\hat{\mathbf{n}}^\star)$ modes per multipole, the inter-multipole alignments could be destroyed or could point out a different candidate for the Axis of Evil. In current work we estimate the size of these effects in section devoted to mentioned uncertainties estimation.

[2] This is without loss of generality since it's always possible to find a reference frame and such rotation so that the phase information is cancelled

[3] We will use the Y_{lm} notation instead of Y_l^m of the spherical harmonics, although we use the definition of the spherical harmonics consistent with Abramovitz & Stegun (1972)

Finally, to marginalize over the mentioned ambiguities we propose a generalization of the analysis of Land & Magueijo (2005b) to include all the $a_{\ell m}(\hat{\mathbf{n}})$ information and all inter-mode correlations in the multivariate calculus. As suggested earlier, additionally to the preferred orientation information of each $a_{\ell m}$ coefficient, we also utilize the percentage of accumulated power information as a statistical weight.

7.2 Search for preferred directions in the CMB sky

7.2.1 Multipole pair angular-momentum axes alignment estimators

In this section introduce statistics optimized for testing the alignments between different planar and zonal multipoles. Following the idea by de Oliveira-Costa et al. (2004) we introduce, a corresponding, yet slightly modified, real-space statistics as:

$$\mathcal{M}_\ell(\hat{\mathbf{n}}^\star) = \max_{\hat{\mathbf{n}}}\Big(\frac{\int d\Omega_{\hat{\mathbf{n}}'} T_\ell^2(\hat{\mathbf{n}},\hat{\mathbf{n}}')\cos^2(b_{\hat{\mathbf{n}}})}{\int d\Omega_{\hat{\mathbf{n}}'} T_\ell^2(\hat{\mathbf{n}},\hat{\mathbf{n}}')}\Big) \tag{7.4}$$

where the $b_{\hat{\mathbf{n}}}$ is the latitude in a coordinate system with z axis aligned along direction $\hat{\mathbf{n}}$ and measured from the plane perpendicular to $\hat{\mathbf{n}}$, and assuming a value of $\pi/2$ in the direction of $\hat{\mathbf{n}}$. $T_\ell(\hat{\mathbf{n}},\hat{\mathbf{n}}')$ is the tested multipole temperature map with north pole oriented at direction $\hat{\mathbf{n}}$. The integral is calculated over a whole sky with the CMB map north galactic pole rotated to a direction $\hat{\mathbf{n}}$ along the shortest 2-sphere geodesic. The exact properties of that rotation are not important, because the rotation of the CMB map around the $\hat{\mathbf{n}}$ axis only changes the phase information of the $a_{\ell m}(\hat{\mathbf{n}})$ coefficients, but does not lead to any transfer of power between different m modes. We assume a convention to store the directions found by the $\mathcal{M}$ estimator on the northern galactic hemisphere.

We implement a search over directions distributed over a hemisphere defined by the centres of the pixels defined in the Healpix pixelization scheme (Górski et al., 2005) of resolution yielding a total number of pixels of 49152 and thus the resulting effective search accuracy yields roughly 1°. The search is performed hierarchically, and exploits the hierarchical nature of the Healpix pixelization scheme, and so we begin the search in a very small resolution over a small number of directions, equally covering a hemisphere, and corresponding to large pixels. Then we choose the one that yields the largest value of $\mathcal{M}_\ell(\hat{\mathbf{n}}^\star)$ and then we increase the resolution for the chosen pixel in a nested way, so that the area of the pixel is equally divided into a new set of pixels defining a new set of directions over which the search with increased resolution is performed. This way we continue to increase the search resolution until the final targeted accuracy is reached.

The corresponding statistics, to the one defined in Eq. 7.4, but optimized for the search of the most zonal harmonics, is defined as:

$$\mathcal{N}_\ell(\hat{\mathbf{n}}^\star) = \min_{\hat{\mathbf{n}}}\Big(\frac{\int d\Omega_{\hat{\mathbf{n}}'} T_\ell^2(\hat{\mathbf{n}},\hat{\mathbf{n}}')\cos^2(b_{\hat{\mathbf{n}}})}{\int d\Omega_{\hat{\mathbf{n}}'} T_\ell^2(\hat{\mathbf{n}},\hat{\mathbf{n}}')}\Big) \tag{7.5}$$

Next, for multipoles ℓ and $\ell' = \ell + j$ we introduce a maximum-momentum axis alignment estimator as:

$$M_{\ell\ell'} \equiv M_\ell^j = \overrightarrow{\mathcal{M}}_\ell(\hat{\mathbf{n}}_\ell^\star) \cdot \overrightarrow{\mathcal{M}}_{\ell'=\ell+j}(\hat{\mathbf{n}}_{\ell'}^\star) \tag{7.6}$$

and a minimum-momentum axis alignment estimator as:

$$N_{\ell\ell'} \equiv N_l^j = \overrightarrow{\mathcal{N}}_\ell(\hat{\mathbf{n}}_\ell^\star) \cdot \overrightarrow{\mathcal{N}}_{\ell'=\ell+j}(\hat{\mathbf{n}}_{\ell'}^\star) \tan(\measuredangle(\hat{\mathbf{n}}_\ell^\star, \hat{\mathbf{n}}_{\ell'}^\star)) \tag{7.7}$$

where

$$\begin{aligned} \overrightarrow{\mathcal{M}}_\ell &\equiv \mathcal{M}_\ell \hat{\mathbf{n}}_\ell^\star \\ \overrightarrow{\mathcal{N}}_\ell &\equiv \mathcal{N}_\ell \hat{\mathbf{n}}_\ell^\star \end{aligned} \tag{7.8}$$

where the $\hat{\mathbf{n}}^\star$ define the directions that respectively maximize and minimize, the momentum estimators as defined in Eqs. 7.4 and 7.5. By varying the j value, the estimators measure degree of correlations between extremal momentum axes of different multipoles. We define corresponding tail "probabilities-to-reject": upper-tail probability for maximum-momentum estimator and lower-tail for minimum-momentum estimator as:

$$\begin{aligned} P_{\rm reject}(M_{\ell\ell'}^{\rm sim} < M_{\ell\ell'}^{\rm data}) &= \sum_{\substack{i, \\ M_{\ell\ell'}^{{\rm sim},i} < M_{\ell\ell'}^{\rm data}}} \frac{1}{N_{\rm sim}} \\ P_{\rm reject}(N_{\ell\ell'}^{\rm sim} > N_{\ell\ell'}^{\rm data}) &= \sum_{\substack{i, \\ N_{\ell\ell'}^{{\rm sim},i} > N_{\ell\ell'}^{\rm data}}} \frac{1}{N_{\rm sim}} \end{aligned} \tag{7.9}$$

The probabilities are inferred using the $N_{\rm sim} = 4000$ MC simulations.

Eq. 7.8 defines the weights $\mathcal{M}_\ell$ and $\mathcal{N}_\ell$ that are used in order to discriminate between cases when two axes are strongly aligned by chance or whether a close alignment is accompanied by large (small) values of maximum (minimum) momentum estimator. Alternatively one could assume the weights $\mathcal{M}_\ell$ and $\mathcal{N}_\ell$ values of a unity and thereby consider only the distributions of spatial alignments between axes assigned to different multipoles. To distinguish these differences we consider the two cases.

7.2.2 Results

Momentum: We first consider only the maximum and minimum momentum values as defined by Eqs. 7.6 and 7.7.

In Fig. 7.1 we show the derived probabilities to reject of the derived from data values of the $\mathcal{M}_\ell$ and $\mathcal{N}_\ell$ estimators. This quantifies anomalous individual multipoles are, compared to simulations, with respect to the shape of the multipole; in particular to the sectoral like multipoles possessing a large values of $\mathcal{M}_\ell$ and imposing a planarity features and to the zonal like multipoles having a small values of $\mathcal{M}_\ell$ and imposing a axial features. Note the correlation between the curves which shows that where $\mathcal{M}_\ell$ values are large and have large $P_{\rm reject} \equiv P(\mathcal{M}_\ell^{\rm sim} < \mathcal{M}_\ell^{\rm data})$, the $\mathcal{N}_\ell$ are small but also have large $P_{\rm reject} \equiv P(\mathcal{N}_\ell^{\rm sim} > \mathcal{N}_\ell^{\rm data})$.

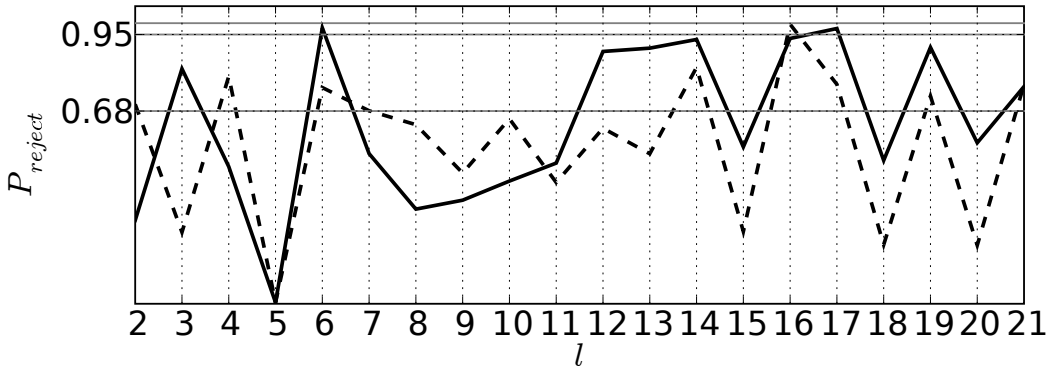


Fig 7.1: Distribution of $P_{\rm reject}$ for the values of $\mathcal{M}_\ell$ (thick-solid line) and $\mathcal{N}_\ell$ (thick-dashed line) in function of multipole number. The horizontal lines indicate the 68%, 95% and 99% CL contours.

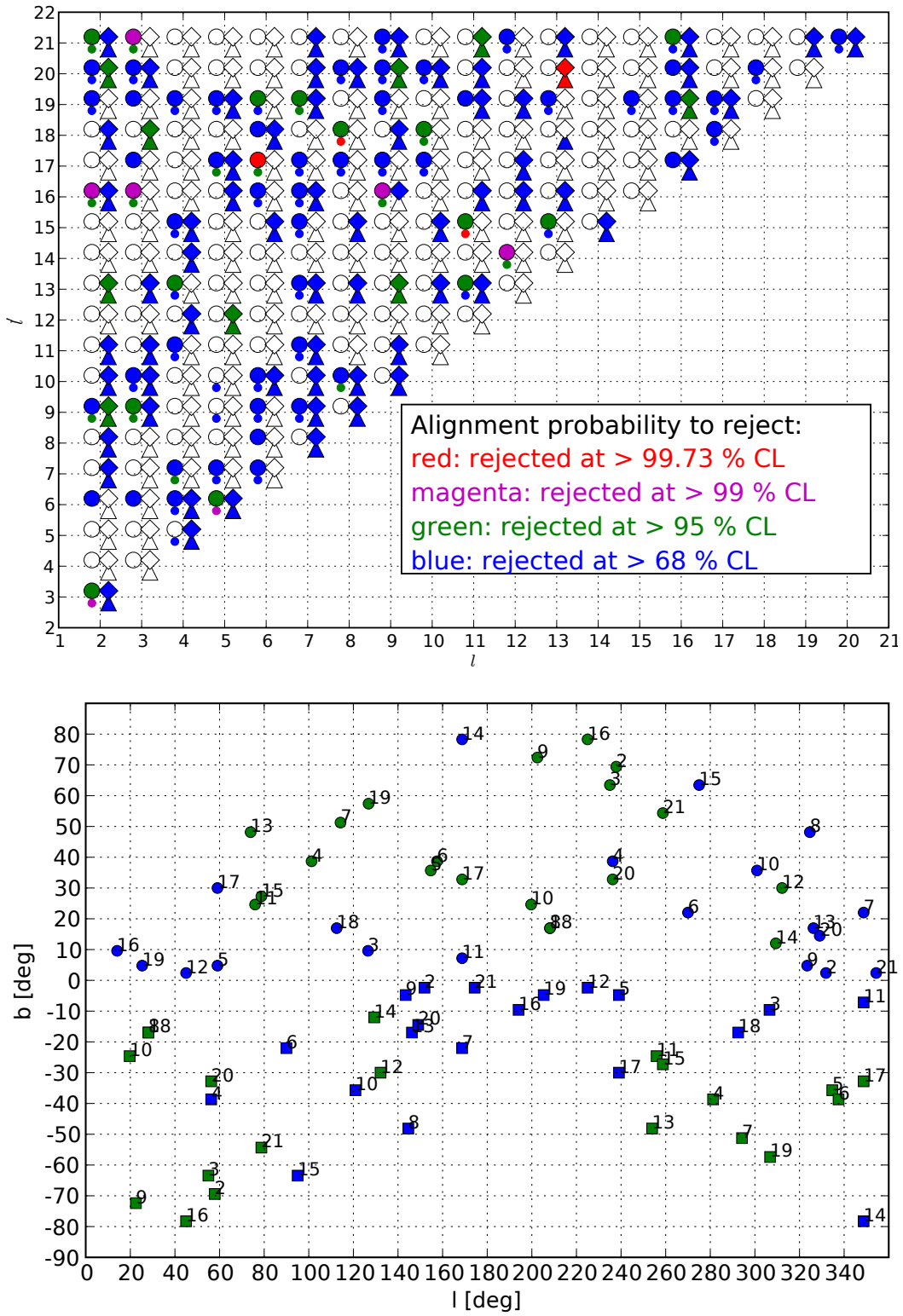


Fig 7.2: Alignments between pairs of multipoles of the three year WMAP ILC map, as measured by the maximum and minimum momentum estimators defined in Eqs. 7.6 and 7.7. In the top panel we plot the significance of alignments which is color-coded as indicated in the legend box of the plot. ($P_{\rm reject} < 68\%$ are plotted in white). For each pair of multipoles a quadruplet of probabilities to reject $P_{\rm reject}$ is given corresponding to maximum (left-hand side doublet of points) and minimum momentum (right-hand side doublet) estimators. The top doublet: circles and diamonds, represents the probabilities derived using weights defined in Eq. 7.8 and the bottom doublet: dots and triangles, represents the probabilities derived using a unit weights ($\mathcal{M}_\ell = 1$, $\mathcal{N}_\ell = 1$). In the bottom panel we plot the distribution of the preferred directions of each multipole as indicated by the numbers in the plot. The green (blue) dots indicate the maximum (minimum) momentum directions in the northern hemisphere; the opposite directions are marked with squares.

Alignments: In Fig. 7.2 (top panel) we plot the distribution of alignments between pair of two different multipoles $\ell < 20$ and $\ell' < 21$ as measured by the estimators $M_{\ell\ell'}$ and $N_{\ell\ell'}$. For each multipole pair, we plot a quadruplet of "probabilities of rejecting" ($P_{\rm reject}$) due to relative alignment between ℓ and ℓ' multipoles as derived from data and compared with the MC simulations. The four probabilities correspond to: maximum(minimum)-momentum estimator (Eq. 7.4), with weights defined either in Eq. 7.8 (large circles(diamonds)) of assuming a unitary weights $\mathcal{M}_\ell = 1$ and $\mathcal{N}_\ell = 1$ (small dots(triangles)). The combinations of multipoles that yield the $P_{\rm reject} \geq 99.73\%$ CL are marked in red. Multipoles that yield $99.73\% > P_{\rm reject} \geq 99\%$ are plotted in magenta. "Probabilities of rejecting" corresponding to 1σ and 2σ confidence levels are plotted in green and blue respectively and probabilities below 1σ significance level are plotted in white for clarity[4].

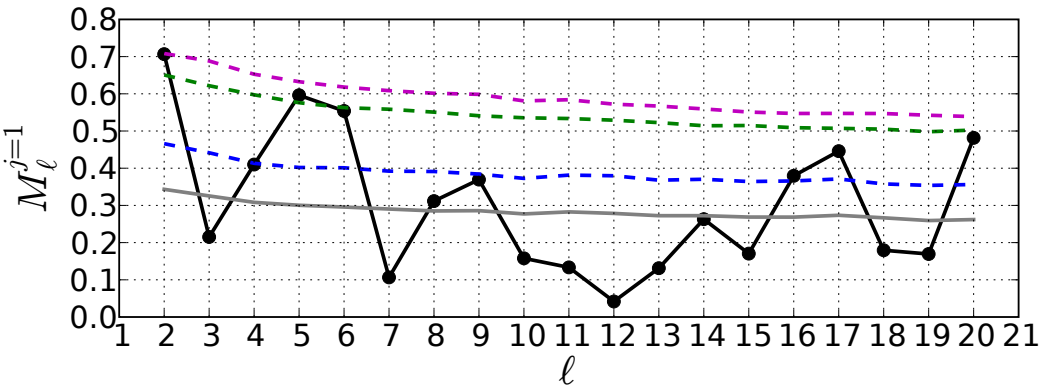


Fig 7.3: Alignments between pairs of multipoles of the three year WMAP ILC map, as measured by the maximum momentum estimator $M_\ell^{j=1}$ defined in Eq. 7.6. This defines a slice through the matrix in Fig. 7.2 with $M_{\ell,\ell'=\ell+1}$. The 68%, 95% and 99% confidence contours are plotted in the same color convention as in Fig. 7.2. The simumations mean is also plotted with gray, solid line.

Clearly the well known alignment between the quadrupole and the octupole is apparent (bottom-left corner of in Fig. 7.2) and is anomalous at 99.3% CL, however the estimator accounting for the proper weights of the derived momentum yields a smaller significance of $P_{\rm reject} \approx 98.87\%$ which translates onto a chance 1:88. The orientation of the maximized momentum is found to be $(l, b) = (238^\circ, 69^\circ)$ and $(l, b) = (235^\circ, 63^\circ)$ for $\ell = 2$ and $\ell = 3$ respectively which yields the angular separation of only 6° ($\cos(6^\circ) \approx 0.994$).

Interestingly, except for the pair $(\ell, \ell') = (13, 20)$, there are no statistically significant alignments detected via the minimum-momentum estimators. In particular we note an alignment between neighboring $\ell = 5$ and $\ell' = 6$, as detected via the maximum-momentum estimator. The angular separation between the two axes is only $\sim 3.8^\circ$ towards direction $(l, b) = (155^\circ, 36^\circ)$ for $\ell = 5$, which remains in no special relation to the quadrupole octupole alignment.

To complete the information gathered in the search in the bottom panel of Fig. 7.2 we

[4]Note that the white dots will not be seen in the plot.

plot the spatial orientation of the most preferred directions in the first 21 multipoles of the ILC three year data. In order to visualize the close alignments between multipoles of which directions in the northern hemisphere are separated roughly by 180° we plot both the directions in the northern hemisphere (dots) and the corresponding opposite directions found in the southern hemisphere (squares). It should be noted that generally the angular separation of the maximum momentum directions and the minimum momentum directions for the same multipole is generally close to 90°.

In particular the close alignment between $\ell = 2$ and $\ell = 3$ is clearly seen for the maximum momentum test with angular separation of only: 6° corresponding to a probability of $\sim 0.068\%$ as inferred from the area enclosed by a spherical cap of aperture 6°. Note also a very close alignment of $\ell = 5$ and $\ell = 6$. with angular separation of the maximum momentum axes of only 3.7°.

Joint significance: We close up this section by performing a joint statistic of all alignments between different pairs of multipoles. In order to robustly quantify whether the measured degree and number of strong alignments are consistent with the GRF expectations we combine the information about all measured alignments into a single χ^2 value using a full covariance matrix estimates from $N_{\rm sim} = 2000$ simulations.
We define the χ^2 value for the set of considered $M_{\ell\ell'}$ estimators as:

$$\chi^2_{(M_{\ell\ell'})} = (M_{\ell\ell'})^T (\mathbf{C}_{\ell\ell' ll'})^{-1} (M_{ll'}) \tag{7.10}$$

where

$$\mathbf{C}_{\ell\ell' ll'} \equiv Cov(M_{\ell\ell'}, M_{ll'}) = \left\langle \left(M^{\rm sim,i}_{\ell\ell'} - \langle M^{\rm sim,i}_{\ell\ell'}\rangle\right)\left(M^{\rm sim,i}_{ll'} - \langle M^{\rm sim,i}_{ll'}\rangle\right)\right\rangle_{\rm sim} \tag{7.11}$$

where $(M_{\ell\ell'})$ is a vector of $M_{\ell\ell'}$ values ordered in an ℓ-major format. Of course for a GRF the covariance matrix should be diagonal, however given an anisotropic noise (and in general any simulated foregrounds) this need not be the case. We constrain the $\ell_{\rm max} = 21$ and hence the total size of the (M) vectors, including every combination of multipoles pair, is 190. We analogically construct the χ^2 values for the minimum-momentum estimator ($N_{\ell\ell'}$) and χ^2 values for the case of unitary weights in these estimators defined earlier. This gives in total four χ^2 values, that for which a joint "probabilities of rejecting" are: $\{0.15, 0.47, 0.91, 0.82\}$ for maximum-momentum (unweighted and weighted) and minimum-momentum (unweighted and weighted) estimators respectively, which remains a good consistency with GRF simulations [5]. This result indicates that the WMAP three year ILC data do not violate the statistical isotropy or Gaussianity by an excessive total number and degree of an accidental inter-multipole alignments if measured in a broad range of multipoles. As was shown earlier this is not the case for the selected pairs of multipoles. The estimation of true oddity then must involve some prejudice or priors on which correlations: i.e. between which pairs of multipoles are to be taken seriously and which can be treated as

[5] We used $N_{\rm sim} = 2000$ independent MC simulations of probe the distribution of the underlying probability distribution functions of the simulated χ^2 values to derive these probabilities.

unimportant. Of course such a prior seem natural in case of closely neighbouring multipoles $\ell = 2$ and $\ell = 3$ or $\ell = 5$ and $\ell = 6$, and seem unnatural in case of $\ell = 13$ and $\ell = 20$.

We note that the increased "probability of rejecting" in case of the unweighted minimal momentum estimator is strongly build up by the contribution to the total χ^2 value from the alignments between multipoles $\ell = 13$ and $\ell = 20$ (see Fig. 7.2). This is clearly seen when looking at contributions per degree of freedom, which in this case yield as much as $\Delta\chi^2 > 10$.

Generalizations: Of course, even in case of a Gaussian random filed simulation, some accidental correlations between pairs of different multipoles are expected. However an analysis presented in the previous section, is not sensitive to cases where the inter-multipole alignments span across range of multipoles, and where all of which could result from the same underlying process. Based on Fig. 7.2 however this doesn't seem to be the case. Nevertheless a more general statistic, optimized for joint measurements of alignments spanning over a range of multipoles could be an interesting extension to the two-multipole case.

We also note that given the quantity defined in Eq. 7.4, and the corresponding alignment estimator defined in Eq. 7.6 have somewhat limited freedom, because we choose to maximize the $\mathcal{M}_l$, instead of the estimator $M_{\ell\ell'}$ itself. It is possible though that the estimator could reach greater values for the two, or in generalized case for a range of multipoles, at cost of a slight decrease of the $\mathcal{M}_\ell$ and $\mathcal{M}_{\ell'}$ values, but covered by an increase due to larger $\cos(\angle(\mathcal{M}_{\ell_i}, \mathcal{M}_{\ell_j}))$ factor. Analogical situation is possible with the minimized $\mathcal{N}_l$ defined in Eq. 7.5 and the corresponding alignment estimator $N_{\ell\ell'}$ (Eq. 7.7). While an accidental alignments are possible in an GRF, an accidental alignments of more than two neighbouring multipoles are way less probable.[6], and hence such an extension applied to this is other modified statistic could be introduced to further tests the statistical isotropy hypothesis.

While we will also return to the more generalized statistics in Sec. 7.3

7.3 Testing the "m-preference"

In this section we perform a search of the mode-to-mode alignments in variety of renditions of the Internal-Linear-Combination maps using statistics defined in equations 7.2 and 7.3. Our aim will be to test the stability of the alignments with respect to different foregrounds cleaning methods, which independently constrain the best approximations of the intrinsic CMB fluctuations.

7.3.1 Data

We choose to test and compare alignments in five different renditions of the WMAP CMB data: the three-year and the five-year Internal-Linear-Combination maps (Hinshaw et al.,

[6] The situation is similar to eg. nuclear reaction rates where the probability of an instant three particles fusion eg. of two protons and neutron into a $^3_2$He nuclei is thousands of times less probable than much more frequent channels of two particle reactions.

2007; Gold et al., 2008) available at
http://lambda.gsfc.nasa.gov/data/map/dr2/dfp/wmap_ilc_3yr_v2.fits and
http://lambda.gsfc.nasa.gov/data/map/dr3/dfp/wmap_ilc_5yr_v3.fits. We will call these maps ILC3 and ILC5 respectively. We also compere the five year edition of the Wiener filtered map provided by Max Tegmark (Tegmark et al., 2003) available at
http://space.mit.edu/home/tegmark/wmap/wiener5yr_map.fits which we will refer to as TOH5. Finally we analyse the so-called harmonic internal linear combination map (HILC5) available at http://www.nbi.dk/ jkim/hilc/HILC5YR_smo.fits (Kim et al., 2008) and the five year KQ75 cut sky version of the inverse noise coadded map from the V channel (Hinshaw et al., 2008) which we will call V5. We convert these data into the spherical harmonics space and extract the multipoles from $\ell = 1$ to $\ell = 6$. This is the range within which we will seek for the mode-to-mode alignments. In the following we describe each of the considered data in more detail:

ILC3 and ILC5 The WMAP Internal Linear Combination method is a model independent, self contained method of estimating the pure CMB component from the foregrounds contaminated maps by a regional variance minimization of the linear combination of one degree pre-smoothed multi-frequency observations. The principal idea is to exploit the frequency dependence of the foregrounds component and frequency non-dependence of the CMB component. The analysis is performed in separate regions exclusively, which are defined according to the spectral properties of the foregrounds emissions. The regional analysis with sharp boundaries however leads to apparent discontinuities of a resulting map. Also the method itself leads to a somewhat biased estimate of the true CMB with variance $\sigma_{\rm CMB} > \sigma_{\rm ILC}$ due to arbitrary non-zero covariance between the foregrounds and the CMB. This bias in the post-processing is accounted for statistically in each region individually.

Since the frequency band maps are pre-smoothed to a common resolution of one degree, the resulting map has an effective resolution of one degree, which when compared to the resolution of the WMAP radiometers ($\sim 0.22^\circ$ for W band) is rather sub-optimal.

It is also important to note that given that the area outside the Kp2 galactic plane cut is treated as one single region, the minimization of a variance in such a large ares might lead to somewhat inaccurate results since the foregrounds emissions may in general vary spatially, although the initial division of sky into regions was performed exactly regarding the spectral properties of these emissions. Also, given that the variance minimization is performed directly in real space, the high-ℓ multipoles of the map ultimately receive too much weight, relative to the amount of CMB signal they carry and the amount of the noise by which they are dominated, as compared and the amount of the CMB variance carried by the lower multipoles.

V5 As mentioned the V5 data is the foregrounds reduced version of the inverse noise coadded (see Section 3.3) map using the five-year observations from the V1 and V2 spectral channels of the WMAP. The map is basically divided into two regions defined by the masked

part of the KQ75 sky mask and the unmasked part. The KQ75 sky mask in turn, which is used in this analysis, is obtained via combining the thresholded, one degree smoothed foreground templates from the K and Q bands at the levels that effectively leave 75% of the sky unmasked. The foregrounds template for each band in turn is obtained either from the three-year version of the ILC map (ILC3), which is a good estimate of the true CMB signal at scales larger than 1°, or from modelling foregrounds based on external observations. The foregrounds reduction outside the KQ75 sky region is performed by subtracting the foregrounds templates for the synchrotron and free-free emissions estimated independently from observations of the 408MHz radio continuum emission (Haslam et al., 1982) and the all sky H_α observations of the Finkbeiner (2003) respectively. For the dust template the 94 GHz radio emission model of Finkbeiner et al. (1999) is used. The point source emission was effectively masked out using point source detections in radio and X-rays observations of Stickel et al. (1994); Hirabayashi et al. (2000); Kuehr et al. (1981); Teräsranta et al. (2001); Perlman et al. (1998); Landt et al. (2001).

The main drawback of these maps as discussed in Delabrouille et al. (2008) is that the regions inside the Galactic plane are heavily contaminated even after foregrounds template removal, which give more priority to cleaning the high latitude regions in order to obtain relatively clean CMB maps, which when masked with the eg. KQ75 sky mask will be sufficiently clean to constrain the CMB power spectrum in the limit of high multipoles. Outside the region defined by the sky mask, the foregrounds removal accuracy is model dependent, and as such imperfections of the model will be imprinted as residual foregrounds. As a result in the limit of low multipoles large differences should be expected between these maps and other full sky CMB map estimates. The noise and smoothing properties however are very well known of which statistical properties are easy to assess via appropriate Monte Carlo simulations.

TOH5 In order to obtain the TOH5 map the five-year observations from the five spectral channels were combined into a "cleaned" maps and wiener filtered in the spherical harmonics space by multiplication of each of the $a_{\ell m}$ coefficients of the input "cleaned" map by the Wiener filter kernel: $\frac{C_\ell b_\ell^2}{C_\ell b_\ell^2 + N_\ell}$ in order to obtain the TOH5 maps $a_{\ell m}$ coefficients, where the b_ℓ is the instrumental beam transfer function and the C_ℓ and N_ℓ are the constrained CMB and noise power spectra respectively.

The foregrounds reduction is performed by linear combination map variance minimization which is performed independently for each multipole and exclusively in 9 selected, different regions in the sky which are defined according to the spectral properties of the cross-band difference maps.

As a result the cleaned, filtered map has an optimal resolution of the WMAP W channel, which is approximately four times better than the resolution of the ILC3/5 maps. It should be noted that the Wiener filter can only be adequately applied if the shape of the underlying power CMB power spectrum is known: i.e. if it is constrained using other methods (like pseudo-C_ℓ method).

HILC5 The HILC5 data is obtained from an implementation of the foregrounds removal method, based on variance minimization, where the weight coefficients are allowed to vary smoothly across the sky, which solves the discontinuities problems of the ILC3/5 maps, however the spatial resolution of these coefficients is constrained by the maximal cut off scale multipole of $\ell_{\max} = 15$, which possibly constrains the ability to resolve the smooth galactic emissions of angular sizes smaller than the effective angular size to which multipoles of $l \leq 15$ are sensitive (roughly 12°).

Due to deconvolution of the single band maps by the instrumental beam prior the cleaning process, the resulting map has increased power in the limit of large multipoles as compared to the TOH5 map, and the variance minimization process is driven mainly by minimization of noise rather than foregrounds.

7.3.2 Statistics

We perform the statistics as defined in equation. 7.2 and 7.3: i.e. for each mode independently, we look for such orientations of the map, so that the amount of power in that mode is maximized. For each maximized mode we record the corresponding direction maximizing the power in that mode and store it for further tests and for comparison purposes.

We perform the search by rotating the maps in the real space and measuring the amounts of power absorbed by a given mode in spherical harmonic space. We use the search direction density corresponding to the Healpix pixelization scheme of resolution parameter $n_s = 128$ which yields the approximated pixel size and consequently the search accuracy of about $\sim 0.45^\circ$. The maps are generated in the resolution $n_s = 64$.

For each pair of mode best fit orientations $\hat{\mathbf{n}}^\star$, we calculate and compare the angular separations. We also maximize according to eq. 7.2 to derive the cross-multipole alignments. In the next sections, we present the details of these searches, discuss them and consider a approaches to significance analyses and their generalizations.

7.3.3 Results

In table 7.1 we present a compilation of the results obtained from the statistics defined in Eq. 7.3 using the considered data set. Note that for the case of dipole there exists the exact degeneracy as it was mentioned in Section 7.1.

Firstly, we notice a strong competition between different modes in terms of the amount of the absorbed power towards different directions. The close alignments between different modes is visualized in figures 7.6 and 7.7 and also in tables 7.2, 7.3 and 7.4 where the alignments within 10° are highlighted in blue. For completeness we also highlighted the perpendicular alignments that are greater than 80° in green. Note however that the perpendicular alignments have additional degree of freedom over the parallel alignments, and as such it is expected, and indeed observed, that they will appear more numerously than the parallel ones. Also, in order to ease the comparison between different data and modes, we present graphically the distribution of the accumulated power in figure 7.8.

Alignments: It is clear that some well known close alignments are found in different renditions of the CMB observations. Most notably, the consistent through all data sets are the close modes-to-mode alignments between: a_{21}-a_{31} and the well known sectoral harmonics alignment a_{22}-a_{33}, which both yield angular separation less than 10°. An exception is the HILC5 map in which the alignment of these modes is larger than 10° in both cases, and in case of pair a_{21}-a_{31} the separation is as large as 64°.

The higher modes also yield some close alignments. In particular, the previously reported in Land & Magueijo (2005b) alignment between a_{53}-a_{40} which is additionally aligned with pair the a_{22}-a_{33}, is confirmed. All of these modes lie roughly in direction $(l, b) = (250^\circ, 50^\circ)$ within accuracy of about 15°. This feature of preferring roughly the same reference frame by independent within the GRF hypothesis modes, was found to be unusual and the direction itself is dubbed the "axis of evil".

It is interesting to see whether the difference between the data sets are strongly multipole or mode dependent. To see this we consider the standard deviation of differences of the orientations obtained from different data sets. In Fig. 7.4 we plot the standard deviations of differences taken in all 10 possible combinations of pairs of data as a function of a mode. Note that while the differences between renditions of the foreground cleaned maps range up to roughly 50°, as measured by the standard deviation, the differences for the most aligned modes: i.e. for the pairs a_{22}-a_{33} and a_{40}-a_{53} and also partially for a_{21} and a_{31} are very small [7]. The latter ones are only small for the case when the V5 data is not considered cross-data in calculating the dispersion of the preferred separations across different data-sets. In fact, given that the V5 data, are extensively masked, they should not be, trusted to be reliable in the limit of the largest scales, and are presented here, and in what follows, only for the comparison purposes.

Although the existence of the alignments seems to be obvious in different data sets, given the notes put in section 7.1, the question of which particular axis of evil is preferred seems less obvious. Namely, the modes a_{22}-a_{33}, which absorb a very large fraction of the total quadrupole and octupole power respectively in an preferred reference frame oriented towards $\sim(l, b) = (240^\circ, 50^\circ)$, very strongly compete with the more zonal-like modes a_{21} and a_{31} which absorb almost the same fraction of the total quadrupole and octupole power but in an completely different reference frame, oriented towards $\sim(l, b) = (115^\circ, 15^\circ)$. This has also been realized in Land & Magueijo (2007) and was referred as "*as instability*". Given however that the a_{31} mode is dubious in light of the fact that different foreground cleaning methods lead to a results which give larger uncertainties between the data-sets, as compared to the aforementioned sectoral mode pair (a_{22}-a_{33}), and also given that the r_ℓ values are slightly larger for that pair, tentatively the "axis of evil" as unveiled by the maximal momentum analysis (section 7.2.1) seems eventually more preferred towards direction roughly $\sim(l, b) = (240^\circ, 50^\circ)$ over the direction preferred by the zonal harmonics pair.

[7] The dispersion within a_{31} is actually about 30°, while the dispersion for r_{21} is large only when the V5 data are considered.

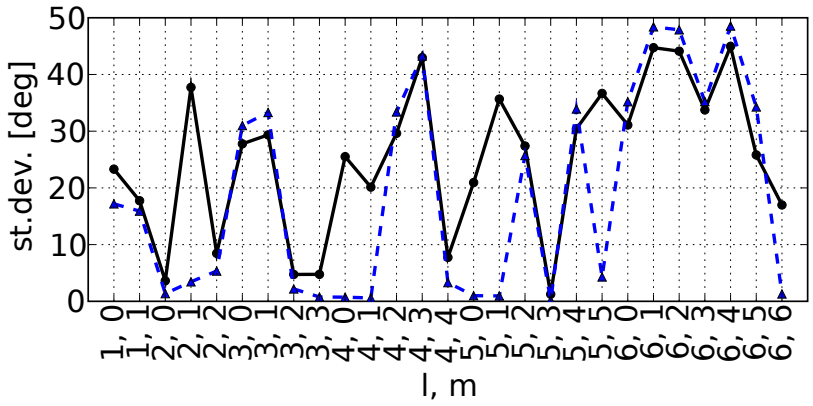


Figure 7.4: Cross-data-set consistency check. For each ℓ, m value we plot with black dots (blue triangles) a standard deviation of all 10(6), possible within all considered (all but V5) data maps, combinations of angular separations between the directions most preferred by a given mode. Note the differences between the curves resulting from inconsideration of the sky-cut V5 data which is most unreliable in the limit of the largest scales.

Alignment reconstruction degeneracy: It is important to mention the issue of certain degree of degeneracy of the statistics as defined in Eq. 7.3 in terms that the resulting maximized value of $r_{\ell m}(\hat{\mathbf{n}}^{\star})$ may actually be one of many possible solutions, which however will drastically differ in regard to the corresponding preferred axis orientation. In order to visualize this, as an example, in figure 7.5, we plot the map of the $r_{\ell m}(\hat{\mathbf{n}})$ coefficients over which the statistics maximize to find the final solution. Note that the maps have a point symmetry resulting from the fact that the maximal possible angle between two axes is 180°. We recall that according to our convention we consider only solutions in the northern galactic hemisphere.

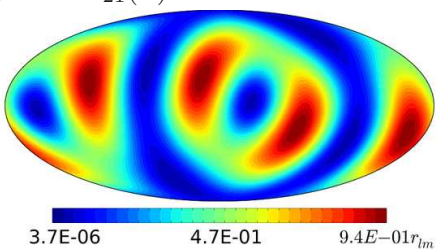
Note that, as indicated in panel d) of the figure 7.5, contrary to what it would seem from the panel b), there is a substantial difference in the amplitude of the $r_{\ell m}$ coefficients in the series of maximums, and hence the maximization does not loose significant alternative solutions in this case. However as shown in the panel c) this is not the case for some of the modes of the lower multipoles. It should be noted however that in general, the alternative solutions could also be accounted for in the analysis devoted to the searching the alignments and appropriately accounted for in the significance analysis by weighting eg. by the amount of absorbed power: i.e. by $r_{\ell m}(\hat{\mathbf{n}}^{\star})$.

These issues has also been previously pointed out in Land & Magueijo (2007).

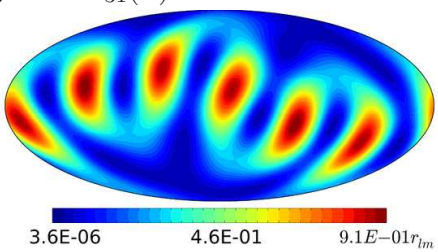
7.3.4 Significance and generalizations

Although we have not studied the issue of the statistical significance of the alignments, we defer this analysis for a future work. For that purpose we have implemented an parallelled algorithm for the general search of the preferred directions, which can be run on an arbitrary number of cluster nodes, allowing thereby for speeding up the significance assessment.

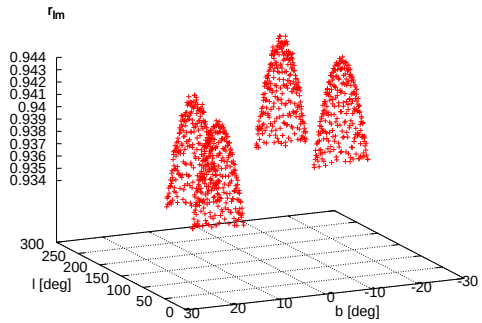
a) ILC5: $r_{21}(\hat{\mathbf{n}})$



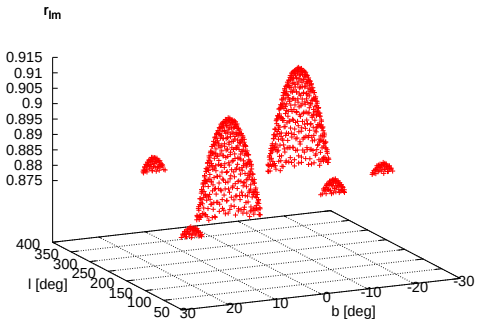
b) ILC5: $r_{31}(\hat{\mathbf{n}})$



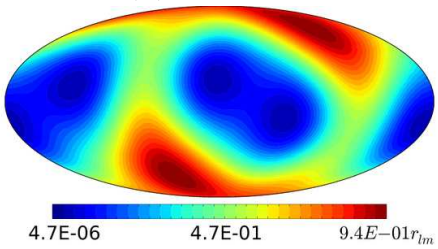
c) thresholded ILC5: $r_{21}(\hat{\mathbf{n}})$



d) thresholded ILC5: $r_{31}(\hat{\mathbf{n}})$



e) ILC5: $r_{22}(\hat{\mathbf{n}})$



f) ILC5: $r_{33}(\hat{\mathbf{n}})$

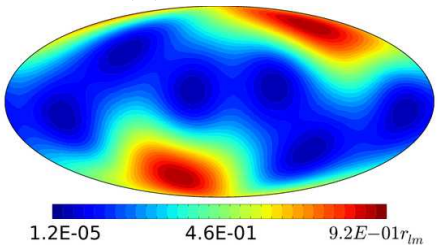


Fig 7.5: Distribution of $r_{\ell m}(\hat{\mathbf{n}})$ coefficients (Eq. 7.3) derived from the ILC5 data. Note that the thresholded maps indicate that for the low multipoles the possible degeneracy in due to multiple directions yielding a very high values of $r_{\ell m}(\hat{\mathbf{n}})$ in their own reference frames, may mislead the alignments analysis if not properly accounted for (panel c). The effect for higher multipoles (eg. panel d) the degeneracy appears to be smaller, although more solutions with close values of $r_{\ell m}(\hat{\mathbf{n}})$ are possible. The distributions in panels e) and f) visualize the well-known quadrupole-octupole alignment problem. Note that in this case there is no degeneracy and there exist only one solution over the entire hemisphere.

In the following we would like to point out a few generalizations that have not yet been pursued. It would be interesting to question the significance of the alignments in an analogical manner as it was put in section 7.2.1. For that purpose an estimator should be defined equivalent to the one defined in equation 7.6. We propose a following mode to mode alignment estimator:

$$R_{\ell m \ell' m'} = \overrightarrow{r}_{\ell m}(\hat{\mathbf{n}}^{\star}) \cdot \overrightarrow{r}_{\ell' m'}(\hat{\mathbf{n}}^{\star\prime}) \tag{7.12}$$

where

$$\overrightarrow{r}_{\ell m}(\hat{\mathbf{n}}^\star) = r_{\ell m}(\hat{\mathbf{n}}^\star)\frac{\hat{\mathbf{n}}^\star}{|\hat{\mathbf{n}}^\star|} \tag{7.13}$$

and the full covariance matrix is defined as:

$$\begin{aligned}\mathbf{C}_{\ell m\ell' m' LML'M'} &\equiv Cov(R_{\ell m\ell' m'}, R_{LML'M'}) \\ &= \left\langle \left(R^{\mathrm{sim,i}}_{\ell m\ell' m'} - \langle R^{\mathrm{sim,i}}_{\ell m\ell' m'}\rangle\right)\left(R^{\mathrm{sim,i}}_{LML'M'} - \langle R^{\mathrm{sim,i}}_{LML'M'}\rangle\right)\right\rangle_{\mathrm{sim}}\end{aligned} \tag{7.14}$$

The corresponding χ^2 test would take into account all possible alignments and their magnitude, proportional to the $r_{\ell m}$ value, and quantify over a range of considered multipoles whether or not the field yields consistency with the GRF expectations, avoiding the problem of the operating only on an maximized statistics given by Eq. 7.2 and the problem of ambiguity of closely competing modes mentioned earlier. The problem of the multiple also strongly competing solutions though should require a separate attention.

Finally, following the idea introduced in section 7.2.1 even further generalisation is possible by changing the statistic to maximize the alignment estimator itself ($R_{\ell m\ell' m'}$), rather than the $r_{\ell m}$ values.

Maximum power test Additionally it could be interesting to introduce a new joint statistics based on results from the m-preference search. If there exists an significant m-preference in power distribution for some reference frames, then the total power as measured by the "maximized variance map" synthesized from the r_{lm} coefficients (Eq. 7.3) should be statistically larger than the one obtained from GRF simulations. Therefore measuring a quantity:

$$\begin{aligned}\sigma_{\ell_{\max}} &= \sum_{\ell=2}^{\ell_{\max}} \sigma_\ell, \qquad \text{where,} \\ \sigma_\ell &= \tfrac{(2\ell+1)}{4\pi}\sum_m |r_{lm}|^2\end{aligned} \tag{7.15}$$

could in principle be also sensitive to this kind of anomalies. Note that such defined estimators are completely independent from the underlying power spectrum properties, and hence from the point of view of Monte-Carlo simulations any large scale discrepancies from the ΛCDM model have no impact on the results of the question of the statistical isotropy. We defer the implementation of this test for possible future work.

			ILC3			ILC5			V5			TOH5			HILC5		
#	ℓ	m	l [deg]	b [deg]	$r_{\ell m}$ [%]	l [deg]	b [deg]	$r_{\ell m}$ [%]	l [deg]	b [deg]	$r_{\ell m}$ [%]	l [deg]	b [deg]	$r_{\ell m}$ [%]	l [deg]	b [deg]	$r_{\ell m}$ [%]
1	1	0	206.0	38.3	100.00	350.8	63.1	100.00	200.0	37.9	100.00	232.4	6.6	100.00	355.8	19.8	100.00
2	1	1	147.7	17.3	100.00	81.6	39.1	100.00	294.6	5.7	100.00	108.3	78.3	100.00	93.2	19.5	100.00
3	2	0	333.3	3.3	93.58	333.6	2.4	92.18	149.8	5.7	82.97	332.9	4.8	96.06	152.9	0.0	82.80
4	2	1	110.7	12.6	93.09	113.5	17.6	94.38	2.1	25.6	99.12	115.3	20.4	89.95	109.3	10.8	99.10
5	2	2	236.1	69.0	93.11	239.1	60.4	94.39	245.9	43.0	99.09	236.6	55.5	89.99	242.4	73.9	99.10
6	3	0	148.0	9.0	65.17	147.3	9.0	64.83	90.7	21.7	71.86	15.8	27.9	64.31	149.8	7.5	69.20
7	3	1	116.0	14.8	91.30	116.0	14.8	91.06	126.9	24.6	82.70	50.6	28.3	92.44	118.5	14.5	93.42
8	3	2	208.8	9.3	72.62	208.8	9.3	72.36	217.6	1.2	92.75	209.2	12.6	62.99	209.5	5.7	73.10
9	3	3	236.6	62.7	91.74	236.6	62.7	91.94	252.3	55.1	68.20	240.2	61.6	96.55	236.7	61.6	90.84
10	4	0	258.5	44.6	77.32	259.2	44.6	76.06	202.5	22.4	71.86	259.1	45.0	72.84	257.3	43.4	81.76
11	4	1	228.2	35.0	90.07	228.9	34.2	90.24	29.2	1.8	76.84	227.8	36.0	88.65	227.5	33.5	83.36
12	4	2	196.2	58.2	80.87	197.3	58.2	82.33	241.2	11.7	87.34	196.9	61.6	83.60	239.8	5.1	81.16
13	4	3	335.7	25.3	72.25	77.0	15.7	73.06	336.8	21.7	85.67	73.8	14.2	76.56	332.9	26.6	76.60
14	4	4	116.7	31.0	64.82	116.7	31.0	64.73	120.9	11.1	64.48	116.0	28.3	69.58	124.5	27.6	59.35
15	5	0	236.9	3.9	62.43	236.9	3.9	62.40	209.2	37.9	63.98	237.3	5.4	56.48	239.4	4.8	61.64
16	5	1	257.7	4.8	57.01	257.7	4.8	56.51	196.0	55.5	74.07	258.4	4.8	58.08	259.4	6.3	65.49
17	5	2	130.8	2.7	66.46	350.2	29.0	65.44	78.4	18.2	76.08	130.4	3.6	67.76	132.9	2.1	68.65
18	5	3	264.0	48.1	85.75	264.0	47.8	85.95	268.8	47.8	82.82	264.8	47.8	91.22	264.7	48.5	90.84
19	5	4	173.0	5.1	64.81	173.0	5.1	65.12	166.6	2.1	77.14	171.2	6.6	76.07	111.8	37.5	67.65
20	5	5	154.6	44.2	37.58	154.6	44.2	37.69	327.7	29.0	59.16	162.2	49.3	37.66	159.2	55.5	42.90
21	6	0	261.9	3.0	52.73	261.9	3.0	53.54	280.6	29.0	66.39	278.8	31.4	54.45	357.2	47.8	54.43
22	6	1	269.6	17.6	78.17	269.6	17.6	78.14	271.4	14.8	76.71	269.3	17.9	75.93	19.4	43.8	75.46
23	6	2	68.2	21.1	68.54	68.2	21.1	69.48	309.9	43.8	66.82	70.3	21.7	67.31	314.6	44.2	68.84
24	6	3	242.6	36.0	56.96	242.2	35.7	57.09	121.3	6.6	68.96	120.6	1.8	61.84	244.0	35.3	55.39
25	6	4	316.1	9.6	61.12	316.1	9.6	60.51	218.3	33.5	79.55	316.1	10.8	66.26	218.3	40.2	50.14
26	6	5	155.4	2.1	55.72	208.3	60.1	56.58	212.3	10.5	64.76	203.6	59.7	66.37	158.2	2.1	56.11
27	6	6	157.5	38.3	79.84	157.8	39.5	78.95	178.9	9.0	57.44	160.3	36.8	70.44	157.2	39.5	85.62

Table 7.1: Comparison of the different renditions of the WMAP foregrounds cleaned CMB data in terms of the mode-to-mode alignments. Note the perfect degeneracy between the two modes of the dipole component.

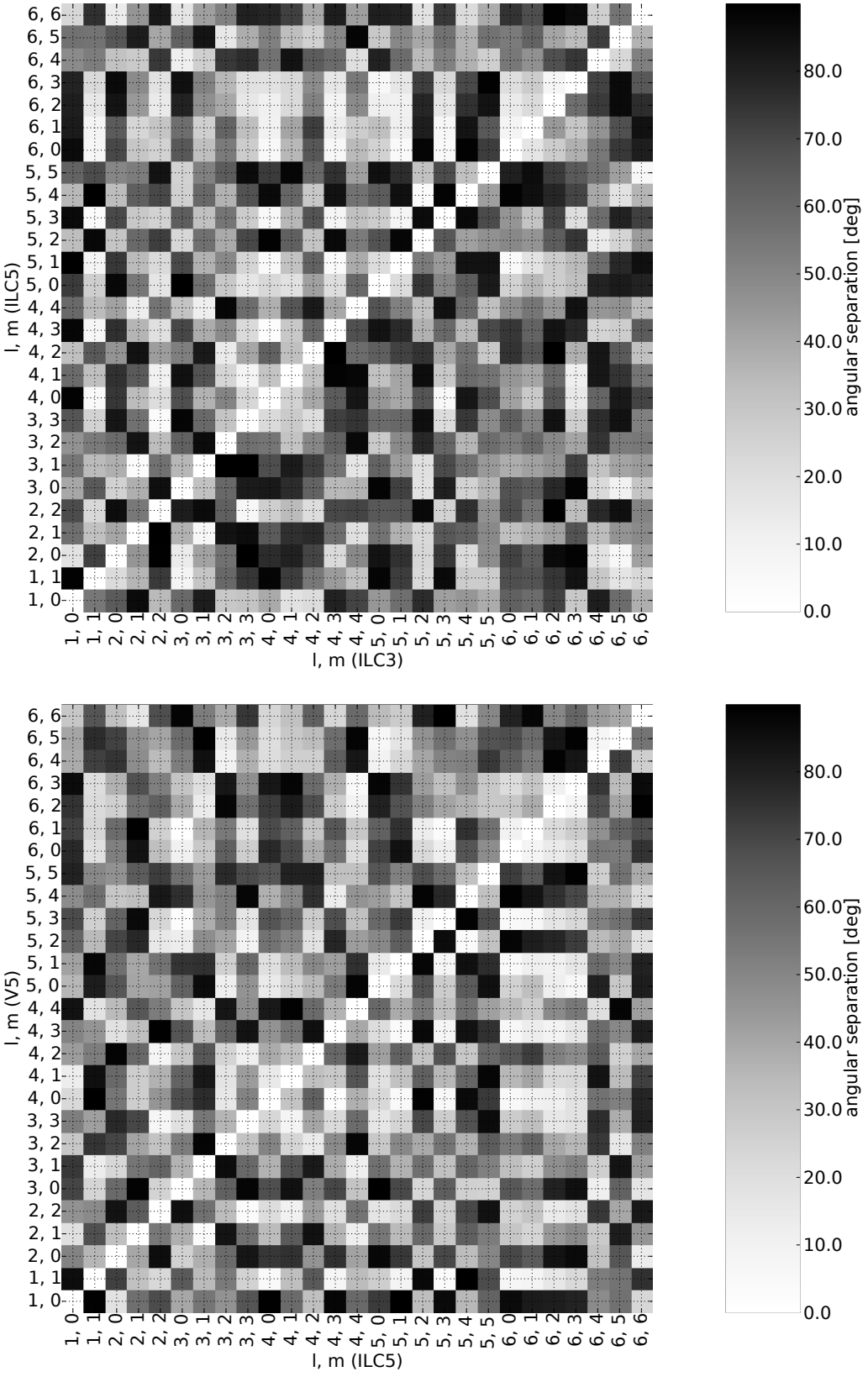


Fig 7.6: Mode to mode angular separation. Maximum power inter-mode ($\vec{r}_{\ell m}(\hat{\mathbf{n}}^{\star})$) alignments comparison in the ILC3-ILC3 (top panel) (below diagonal), and ILC5-ILC5 (above diagonal) data, and (bottom panel) ILC5-ILC5 (below diagonal), and V5-V5 (above diagonal) data.

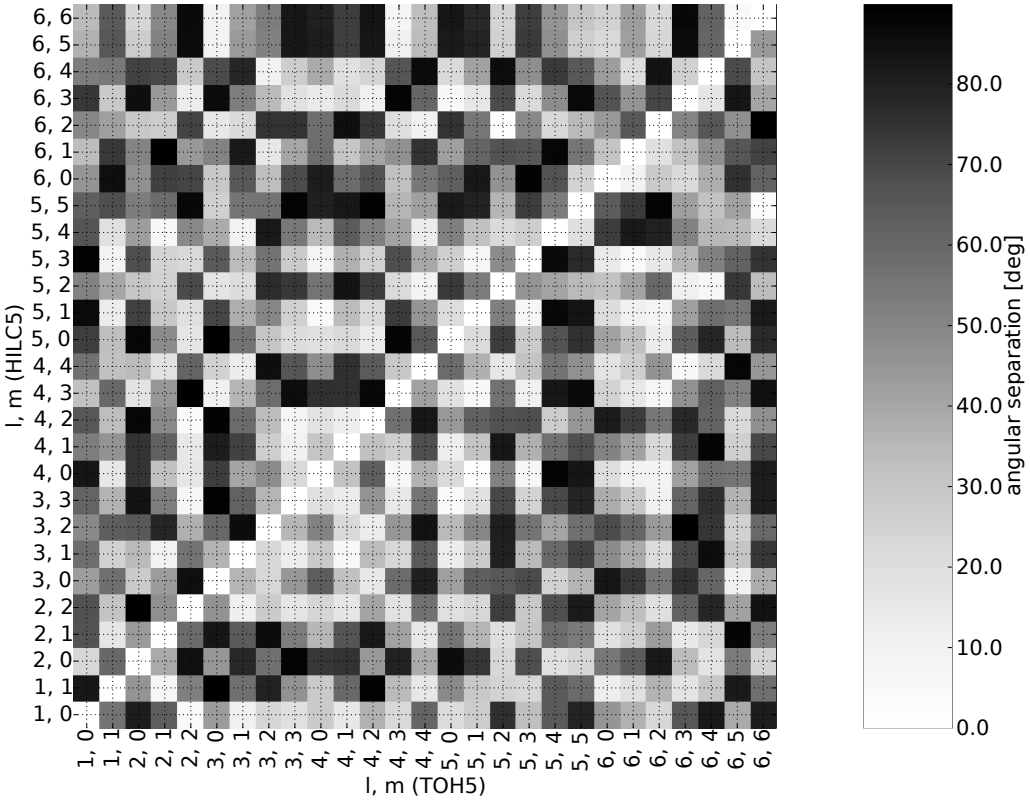


Fig 7.7: Mode to mode angular separation. Maximum power inter-mode ($\vec{r}_{\ell m}(\hat{\mathbf{n}}^\star)$) alignments comparison in the TOH5-TOH5 (below diagonal), and HILC5-HILC5 (above diagonal).

	l	1	1	2	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6	6	6
l	m	0	1	0	1	2	0	1	2	3	0	1	2	3	4	0	1	2	3	4	5	0	1	2	3	4	5	6
1	0	**0.0**	54.7	64.0	**86.3**	34.7	59.5	**80.9**	29.1	30.6	39.2	18.1	20.9	79.1	70.9	44.4	57.5	76.7	42.4	44.7	38.5	61.9	58.7	71.4	29.1	**80.6**	58.6	37.6
1	1	**89.2**	**0.0**	21.3	36.0	73.3	**8.3**	30.5	59.8	74.2	**88.1**	72.6	54.1	43.3	31.3	**88.1**	72.5	22.1	**86.5**	27.6	27.5	68.0	67.0	74.4	**83.8**	29.2	17.0	22.7
2	0	18.2	72.0	**0.0**	45.2	**89.5**	13.4	41.1	56.7	**89.9**	76.9	79.6	70.3	22.1	48.9	**83.8**	75.4	23.3	73.8	21.4	47.5	71.3	64.0	**86.6**	**88.6**	18.2	**5.8**	41.8
2	1	57.5	31.9	40.1	**0.0**	**89.9**	36.8	**5.6**	**84.2**	**86.1**	64.2	75.9	77.0	58.0	19.2	55.9	37.1	22.2	65.2	61.9	49.0	32.6	36.7	41.4	66.5	33.6	45.5	48.7
2	2	68.5	22.6	**85.5**	54.4	**0.0**	**80.9**	**86.3**	62.3	**6.3**	26.9	34.3	20.2	69.8	70.7	65.1	65.8	**87.1**	25.0	75.8	46.5	68.2	55.5	**89.5**	33.2	77.5	**84.7**	50.6
3	0	39.8	64.3	25.4	37.6	**85.3**	**0.0**	31.8	60.0	**81.4**	**82.2**	76.8	61.3	35.1	36.5	**88.3**	71.4	18.2	**80.0**	25.1	35.7	67.0	63.5	77.3	**88.4**	22.1	10.1	30.5
3	1	55.1	34.3	37.6	**2.6**	56.9	35.2	**0.0**	**89.7**	**89.9**	68.5	**81.2**	72.3	55.7	16.2	61.5	42.7	19.0	69.1	56.8	44.0	38.2	41.6	45.8	71.6	31.5	40.9	43.4
3	2	46.3	53.8	58.8	**83.4**	33.4	61.9	**86.0**	**0.0**	57.0	55.4	31.2	49.9	62.2	**87.0**	28.4	48.7	77.7	60.2	35.8	58.2	53.1	59.5	49.2	40.7	74.8	53.6	54.2
3	3	66.2	25.1	**83.0**	56.9	**2.6**	**87.9**	59.4	30.9	**0.0**	22.0	28.2	20.1	71.7	74.8	58.8	60.0	**85.2**	21.0	73.6	48.3	62.6	50.5	**84.3**	27.0	76.7	**84.1**	51.8
4	0	**89.2**	**4.4**	74.0	34.0	20.7	67.9	36.6	50.4	23.1	**0.0**	25.0	39.5	63.7	**83.2**	44.9	39.8	66.3	**5.2**	**83.2**	68.5	41.7	28.6	66.4	14.8	60.4	**82.2**	70.8
4	1	58.8	32.9	75.4	64.6	10.3	**84.6**	67.1	24.1	**7.8**	30.6	**0.0**	31.4	**88.7**	**87.8**	32.1	40.7	**85.5**	29.5	58.9	55.5	44.7	40.6	59.2	11.8	**82.8**	74.8	55.4
4	2	32.5	64.8	45.4	**84.3**	42.8	52.2	**81.7**	13.8	40.1	61.9	32.5	**0.0**	**90.0**	58.7	62.9	71.2	75.0	40.0	56.1	28.9	74.9	66.4	**89.8**	37.5	**83.3**	64.5	31.8
4	3	**87.7**	**6.8**	75.8	35.9	19.3	70.2	38.5	48.2	21.5	**2.4**	28.9	59.9	**0.0**	67.5	**83.8**	77.1	37.0	59.5	34.7	69.5	74.1	61.4	**83.3**	77.8	24.4	27.4	63.6
4	4	59.6	33.9	41.4	12.3	55.7	30.6	11.7	**87.7**	58.3	37.3	66.0	**81.6**	39.6	**0.0**	66.7	51.6	31.3	**84.3**	58.7	32.4	47.5	55.1	44.3	**84.0**	44.7	46.6	34.1
5	0	72.9	26.9	**87.4**	54.5	16.3	**89.6**	57.1	29.0	15.9	22.8	18.3	42.0	20.4	59.8	**0.0**	20.8	74.2	49.9	63.7	**81.8**	25.0	34.8	27.3	32.5	78.7	**81.4**	79.3
5	1	**89.9**	**8.4**	73.7	34.2	22.1	69.5	36.8	49.4	24.2	**4.3**	31.4	61.5	**3.4**	38.9	20.9	**0.0**	53.6	43.6	**84.3**	**84.0**	**4.6**	17.3	27.5	34.1	58.1	77.9	**85.0**
5	2	33.1	**87.2**	30.5	59.5	72.2	22.9	57.0	39.1	69.6	**89.2**	62.3	30.6	**86.9**	53.5	66.8	**87.7**	**0.0**	65.1	42.2	46.5	49.2	45.4	63.5	74.2	13.4	24.6	43.1
5	3	**86.6**	**3.0**	69.5	29.5	25.0	62.9	32.0	55.4	27.4	**4.9**	35.1	66.7	**7.3**	32.4	27.7	**7.6**	**85.9**	**0.0**	**86.9**	68.9	45.1	30.8	70.6	19.9	58.1	79.3	71.8
5	4	34.9	**89.9**	33.5	62.6	69.9	25.7	60.1	36.6	67.3	**86.5**	60.1	29.0	**84.1**	56.3	63.9	**84.8**	**3.2**	**88.6**	**0.0**	42.3	**88.6**	**85.3**	78.2	70.5	39.6	17.8	36.1
5	5	61.2	68.7	50.0	53.3	**83.9**	25.1	51.7	63.8	**86.2**	73.1	**87.8**	60.0	75.5	41.9	**84.9**	76.2	31.5	68.7	31.8	**0.0**	79.8	**85.5**	73.0	64.5	56.4	42.1	**6.3**
6	0	**86.3**	**6.6**	69.8	30.1	25.2	65.2	32.7	53.6	27.5	**4.8**	35.0	65.6	**6.0**	34.7	25.1	**4.3**	**88.1**	**4.5**	**89.1**	72.4	**0.0**	16.4	27.6	37.5	54.2	73.7	**80.7**
6	1	**81.2**	**8.0**	64.0	23.9	30.5	57.7	26.4	60.8	33.0	10.4	40.7	72.3	12.6	27.1	32.7	11.9	**80.6**	**5.6**	**83.4**	64.9	**7.7**	**0.0**	44.0	30.1	45.8	67.7	**84.6**
6	2	**80.5**	15.2	**83.6**	44.2	14.0	79.1	46.8	39.5	15.6	11.2	22.1	51.5	**8.9**	48.5	11.6	10.0	78.0	16.2	75.2	**83.7**	14.2	21.4	**0.0**	57.4	73.4	**86.6**	76.6
6	3	78.8	22.4	**86.7**	48.6	17.7	**84.9**	51.2	35.0	18.3	18.0	22.6	47.9	15.6	54.4	**6.0**	15.5	72.4	22.7	69.5	**89.0**	19.7	27.4	**7.7**	**0.0**	71.1	**86.5**	65.3
6	4	48.4	52.8	31.5	28.4	73.7	11.6	26.3	73.5	76.3	56.5	**83.8**	63.8	58.9	19.5	79.2	58.4	34.4	51.6	37.0	26.6	54.1	46.5	67.8	73.8	**0.0**	22.5	51.9
6	5	56.9	55.9	65.7	**81.0**	40.0	61.7	**83.6**	14.8	38.0	51.8	33.4	25.6	49.3	**88.1**	29.0	49.4	39.5	56.5	36.4	56.5	53.6	61.3	40.6	33.9	72.4	**0.0**	36.2
6	6	23.3	75.7	13.2	45.1	**82.0**	16.7	42.5	51.5	79.4	78.6	71.7	39.5	**80.7**	42.3	**80.5**	79.2	17.8	73.8	21.0	38.6	75.0	68.2	**89.2**	**86.3**	26.6	55.6	**0.0**

Table 7.2: Mode to mode angular separation. Maximum power inter-mode ($\vec{r}_{\ell m}(\hat{\mathbf{n}}^{\star})$) alignments comparison in the: ILC3-ILC3 data (below diagonal) and ILC5-ILC5 data (above diagonal). The alignments below 10° were highlighted in blue while the orthogonal alignments ($\geq 80^{\circ}$) were highlighted in green.

	l	1	1	2	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6	6	6
l	m	0	1	0	1	2	0	1	2	3	0	1	2	3	4	0	1	2	3	4	5	0	1	2	3	4	5	6
1	0	**0.0**	**89.2**	18.2	57.5	68.5	39.8	55.1	46.3	66.2	**89.2**	58.8	32.5	**87.7**	59.6	72.9	**89.9**	33.1	**86.6**	34.9	61.2	**86.3**	**81.2**	**80.5**	78.8	48.4	56.9	23.3
1	1	**85.4**	**0.0**	72.0	31.9	22.6	64.3	34.3	53.8	25.1	**4.4**	32.9	64.8	**6.8**	33.9	26.9	**8.4**	**87.2**	**3.0**	**89.9**	68.7	**6.6**	**8.0**	15.2	22.4	52.8	55.9	75.7
2	0	50.8	35.5	**0.0**	40.1	**85.5**	25.4	37.6	58.8	**83.0**	74.0	75.4	45.4	75.8	41.4	**87.4**	73.7	30.5	69.5	33.5	50.0	69.8	64.0	**83.6**	**86.7**	31.5	65.7	13.2
2	1	18.7	67.6	32.4	**0.0**	54.4	37.6	**2.6**	**83.4**	56.9	34.0	64.6	**84.3**	35.9	12.3	54.5	34.2	59.5	29.5	62.6	53.3	30.1	23.9	44.2	48.6	28.4	**81.0**	45.1
2	2	46.0	48.6	**83.9**	63.8	**0.0**	**85.3**	56.9	33.4	**2.6**	20.7	10.3	42.8	19.3	55.7	16.3	22.1	72.2	25.0	69.9	**83.9**	25.2	30.5	14.0	17.7	73.7	40.0	**82.0**
3	0	70.6	24.0	59.2	**88.5**	24.7	**0.0**	35.2	61.9	**87.9**	67.9	**84.6**	52.2	70.2	30.6	**89.6**	69.5	22.9	62.9	25.7	25.1	65.2	57.7	79.1	**84.9**	11.6	61.7	16.7
3	1	73.5	13.1	22.9	55.3	61.0	36.3	**0.0**	**86.0**	59.4	36.6	67.1	**81.7**	38.5	11.7	57.1	36.8	57.0	32.0	60.1	51.7	32.7	26.4	46.8	51.2	26.3	**83.6**	42.5
3	2	29.8	74.8	70.3	40.9	31.0	53.1	**87.9**	**0.0**	30.9	50.4	24.1	13.8	48.2	**87.7**	29.0	49.4	39.1	55.4	36.6	63.8	53.6	60.8	39.5	35.0	73.5	14.8	51.5
3	3	52.5	42.2	77.5	70.2	**6.5**	18.3	54.6	36.1	**0.0**	23.1	**7.8**	40.1	21.5	58.3	15.9	24.2	69.6	27.4	67.3	**86.2**	27.5	33.0	15.6	18.3	76.3	38.0	79.4
4	0	22.4	**89.8**	54.9	27.1	45.0	68.2	76.7	15.4	50.6	**0.0**	30.6	61.9	**2.4**	37.3	22.8	**4.3**	**89.2**	**4.9**	**86.5**	73.1	**4.8**	10.4	11.2	18.0	56.5	51.8	78.6
4	1	12.8	**85.1**	59.4	27.4	36.8	61.4	**82.3**	17.1	43.1	14.4	**0.0**	32.5	28.9	66.0	18.3	31.4	62.3	35.1	60.1	**87.8**	35.0	40.7	22.1	22.6	**83.8**	33.4	71.7
4	2	43.0	52.7	**88.2**	59.8	**7.1**	29.5	65.5	24.4	11.7	38.9	32.4	**0.0**	59.9	**81.6**	42.0	61.5	30.6	66.7	29.0	60.0	65.6	72.3	51.5	47.9	63.8	25.6	39.5
4	3	50.6	45.6	18.4	32.7	**89.4**	66.1	32.9	60.8	**83.6**	45.8	55.0	**84.7**	**0.0**	39.6	20.4	**3.4**	**86.9**	**7.3**	**84.1**	75.5	**6.0**	12.6	**8.9**	15.6	58.9	49.3	**80.7**
4	4	**84.2**	17.3	34.4	65.5	53.0	30.2	15.7	**83.1**	47.3	**82.1**	**89.2**	58.9	36.4	**0.0**	59.8	38.9	53.5	32.4	56.3	41.9	34.7	27.1	48.5	54.4	19.5	**88.1**	42.3
5	0	35.2	**81.0**	65.8	40.7	41.6	61.5	**86.0**	11.8	45.9	14.0	24.2	34.6	53.4	**88.3**	**0.0**	20.9	66.8	27.7	63.9	**84.9**	25.1	32.7	11.6	**6.0**	79.2	29.0	**80.5**
5	1	41.4	**87.7**	57.6	39.8	55.9	74.7	75.2	25.7	60.0	19.6	33.7	48.8	42.4	75.4	14.3	**0.0**	**87.7**	**7.6**	**84.8**	76.2	**4.3**	11.9	10.0	15.5	58.4	49.4	79.2
5	2	60.6	34.7	70.3	77.8	15.0	12.3	47.6	40.9	**8.9**	56.1	50.4	18.0	78.3	42.5	49.2	62.5	**0.0**	**85.9**	**3.2**	31.5	**88.1**	**80.6**	78.0	72.4	34.4	39.5	17.8
5	3	68.8	25.8	61.0	**86.7**	22.9	**1.9**	38.1	51.2	16.5	66.3	59.6	27.6	68.0	32.0	59.7	73.0	10.5	**0.0**	**88.6**	68.7	**4.5**	**5.6**	16.2	22.7	51.6	56.5	73.8
5	4	47.9	57.1	30.1	32.9	**82.2**	75.9	44.7	51.5	**87.5**	37.5	49.0	75.8	12.4	45.7	42.7	30.6	**88.2**	77.7	**0.0**	31.8	**89.1**	**83.4**	75.2	69.5	37.0	36.4	21.0
5	5	79.3	49.5	46.1	63.6	72.7	57.1	43.1	77.0	69.0	65.9	79.2	79.7	32.4	32.3	65.8	51.5	68.1	58.5	31.4	**0.0**	72.4	64.9	**83.7**	**89.0**	26.6	56.5	38.6
6	0	77.5	18.8	51.8	**84.1**	32.8	**9.9**	29.1	62.5	26.8	77.3	69.5	38.4	56.9	20.6	71.3	**84.6**	22.1	11.6	66.2	47.6	**0.0**	**7.7**	14.2	19.7	54.1	53.6	75.0
6	1	71.3	23.3	58.4	**89.2**	25.4	**0.7**	35.5	53.8	19.1	68.9	62.2	30.2	65.4	29.5	62.1	75.3	13.0	**2.6**	75.2	56.6	**9.3**	**0.0**	21.4	27.4	46.5	61.3	68.2
6	2	75.1	21.9	26.2	56.4	62.4	39.2	13.7	**87.7**	56.6	72.8	**81.4**	68.2	27.1	**9.4**	79.5	67.1	51.5	41.0	36.9	29.4	29.8	38.5	**0.0**	**7.7**	67.8	40.6	**89.2**
6	3	**86.2**	20.7	37.1	67.6	52.3	30.6	19.6	**82.9**	46.8	**82.7**	**87.9**	58.5	37.3	**4.0**	**88.0**	74.7	42.8	32.4	45.7	29.4	20.7	30.0	11.2	**0.0**	73.8	33.9	**86.3**
6	4	38.6	71.9	74.6	47.6	33.7	52.4	**84.8**	**9.4**	37.4	20.7	26.2	26.6	62.6	**82.5**	**9.2**	22.6	40.1	50.6	51.9	73.7	62.2	53.0	**88.8**	**83.0**	**0.0**	72.4	26.6
6	5	40.6	76.7	71.0	46.6	40.5	58.4	**89.5**	13.9	44.1	19.8	29.0	33.4	57.7	**88.3**	**5.9**	16.3	46.3	56.7	46.4	67.0	68.3	59.0	**83.2**	**89.0**	**6.8**	**0.0**	55.6
6	6	29.5	66.0	31.7	15.8	67.9	**88.2**	53.0	39.2	73.9	23.8	32.2	62.4	22.8	59.1	34.2	28.0	79.8	**89.9**	18.4	49.8	79.3	**87.5**	49.7	60.1	42.9	39.5	**0.0**

Table 7.3: Mode to mode angular separation. Maximum power inter-mode ($\vec{r}_{\ell m}(\hat{\mathbf{n}}^{\star})$) alignments comparison in the: ILC5-ILC5 data (below diagonal) and V5-V5 data (above diagonal). The alignments below 10° were highlighted in blue while the orthogonal alignments ($\geq 80^{\circ}$) were highlighted in green.

	l	1	1	2	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6	6	6
l	m	0	1	0	1	2	0	1	2	3	0	1	2	3	4	0	1	2	3	4	5	0	1	2	3	4	5	6
1	0	**0.0**	55.9	**80.9**	62.9	**4.2**	42.8	10.9	23.8	19.1	28.4	15.5	36.9	22.2	60.4	21.0	27.6	76.0	32.7	64.2	78.6	45.1	37.0	23.4	65.4	**82.1**	39.3	**81.0**
1	1	**82.6**	**0.0**	45.6	**7.0**	51.7	**89.5**	57.1	79.0	46.3	28.4	59.4	**89.0**	34.2	15.1	48.6	29.2	24.6	23.3	64.5	58.9	12.0	18.9	36.4	17.4	29.5	**81.7**	57.5
2	0	23.2	59.7	**0.0**	39.0	**84.8**	44.2	77.7	57.1	**88.1**	73.6	75.1	44.2	79.1	38.8	**85.9**	74.4	23.1	68.1	19.1	23.8	54.5	63.6	**81.9**	33.7	17.0	53.7	24.5
2	1	66.8	16.2	43.7	**0.0**	58.7	**83.1**	64.0	**86.0**	53.0	35.3	66.2	**82.1**	41.2	14.8	55.1	36.1	19.4	30.3	58.0	54.0	18.5	25.9	43.2	14.9	23.5	**87.8**	52.8
2	2	67.0	30.8	**89.6**	46.9	**0.0**	45.7	11.4	27.8	16.4	24.2	16.1	40.7	18.0	56.5	18.9	23.4	72.1	28.4	67.9	**81.6**	41.0	32.8	19.5	61.6	78.2	41.3	**83.9**
3	0	42.5	57.2	27.9	42.9	**85.1**	**0.0**	35.9	23.4	44.5	63.3	32.3	14.9	58.5	**80.0**	41.9	62.6	65.5	69.1	25.1	35.9	**83.0**	73.5	54.6	75.2	60.1	10.3	38.2
3	1	57.6	25.3	34.5	**9.2**	56.1	36.3	**0.0**	22.7	12.6	28.8	**4.7**	33.9	23.2	64.5	12.8	28.0	79.6	34.2	59.6	71.5	47.9	38.7	21.0	69.4	**85.5**	30.4	73.7
3	2	48.9	63.1	64.8	78.0	37.3	59.7	**85.9**	**0.0**	35.1	50.7	22.3	13.3	44.8	**84.2**	34.5	49.9	**80.2**	55.7	40.9	57.1	68.8	60.1	43.6	**89.2**	74.1	25.6	59.6
3	3	61.0	36.5	**83.8**	52.5	**6.6**	**88.9**	61.8	35.9	**0.0**	18.9	13.4	45.5	15.1	55.8	**3.6**	18.3	70.0	24.8	69.4	78.3	38.6	29.1	10.1	60.4	75.5	36.7	**80.4**
4	0	**82.9**	15.6	74.7	31.6	15.9	72.4	40.7	48.0	22.2	**0.0**	31.4	62.6	**6.2**	36.8	21.9	**0.8**	51.3	**5.9**	**88.0**	**82.8**	19.7	10.2	**8.9**	41.5	56.9	55.6	**80.9**
4	1	52.7	45.6	74.9	61.8	14.9	**80.7**	71.0	26.7	10.4	30.4	**0.0**	32.1	26.3	67.8	12.3	30.6	**82.6**	37.0	56.6	67.5	50.9	41.5	23.0	72.6	**88.3**	26.0	69.7
4	2	65.3	33.3	**87.5**	49.4	**4.0**	**88.6**	58.5	33.4	**7.5**	17.9	12.7	**0.0**	56.9	**82.8**	44.0	61.8	67.2	67.9	27.6	44.4	**81.3**	72.4	54.9	77.8	61.2	22.2	46.9
4	3	32.0	60.0	17.3	44.3	**89.8**	11.2	36.0	58.1	**85.5**	75.6	75.4	**87.1**	**0.0**	41.5	18.5	**5.4**	56.4	11.0	**82.8**	**86.8**	24.7	15.5	**6.7**	46.3	62.2	51.4	**84.7**
4	4	56.8	31.8	34.2	17.6	61.0	25.6	13.3	**85.2**	65.7	47.2	75.8	64.2	28.9	**0.0**	58.7	37.5	15.7	31.0	55.9	46.3	17.2	26.7	45.7	**5.1**	21.8	**87.6**	44.5
5	0	72.2	33.0	**87.5**	47.8	16.4	**89.8**	56.0	30.2	21.1	18.7	22.2	13.6	**88.1**	64.6	**0.0**	21.4	72.6	27.8	67.0	75.2	41.5	32.0	13.1	63.2	78.0	33.7	77.2
5	1	**86.3**	13.2	71.4	28.8	19.4	70.3	37.7	49.9	25.8	**3.6**	33.7	21.1	73.0	44.9	20.0	**0.0**	52.0	**6.5**	**87.2**	**83.4**	20.4	10.9	**8.3**	42.2	57.7	55.0	**81.4**
5	2	51.1	40.2	29.9	25.9	69.1	17.1	20.1	76.7	73.5	55.6	**83.7**	72.3	21.2	**8.5**	73.1	53.4	**0.0**	45.6	40.8	34.6	31.8	41.1	59.9	10.6	**6.2**	73.9	33.4
5	3	**89.0**	**8.5**	68.2	24.6	22.3	64.5	33.8	55.9	28.0	**7.8**	37.2	25.0	68.0	39.5	26.4	**6.8**	47.8	**0.0**	**86.4**	77.0	13.9	**4.5**	14.8	35.7	51.3	61.5	75.0
5	4	66.0	18.9	42.8	**5.5**	49.0	38.6	**9.6**	**81.8**	54.2	34.5	63.9	51.9	41.1	13.1	51.6	32.0	21.5	26.9	**0.0**	19.6	72.6	**81.9**	79.4	51.0	35.2	34.6	21.9
5	5	62.7	67.7	53.3	58.4	**87.1**	25.7	55.3	55.9	**88.2**	79.7	**82.2**	**88.8**	36.0	42.1	**81.7**	79.2	35.8	72.4	53.0	**0.0**	63.5	72.9	**88.3**	42.2	31.2	41.8	**2.5**
6	0	45.5	**85.0**	46.6	71.7	70.5	28.9	65.3	33.3	68.5	**80.8**	58.5	66.6	32.2	54.1	62.2	**82.2**	45.8	**88.5**	67.0	24.3	**0.0**	**9.5**	28.5	21.8	37.6	75.2	61.6
6	1	33.4	73.5	50.4	**89.4**	44.1	50.9	**82.0**	15.5	40.3	57.9	29.9	40.8	46.5	75.1	42.4	60.5	66.9	65.5	**87.6**	55.0	30.6	**0.0**	19.0	31.3	46.8	65.7	71.0
6	2	50.1	41.9	29.3	27.6	70.7	15.4	21.7	75.0	75.0	57.2	**85.2**	74.0	19.7	10.2	74.8	55.1	**1.7**	49.5	23.2	34.6	44.1	65.3	**0.0**	50.3	65.4	46.7	**89.8**
6	3	73.9	28.6	**85.0**	43.9	12.7	**85.8**	52.4	34.5	18.4	13.7	22.0	11.0	**88.2**	60.4	**5.2**	15.4	68.8	21.5	47.4	**86.9**	67.0	45.6	70.5	**0.0**	16.8	**83.1**	40.6
6	4	52.7	54.4	71.1	69.7	28.1	68.5	78.0	**9.2**	27.1	39.1	18.7	24.1	66.1	**86.2**	22.0	41.2	**85.4**	46.9	73.2	64.9	42.5	20.7	**83.7**	25.8	**0.0**	68.8	30.4
6	5	37.0	65.5	27.5	50.8	**86.3**	**8.6**	43.6	51.4	**82.5**	**80.8**	72.1	**82.8**	10.7	33.8	**81.6**	78.7	25.3	73.0	46.8	26.7	22.2	42.2	23.6	**85.9**	60.1	**0.0**	43.9
6	6	34.1	64.4	23.6	49.2	**86.6**	**9.2**	41.5	52.8	**82.5**	79.9	72.2	**83.3**	**7.0**	32.7	**83.0**	77.6	24.4	72.2	45.5	30.4	25.3	42.5	22.7	**87.1**	61.2	**3.9**	**0.0**

Table 7.4: Mode to mode angular separation. Maximum power inter-mode ($\vec{r}_{\ell m}(\hat{\mathbf{n}}^{\star})$) alignments comparison in the: TOH5-TOH5 data (below diagonal) and HILC5-HILC5 data (above diagonal). The alignments below 10° were highlighted in blue while the orthogonal alignments ($\geq 80^{\circ}$) were highlighted in green.

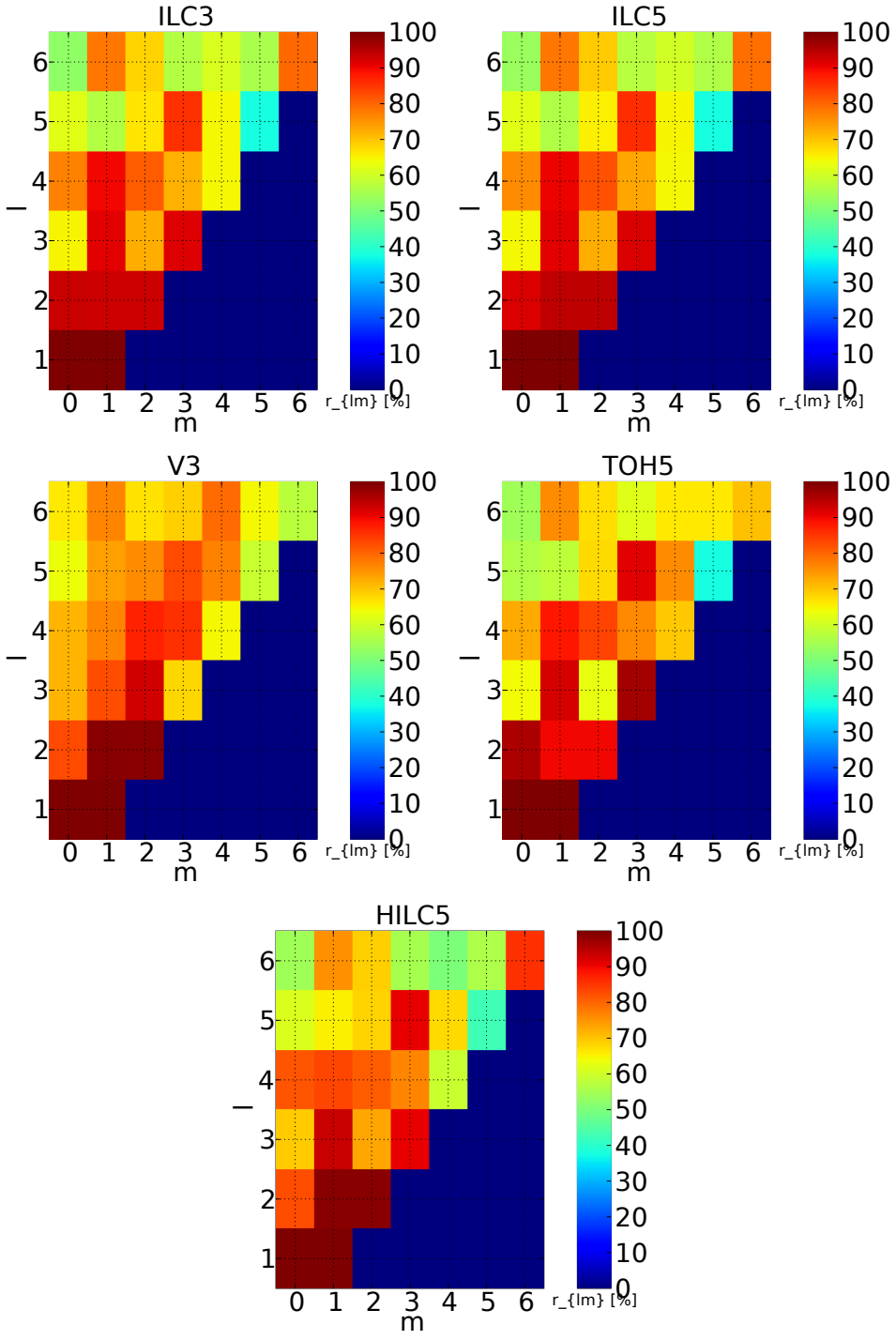


Figure 7.8: Results of preferred reference frame search. In each panel, for each of the considered data we plot the distribution of power accumulated in individual modes as described in table 7.1.

Chapter 8

Tests of the Poincaré dodecahedral space topology hypothesis

The matherial presented in this chapter was published in Astronomy and Astrophysics (Lew & Roukema, 2008) as a follow-up paper to the Roukema et al. (2004).

8.1 Abstract

It has been suggested by Roukema and coworkers (hereafter R04) that the topology of the Universe as probed by the "matched circles" method using the first year release of the WMAP CMB data, might be that of the Poincaré dodecahedral space (PDS) model. An excess in the correlation of the "identified circles" was reported by R04, for circles of angular radius of $\sim 11^\circ$ for a relative phase twist -36°, hinting that this could be due to a Clifford translation, if the hypothesized model were true. R04 did not however specify the statistical significance of the correlation signal.

We investigate the statistical significance of the signal using Monte Carlo CMB simulations in a simply connected Universe, and present an updated analysis using the three-year WMAP data. We find that our analyses of the first and three year WMAP data provide results that are consistent with the simply connected space at a confidence level as low as 68%.

8.2 Introduction

If the topology of the Universe were multiply connected, as opposed to simply-connected, and if the comoving size of the fundamental domain (FD) were smaller than the comoving distance to the surface-of-last-scattering (SLS), then it should be possible to detect repeating patterns in the CMB fluctuations using full-sky data of sufficient signal-to-noise ratio. These fluctuations would be those lying along pairs of circles defined by points of intersection between different copies of the SLS in the covering space (Cornish et al., 1998b).

These patterns although found in different directions of the sky, would constitute so-called "matched circles", as they would represent the same physical points, but observed from different directions due to topological lensing.
While this principle is true for any 3-manifold model of space, the number of pairs of "matched circles" or their sizes and relative spatial orientations, as well as their handedness, or phase shift, depend significantly on the assumed 3-manifold and its topological properties, thus providing a way to observationally distinguish between models.

While the positive correlation signal from matched pairs is expected directly from the metric perturbations, via the Sachs-Wolfe Effect (Sachs & Wolfe, 1967), there are many other cosmological effects (e.g. the Doppler effect, the Integrated Sachs-Wolfe Effect (ISW)) (Kunz et al., 2008), astrophysical foregrounds (Bennett et al., 2003b) and instrumental effects that constitute noise, from the point-of-view of a matched circles search, and the magnitude of the effects depends on the angular scale.

Although the CMB data have been analyzed to detect topological lensing signals since the availability of the COBE data (Roukema, 2000), the release of the WMAP observations has provided full-sky data of unprecedented accuracy and resolution, opening up more promise for direct tests of the topology of the Universe. Although the "matched circles" test is straightforward, it is limited due to noise and FD size constraints. Additional theoretical predictions can be used as independent tests that involve predictions of CMB temperature and polarization fluctuations for the case that the Universe is multiply connected, both in real and spherical harmonic spaces, or topological effects on the CMB power spectrum (Aurich et al., 2004; Weeks, 2003; Gomero & Rebouças, 2003; Dineen et al., 2005; Riazuelo et al., 2004a; Lehoucq et al., 1999; Inoue, 1999; Niarchou & Jaffe, 2007; Cornish et al., 1998a; de Oliveira-Costa & Smoot, 1995; Kunz et al., 2006; Luminet et al., 2003; Niarchou & Jaffe, 2006; Phillips & Kogut, 2006; Caillerie et al., 2007). Although a successful "matched circles" test would provide strong support for the Universe being multiply connected, no statistically-significant evidence has been found (Cornish et al., 2004; Key et al., 2007).

In Roukema et al. (2004) we performed a "matched circles" search using the first year WMAP ILC map (Bennett et al., 2003b) and found an excess correlation, which one would expect under the PDS hypothesis for circles of angular radii $\alpha \sim 11^\circ$ with centers towards $(l^\circ, b^\circ) = \{(252, 65), (51, 51), (144, 38), (207, 10), (271, 3), (332, 25)\}$ and their opposites (Fig. 8.1).

In this present work we have two key objectives. We revisit those results, verify the existence of the excess correlations and quantify their statistical significance. Secondly, we update the search with the WMAP three year data release, extend it to probe three different resolutions (smoothing lengths) and define the detection confidence thresholds. We also discuss the effects of underlying 2-point correlations, smoothing length and incomplete sky coverage on the value of correlation coefficient.

In section 8.3 we introduce the datasets used in the analysis, provide details of their preprocessing, and describe simulations that we use to complete a statistical significance

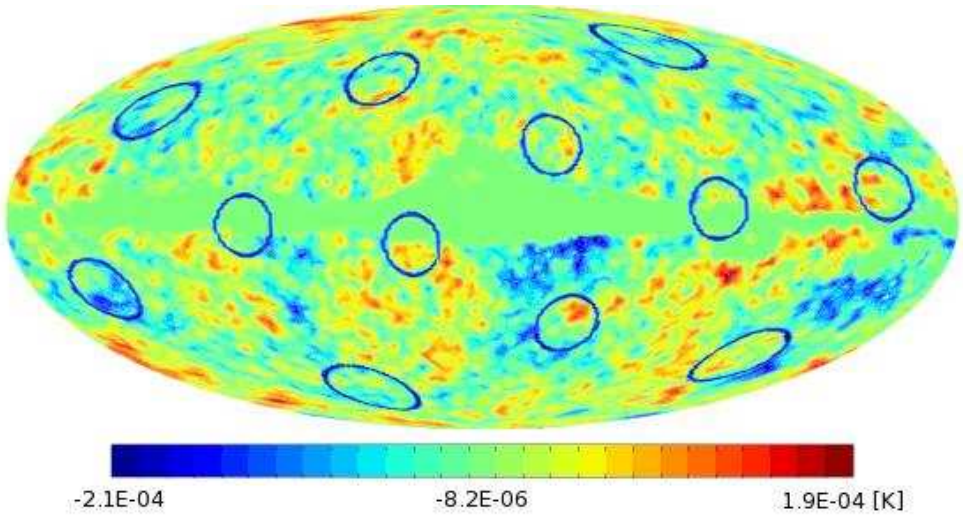


Fig 8.1: Visualization of the matched circles solution reported in Roukema et al. (2004) and reproduced to constrain its statistical significance, over-plotted on the first year ILC map masked with Kp2 sky mask.

analysis. In section 8.4 we introduce the details of the statistics being performed and our confidence-level analysis. Results are presented in section 8.5. We conclude in section 8.8.

8.3 Data and simulations

We perform a "matched circles" search using two sets of data. Firstly, for the sake of compatibility, we choose the same data as in Roukema et al. (2004) – i.e. the first year WMAP ILC map.

The topologically interesting signal generally dominates over the Doppler (and other) components on large scales (Riazuelo et al., 2004a): this is a motivation for using a large smoothing length. However, extended flat fluctuations that happen to have a similar large scale trend can lead to false positives on large scales (Key et al., 2007). This implies a trade off in the choice of the smoothing length of the data, between large smoothing lengths preferred by the topologically interesting content, and small smoothing scales which avoid false positives induced by chance correlations of extended flat fluctuations.

We choose to test three different smoothing scales: $FWHM \equiv \lambda \in \{1^\circ, 2^\circ, 4^\circ\}$.

The ILC map was obtained from a linear combination of one degree smoothed maps in the five frequency bands, by inverse noise co-adding them, hence its resolution is consistent with a one degree smoothing scale.

We further Gaussian smooth this map in spherical harmonic space by convolving it with Gaussian beam response kernels of FWHM corresponding to 2° and 4° respectively, to obtain the first set of data for matched circles tests.

Secondly, we choose the three year foreground reduced WMAP data from individual frequency bands Q[1/2],V[1/2] and W[1/2/3/4] and co-add them into one map, according

to the inverse noise weighting scheme used in Hinshaw et al. (2003b). We call the resultant map the "INC map". We smooth the INC map using a Gaussian convolution kernel to four different FWHM smoothing lengths: $\lambda \in \{0.5^\circ, 1^\circ, 2^\circ, 4^\circ\}$.

We downgrade all data from the initial resolution defined by the Healpix pixelization scheme (Górski et al., 2005) with resolution parameter $n_s = 512$ (res. 9) to a resolution parameter of $n_s = 256$ (res. 8). We remove the the residual monopole and dipole components ($\ell = 0, 1$) in spherical harmonic space[1], because these components are of no cosmological interest. At the final stage of preprocessing, we remove the residual monopole by offsetting the maps in real (2-sphere) space so that $\langle T \rangle = 0$ outside the Kp2 sky mask.

Throughout the analysis, i.e. for both the ILC and INC maps, we use the Kp2 sky mask, which masks $\sim 15\%$ of the sky including the brightest resolved point sources. The Kp2 sky mask is different from the sky mask used in Roukema et al. (2004). While the ILC map is best suited e.g. for the full sky low multipoles alignment analysis, for the purpose of the matched circles test, the residual galactic contamination should be masked out, although we realize that the use of the Kp2 mask may be too conservative. In App. 8.6.2 we compare the impact of different sky masks and demonstrate that our results are not very sensitive to the precise characteristics of the sky mask.

For each of the two data sets, we produce $N_{\rm sim} = 100$ realistic Gaussian random field (GRF) signal and noise simulations of the WMAP data to quantify the statistical significance of plausible detections, to discard false positives, and to resolve the 2σ-confidence levels.

Therefore, for the first dataset we simulate the first year ILC map, inside "region 0" defined outside the Kp2 sky mask of Bennett et al. (2003b), and for the second dataset we simulate the three year WMAP INC map.

As will be shown in Sect. 8.4, the matched circles correlation coefficient depends on the monopole value in the map. Also, in principle it is sensitive to the shape of the two-point correlation function, since the correlator is a two-point statistic, by construction, and so it becomes a measure of the underlying intrinsic two-point correlations in the CMB (albeit via a specially selected subset of pairs of points on the matched circles). Therefore it is necessary to take into account possible variations in the underlying two-point correlation function with varying angular separation, which if not properly accounted for in simulations may lead to under(over)-estimation of the confidence level thresholds.

Given that the concordance best fit LCDM cosmological model (Spergel et al., 2007) yields a very poor fit to the CMB data at large angular scales, due to lack of correlations in the two-point correlation function of the data with respect to the LCDM model at scales $> 60^\circ$, and that the correlation statistic is sensitive to the details of the intrinsic two-point CMB correlations (and in particular to any large scale anomalies), we do not assume the LCDM model to help create our simulations. Instead, we take a model independent approach. As the CMB reference power spectrum in our GRF simulations of the expected signal, we use the reconstructed power spectrum from the three-year WMAP data (Hinshaw et al., 2007)

[1]Since the residual WMAP maps foregrounds are strong, we perform this step using the Kp2 sky mask to keep the compatibility between the data and simulations.

[2]. Furthermore, we neglect the effects of cosmic variance, and only randomize the phases (and noise realizations) in our simulations. We remove the $C_{\ell=0,1}$ (i.e. the monopole and dipole) components from our simulations.
We use the same set of a_{lm}s representing the CMB signal for a single simulation of the two datasets, followed by convolution with instrumental beam profiles.
For each differential assembly (DA), we simulate the noise according its properties and scanning strategy (number of observations per pixel in map) using uncorrelated, Gaussian noise.

The simulations are preprocessed in exactly the same way as the observational data. We neglect the impact of the (resolved or unresolved) point sources which is negligible, since we apply relatively large smoothing and use the Kp2 sky mask for the analysis.

In Section 8.6, we discuss the sensitivity of our results to the degree of smoothing, the sky mask applied, and the assumed statistical approach in greater detail.

8.4 Statistics

We describe our correlator statistics, parameter space, search optimization and approach for assessing the statistical significance.

8.4.1 Matched circles test

As in Cornish et al. (2004) and Roukema et al. (2004) we use a correlation statistic of the form

$$S = 2\frac{\langle T_i m_i T_j m_j \rangle}{\langle T_i^2 m_i m_j \rangle + \langle T_j^2 m_i m_j \rangle} \tag{8.1}$$

where the index i defines a set of all points in the "first" set of six circles related to the orientation of a fundamental dodecahedron; index j is the set of corresponding points along the matched six circles; and m_i, m_j are cut sky weights of the Kp2 sky mask, which can have a value of either 0 for a masked pixel or 1 for an unmasked pixel. Clearly, perfectly matched circles would yield $S = 1$, which, due to non-zero noise contributions, is not possible in reality.

The dispersion of the correlation coefficient as defined in Eq. 8.1 is statistically enhanced in the small circles regime, due to the joint effect of the reduced number of points probing the matched circles as compared to larger circles, the accidental correlations of large (w.r.t. the smoothing scale) flat fluctuations that happen to have similar (or opposite) large scale trends, as well as due to the fact that the r.m.s. values necessarily shrink (down to zero in case of zero mean fluctuations) for circles of size comparable or smaller than the smoothing length.

[2] In the high l end (noise dominated range) of the reconstructed power spectrum, the unphysical negative values are zeroed to have a zero contribution to the total variance of the map. This approximation has a negligible effect due to small statistical weight of the large l multipoles, and large exponential Gaussian smoothing that we apply to the data. In practice, this approximation has a negligible effect on the variance of the resulting simulation. Moreover, it can at most only make our analysis more conservative.

As shown in Sect. 8.5, this reduces the ability to robustly determine the degree of consistency or inconsistency of data with simulations, due to the finite accuracy of the S values and significant steepening of confidence-level contours in this regime.

We note that the S statistic value would tend to unity, regardless of the shape of the underlying CMB fluctuations, as the monopole increases in the CMB maps. One could expect a similar effect for the dipole component, for small circles. This effect would affect the simulations and the data to the same extent. The sensitivity of the test would however be significantly weakened and as such we remove the monopole and dipole components from the datasets for the analysis, and defer study of the impact of other small ℓ multipoles to section 8.6.

8.4.2 Parameter space

We perform a resolution-limited, full parameter-space search over the orientation of the fundamental dodecahedron, and over a limited range of the identified circles sizes of up to 20°. The parameters are defined as follows: l, b – galactic longitude and latitude of the first circle, g – the angle of rotation of the dodecahedron about the axis determined by (l, b), a – the angular radius of the matched circle, and s – the twist parameter defining the relative phase offset of the matched circles.
We use the following parameter space:

$$\begin{aligned} l &\in [0^\circ, 72^\circ) \\ b &\in [\,26.57^\circ, 90^\circ) \\ g &\in [0^\circ, 72^\circ) \\ a &\in [1^\circ, 20^\circ] \\ s &\in \{-36^\circ, 0^\circ, 36^\circ\} \end{aligned} \tag{8.2}$$

The boundaries in (l, b) conservatively cover a larger region than the one twelfth of the sphere from which a "first" circle centre can be chosen non-redundantly. The range of angle g is 72°, to cover all possible orientations of the fundamental dodecahedron for a chosen "first" circle centre. Values larger than 72° would yield the same set of 12 circle centres as a rotation by that angle modulo 72°. The interval in circle size a is chosen to be roughly symmetric and centered about the 11° value suggested by Roukema et al. (2004). The three twists s are chosen as in Roukema et al. (2004).

For all datasets we use the same resolution of 1° in probing the parameter space, except for the data with smoothing length $\lambda = 0.5^\circ$, in which case we use a resolution of 0.5°.

8.4.3 Accuracy and search optimization

The resolution of the data that we analyze is spatially constant and is limited by the finite pixel size, so circles of different sizes are probed by different numbers of pixels.
As the parameter space of the search is large, it is important to consider the trade-off between the accuracy of the estimates of S (directly related to the number of pixels probing

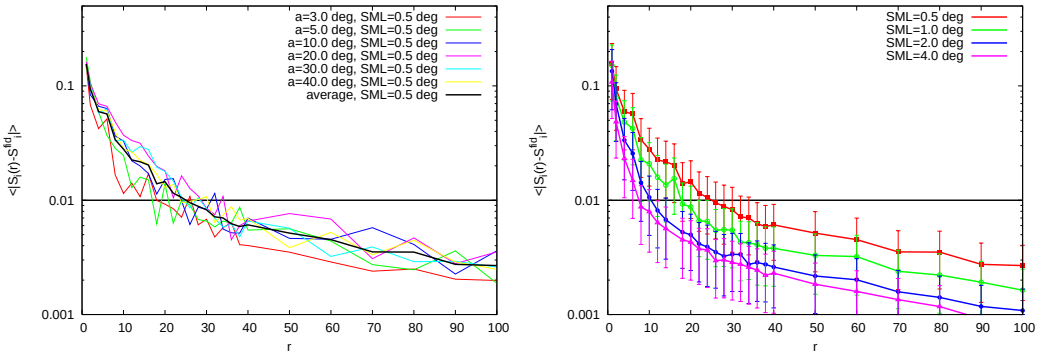


Fig 8.2: Left panel: Convergence of S values to the "ideal" fiducial value $S^{\rm fid} = S(r = 1000)$ as a function of resolution parameter r (a sampling density resolution parameter, defining the number of pixels to be used, to probe the CMB fluctuations along circles through Eq. (8.3)) and as a function of circle size a. The $N_{pix}(a)$ function shape for a given smoothing length is fitted linearly, so that the accuracy in S values was approximately constant for all considered circles radii. The assumed working precision level of $\Delta S = 0.01$ is marked with thick horizontal line. For clarity, only ΔS relation derived for data smoothed with Gaussian $\lambda = 0.5^\circ$ is shown. Similar relations are obtained for the remaining three smoothing lengths. The average value (black thick line) from all tested circles radii is used to define the required value of r parameter for the circle search with data smoothed to 0.5°, in order to achieve the targeted accuracy on S value. Right panel: Average S convergence relations derived for data with different smoothing lengths $\lambda \in \{0.5^\circ, 1^\circ, 2^\circ, 4^\circ\}$, along with 1σ error bars from 20 simulations. The intersections of these with the horizontal (black thick) line give the required values of r for each smoothing length in order to obtain the assumed working precision of $\Delta S = 0.01$.

the underlying fluctuations) and the numerical computational time needed to obtain better accuracy. However, the speed of the search can be substantially increased, since the effective resolution of the data in our case is not limited by the pixel size, but rather by the smoothing length.

In this section, we focus on the density of points (probing the fluctuations along circles in the sky) required to obtain a given accuracy in estimating S, and its dependence on the angular radius of the circles a, a circle sampling density parameter r, the map resolution parameter $n_{\rm S}$, and the smoothing length properties of the data.

Maps smoothed with larger smoothing lengths have fewer significant, high-spatial frequency Fourier modes, and there is no need for fine sampling in order to fully encode the information content along the circles. Assessing the same level of precision for smaller circles also requires a smaller number of pixels than for larger circles.

We perform a series of tests to determine the sampling density required to achieve our desired accuracy level. The tests rely on measuring the speed of convergence to the "ideal" fiducial $S^{\rm fid}$ value, derived using far more points in the circle than the number of available

pixels along the circle in our datasets[3], as a function of the increasing sampling density.

We empirically model the circle sampling-density function in such a way that for a given r value parameter, and for a given smoothing length of the data, the accuracy of the resulting S values (i.e. the statistical size of the departure from the fiducial value) is approximately the same for all circle sizes (Fig. 8.2 left panel). We use the following fitted function:

$$N_{pix} = \Big(3.40a[deg] + 76.85\Big)\Big(\frac{r}{32}\Big)\Big(\frac{n_s}{256}\Big) \tag{8.3}$$

where N_{pix} is the number of pixels used for calculation of S, for a circle of angular size a, and for a map of resolution $n_{\rm S}$. The resolution parameter r controls the sampling density. In practice, we choose the closest, even integer as an N_{pix} value for the calculations. This empirically-devised formula yields approximately the same accuracy of derived values of S for all circle sizes (Fig. 8.2 left panel), and holds for all smoothing lengths. The aim is to find a value of r, for each smoothing length, which will provide sufficient accuracy.

We therefore calculate the deviation $\Delta S(r)$

$$\Delta S(r) = \langle |S_i(r) - S_i^{\rm fid}(r = 1000)| \rangle \tag{8.4}$$

where the $\langle\rangle$ averaging is performed over all curves derived from Eq. 8.3 for circle radii $a[deg] \in \{3, 5, 10, 20, 30, 40\}$.

We assume the working accuracy for S values to be $\Delta S = 0.01$ throughout the analysis. This defines the required values of the sampling density parameter r (Fig. 8.2 right panel) and the corresponding number of pixels to be used (Eq. 8.3) to achieve the targeted accuracy. For the smoothing lengths $\lambda[deg] \in \{0.5, 1, 2, 4\}$, the required resolution parameter values are $r \in \{26, 18, 12, 8\}$. We use these values throughout the analysis with both the data and the simulations.

8.4.4 Statistical significance

In this section we discuss our statistical approach for quantifying the confidence intervals.

Since our simulations simulate CMB fluctuations in an isotropic, simply-connected Universe, we test the consistency of the WMAP data with the null hypothesis that the CMB is an arbitrary realization of the GRF in a simply-connected space. We quantify the degree of consistency via S correlator values obtained from the data, and compared with those of simulated distributions from $N_{\rm sim} = 100$ GRF simulations (Sect. 8.3). As an alternate hypothesis we choose the PDS topological model. The inconsistency of the data with the simulations, at high significance level, would then be considered as consistency in favor of the alternative hypothesis (PDS model).

Since we are interested only in the highest positive S correlations, we build probability distribution functions (PDFs) of $S^{\rm max}(a)$, the maximal value of the correlation $S(a)$ found in the matched circle search in the parameter space (l, b, g, s) (Eq. 8.2), using $N_{\rm sim} = 100$

[3] For all directions pointing inside a single pixel, the same temperature value of that pixel is used.

simulations[4]. We probe the underlying PDFs of $S^{\max}(a)$ at 8 different values of a, i.e. for $a \in \{1, 2, 5, 8, 11, 14, 17, 20\}$ in degrees.

We reconstruct the confidence intervals $[c(a), d(a)]$, for the 68% and 95% confidence levels defined by the (cumulative) probability P of finding a GRF, simulated, CMB realization that yields $S^{\max}_{\rm sim} > S^{\max}_{\rm data}$:

$$\begin{aligned} P(S^{\max}_{\rm sim} > S^{\max}_{\rm data})(a) &= 1 - \int_{c(a)}^{d(a)} f(S^{\max}, a) dS^{\max} \\ &= 1 - \sum_{i=1, S^{\max}_{{\rm sim},i} \leq S^{\max}_{\rm data}}^{i=N_{\rm sim}} 1/N_{\rm sim} \end{aligned} \tag{8.5}$$

where $c(a) = \min(S^{\max})(a)$ and $f(S^{\max}, a)$ is the MC probed PDF of the $S^{\max}$ values.

We interpolate confidence interval contours for the remaining a values of the parameter space using 4th order polynomial fit.

In the next section we apply this procedure to the considered WMAP datasets and simulations and present our results.

8.5 Results

In Fig. 8.3 we present results of the all-parameter-space search for the WMAP first year ILC map (left panel), and the three year WMAP INC map (right panel).

The signal at $\sim 11^\circ$ in Fig. 4 of Roukema et al. (2004) is reproduced and plotted with red crosses in Fig. 8.3 (middle-left).

Clearly, it is not necessary to process large number of simulations to resolve high confidence level contours, since all the datasets are consistent with the simply connected space GRF simulations at a confidence level as low as about 68% at all smoothing scales.

Is is easily seen that as the circle size shrinks to zero ($a \leq 2^\circ$), it is difficult to estimate precisely the significance of the detections since the CL contours steepen, while the accuracy of the S value determination is fixed at $\Delta S \sim 0.01$. This effect is most severe for large smoothing scales, as expected.

We note that the correlations S tend to increase in relation to the smoothing length applied to the data.

In particular, the signal reported in Roukema et al. (2004) is sensitive to increases in the smoothing length. While at the smoothing length of $\lambda = 1^\circ$ there is practically no excess maximum in $S^{\max}(a \approx 12^\circ)$ for $s = -36^\circ$ relative to the $S^{\max}$ values for $s = 0^\circ$ and $s = +36^\circ$, on the other hand, the excess is clearly seen at the smoothing length of $\lambda = 4^\circ$, where its significance increases almost up to the 95% CL.

[4] Statistically there are some small differences in the S values resulting from probing slightly different angular separations (arising due to different separations of pairs of points, for the same pair of matched circles, when calculated with two different phase twists: $s = 0$ and $= \pm 36$), due to the dependence of the S value on the underlying CMB two-point correlation function. For such small twists, this is found to be of the same magnitude as the statistical error on the $\Delta S (\approx 0.01)$ for all considered circle radii.

The results for the three year INC data are consistent with the first year data, in the sense that no statistically important excess correlations are found.

In addition to analysing our two primary datasets, we also carried out the following complementary searches.

We completed an all-parameter-space search using the three year WMAP ILC data. We find that the excess correlation corresponding to the hypothesized PDS model is weakened for all smoothing lengths (for $s \in \{0^\circ, +36^\circ\}$), and basically indistinguishable from the noise of what would be false positive detections if we were to define the 68% confidence level as a detection threshold. For $s = -36^\circ$, we plot the $S^{\max}(a)$ values for the three year ILC data with a black line in the left column of Fig. 8.3.

Our other complementary test was that we performed a 0.5° resolution all parameter space search, using first and three year INC data and did not find extra strong localized correlations. However since the computation time increases with the power of the increased resolution (i.e. increasing the resolution by a factor of two increases the calculation time by a factor of 2^n where $n = 4$ is the number of parameters in parameter space) we haven't performed the significance analysis with simulations, and therefore we do not present these results.

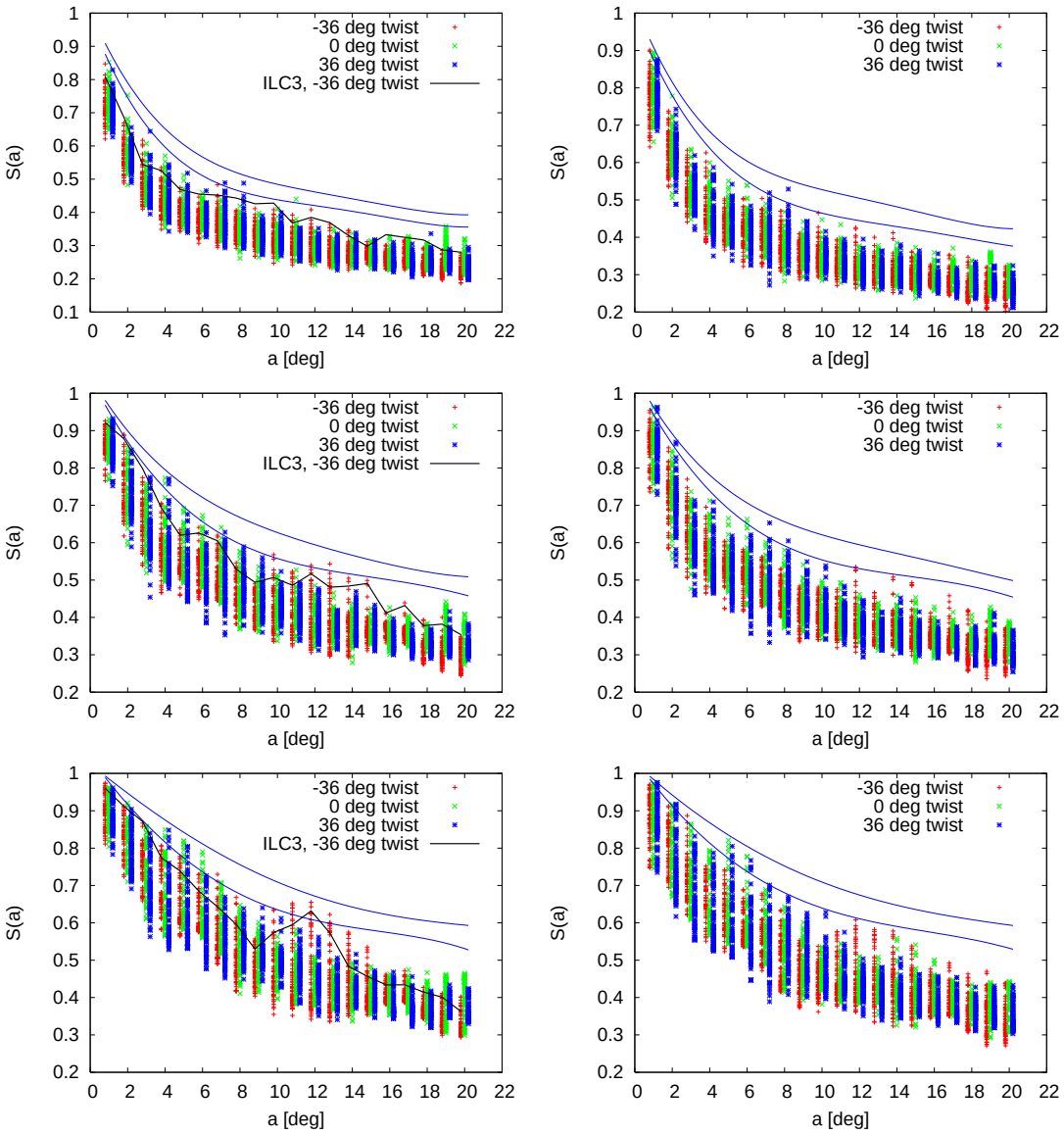


Fig 8.3: Results of the search in the parameter space (see Sect. 8.4.2) for the highest S correlations in the first year WMAP ILC map (left column) and three year INC map (right column) smoothed to $\lambda = 1^\circ$ (top), $\lambda = 2^\circ$ (middle), $\lambda = 4^\circ$ (bottom). For clarity only the highest 72 $S(a)$ statistic values are plotted for each of the three considered phase shifts: $-36^\circ, 0^\circ, 36^\circ$ marked with red ($+$), green ($\times$) and blue ($*$) respectively, and separated by 0.2° offset for better visualization and comparison. The red crosses correspond to the PDS model. The thick solid line in the left column show the $S^{\max}$ values for a search in the three-year ILC data with a phase shift of -36°. The 68% and 95% confidence level contours from $N_{\rm sim} = 100$ simulations are over-plotted. Clearly we reproduce the results of Roukema et al. (2004). Most of the points with the highest S values in the range of $10^\circ \leq a \leq 12^\circ$ closely correspond to the solution depicted in Fig. 8.1. It is easily seen that much higher correlation coefficients would have been required in order to significantly reject the null hypothesis that the Universe is simply connected in favor of the PDS model alternative hypothesis.

8.6 Power spectrum, cut sky and smoothing length dependence on S correlations.

8.6.1 S dependence on the CMB power spectrum

As mentioned in Sect. 8.3 the S correlation value depends not only on the particular alignment of CMB fluctuations, but also on the underlying CMB two-point correlation function (or its Legendre transform – power spectrum), as it is also a two point statistic. Therefore any discrepancies of the data from the assumed model will be statistically imprinted onto the S values evaluated from simulated maps (generated according to the assumed model). This will lead to biases in estimates of the confidence level thresholds. Given that the nature of the large scale (mostly quadrupole and octupole) anomalies of the WMAP data w.r.t. concordance LCDM model is unknown, we assumed a model independent approach for generating simulations as described in Sect. 8.4. In this section we show the impact of biasing of S values due to large scale uncertainties in the assumed CMB power spectrum.

In order to quantify the impact of the underlying CMB power spectrum on the S correlator values, we perform the following exercise.
We create two INC simulations of the WMAP three year data, for two different power spectra, yet keeping exactly the same phase information in both cases.
One simulation is made using exactly the power spectrum reconstructed from the WMAP observations as in Sect. 8.3 (i.e. neglecting cosmic variance effects).
The other simulation is made using a random realization of the best fit LCDM model (Hinshaw et al., 2003b) but with the same phase information as in the first simulation. We Gaussian smooth them to common resolution with beam of FWHM$= 1^\circ$. This guarantees that the differences in S values will only be due to different underlying power spectra[5]. More specifically, since the LCDM model yields a good fit to the WMAP data for large l's, the discrepancies will be only due to the large scale anomalies. Using 20 random orientations of the fundamental dodecahedron, we then calculate the average of differences of the S correlator values between the two maps. That is, we calculate: $\langle \Delta S \rangle = \langle S_{LCDM} - S_{WMAP} \rangle$. Hence, the positive $\langle \Delta S \rangle$ values show the excess correlations that one would additionally get if the LCDM model was assumed, as compared to the "exact" realization of the reconstructed WMAP power spectrum (and *vice-versa* for the negative values).

In Fig. 8.4 the $\langle \Delta S \rangle$ relation is shown with red line. Clearly, the strongest additional correlations appear for the smallest circles, i.e. for the largest angular separations in the two-point correlation function, and are as large as ~ 0.15. (These scales are additionally contaminated due to large smoothing scales as compared to the circle size). While for the circle sizes of about $a \sim 10^\circ$ the effect is small, it is obvious that at smaller circle sizes the confidence level contours obtained from the simulations performed according to the LCDM model would be too conservative than it is needed. Therefore, in this regard, given that the LCDM model is currently widely accepted, we consider our analysis to be very

[5]The noise component is negligible at the considered smoothing scales of 1°

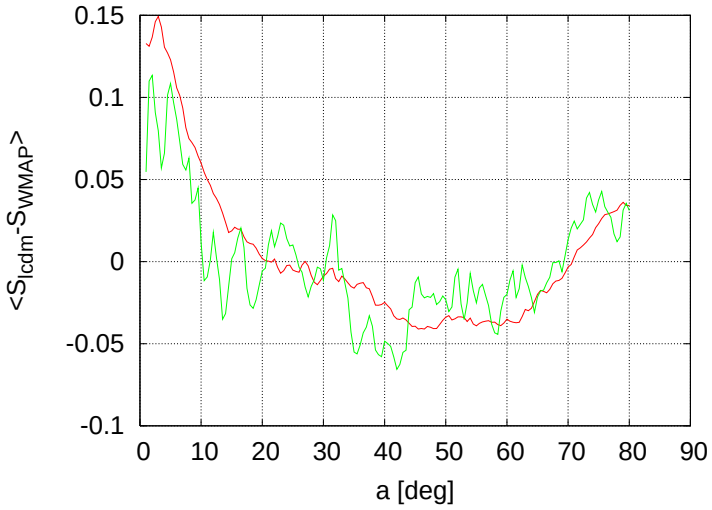


Fig 8.4: Average difference $\langle \Delta S \rangle = \langle S_{LCDM} - S_{WMAP} \rangle$ correlator values, as a function of circle size, statistically probing different regions of the underlying power spectrum (red line). With green line also the difference relation is plotted for the orientation of the dodecahedron depicted in Fig. 8.1.

conservative. This plot includes the dependence of $\langle \Delta S \rangle$ for larger circle sizes than those which we analyzed. And it is easily seen that for scales of about 40° the confidence level intervals would be underestimated.

We therefore conclude that the simply connected space hypothesis would be consistent with the data at yet even smaller CL, if the LCDM model were assumed for the generation of simulations.

8.6.2 S dependence on galactic sky cut

In this section we show the effects of different galactic sky cuts on the resulting amplitude of the correlation signal. For this purpose we work with the first year WMAP ILC map, smoothed to a resolution of 2° without mask.

We compute the the correlation statistics for the orientation of the dodecahedron found in Roukema et al. (2004) and confirmed in this work (i.e. for the dodecahedron with the following face centers
$(l^\circ, b^\circ) = \{(252, 65), (51, 51), (144, 38), (207, 10), (271, 3), (332, 25)\}$ and their opposites). We then apply different sky masks and show the $S(a)$ relations in Fig. 8.5. We used three different sky masks: which we call “bgc”, “Kp2”, and “kp03”. The first of these is the one that was used in Roukema et al. (2004) (i.e. the galactic plane is masked for $|b| < 2^\circ$, and points within 20° from the Galactic Center are masked). The second and third sky masks correspond to the Kp2 and the third year Kp0 sky masks (Bennett et al., 2003b).

Clearly, the fine details on the sky masking in case of ILC maps do not have large impact on the resulting S correlation value, except for the “no mask” case where the correlation peak is lower (most likely due to some residual foreground contaminations of the Galactic

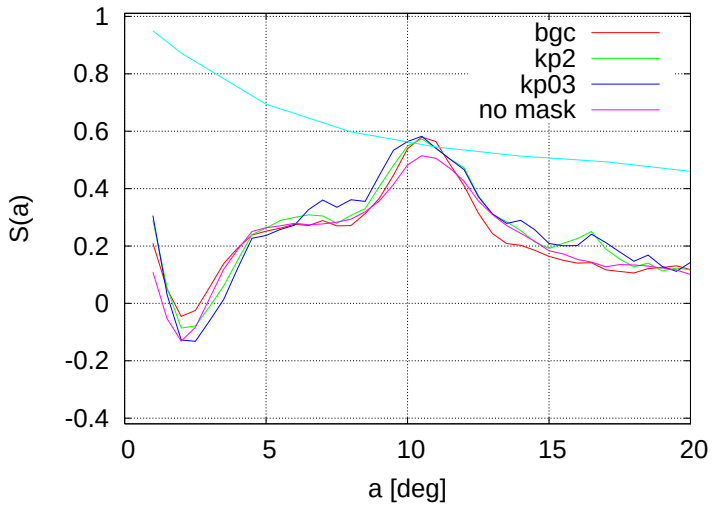


Fig 8.5: Effects of different sky masks on the $S(a)$ correlation statistic. See text for mask abbreviations definition (Sect. 8.6.2). One sigma CL contour derived for the case of Kp2 is also plotted.

plane).

We do not show the galactic sky cut dependence for the INC dataset. This is because there are stronger foregrounds present than for the ILC dataset, so that it is not straightforward to presmooth the map without contaminating the "clean" regions of the sky by the Gaussian tails of smoothing kernel.

8.6.3 S dependence on smoothing length

In this section we present the impact of different smoothing lengths used during map smoothing process on the S correlation values. We show the dependence using the dodecahedron orientation corresponding to the highest correlation value ($S^{\rm max}$ value) at $a = 12^\circ$ in Fig. 8.3 (middle-right) which corresponds to the dodecahedron with faces $(l^\circ, b^\circ) = \{(49, 51), (91, -2), (144, 40), (256, 64), (333, 23), (28, -10)\}$ and their opposites. This closely matches the dodecahedron depicted in Fig. 8.1. We use the WMAP three year INC map and the Kp2 sky mask.

In Fig. 8.6 we show

the $S(a)$ values for three different smoothing lengths used in the analysis. The characteristic trend towards higher correlation coefficients as smoothing length increases is clearly apparent, but the the 68% CL contours increase by about the same amount.

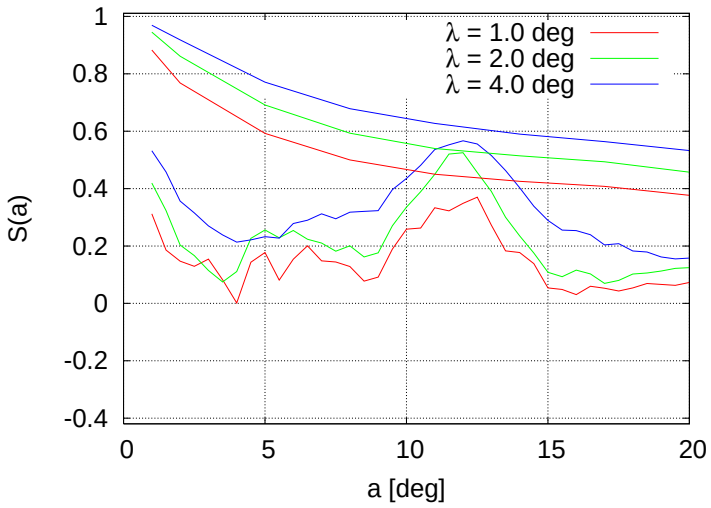


Fig 8.6: Effect of different smoothing lengths on the $S(a)$ correlation statistic. We plot $S(a)$ dependence for the three consecutive smoothing lengths used in the analysis. Corresponding 1σ confidence thresholds for each smoothing scale are also shown.

8.7 Discussion

The analysis of the correlations derived from the data and presented in the previous section finds no statistically-significant detections. The cross-correlations of the S values, obtained for different angular radii of the matched circles, were however neglected. It is of course faster to compute confidence intervals for a sparse parameter space and interpolate in between. However, the significance of any detections found this way (i.e. conditionally to the *a priori* assumed circle radii) might be overestimated, compared to the case when all possible correlations were accounted for in the full covariance matrix analysis. In the present work, since we do not find any significant deviations from the null hypothesis (i.e. we do not find any strong outliers in the S correlations) in any individually probed value of the "a" parameter, we find no need for any further extensions to the significance analysis already pursued.

We note that these cross-correlations are present not just in the data, but also in the Monte-Carlo simulations, so they affect the analyses of simulations and data to the same extent.

Finally, we note that our statistical approach of considering only the maximal S correlation values could be altered to consider the full distributions of S correlations. However, we are especially interested in viable candidates for non-trivial topology (especially in the proposed correlation signal around angular circle radii of 11°) and as such, the models with the largest S values are the best candidates. Given that there can be only one correct orientation of the fundamental dodecahedron and hence only one S correlation value corresponding to it (most likely the largest locally found S value), in the alternative way involving the full distributions of S correlations, the test would be heavily dominated by numerous values that will not be associated with the true topological correlation signal. As a result, the test

would mostly measure a degree of consistency between the simulations and data with respect to underlying two-point correlations via the circles on the sky, rather than candidates for non-trivial topology. We therefore build and rely on the statistics of specifically selected (in the full parameter space search) S values, one for a given simulation at a given circle radius, to build our statistics and reconstruct the confidence thresholds. Although we are aware that relying on the distributions of maximal values of random variates may lead to asymmetrical distributions with enhanced tails, we note that in the case of the S statistics, the possible values are by definition restricted to within the range $[-1, 1]$. We also note that it is unnecessary to resolve high confidence level contours, since the data are consistent with simulations mostly to within $\sim 1\sigma$ confidence contour.

8.8 Conclusions

In Roukema et al. (2004) it has been suggested that the shape of the space might be consistent with the Poincaré dodecahedral space (PDS) model (Fig. 8.1). This suggestion was due to an excess positive correlation in the matched circles test (Cornish et al., 1998b) of the first year WMAP ILC map, however the statistical significance of this excess was not specified.

We have revisited those results and found consistent correlation excess corresponding to the same orientation of the fundamental dodecahedron using independent software.

We extended and updated the matched circles search with the WMAP three year ILC data and the three year foreground reduced, inverse noise co-added map, and tested these at three different smoothing scales $(1^\circ, 2^\circ, 4^\circ)$.

We performed an analysis of the statistical significance of the reported excess, based on realistic and very conservative MC GRF CMB simulations of the datasets.

We find that under "matched circles" tests, both the first and three year WMAP data are consistent with the simply-connected topology hypothesis, for all smoothing scales, at a confidence level as low as 68%, apart from the first year ILC data smoothed to 4°, which are consistent at 95% CL.

Chapter 9

Summary, conclusions and outlook

The current concordance cosmological model offers number of predictions that allow, or will allow experimentally verify its properties and concretize the inflationary physics that is now embedded into the standard cosmological model as a generic solution to the number of long-standing cosmological problems, and as a generator of a nearly-adiabatic, and nearly-Gaussian, density perturbations imprinted over nearly-flat, homogeneous and isotropic background, with nearly scale invariant power spectrum, and with small amount of tensor perturbations in a form of gravitational wave background. Some of these predictions, most notably, the curvature of spatial sections, statistical isotropy, Gaussianity and tilt of the primordial power spectrum are now currently accessible cosmological observables via the observations of the cosmic microwave background radiation (CMBR) and they serve as valuable discriminators between viable cosmological models. Violations of any of these predictions would escape beyond the scope of the standard model and would rise a need for its revision.

The work pursued in this thesis aimed at direct testing two of these predictions: the Gaussianity, and the statistical isotropy. Interestingly there exist few puzzling anomalies in the available observations of the CMBR, that are tentatively incompatible with the generic ΛCDM cosmological model. We performed a number of statistical tests: both in real, and in spherical harmonics space, utilizing the recent CMBR observations of the Wilkinson Microwave Anisotropy Probe (WMAP). Apart from what has already been known, we attempted to either seek for new deviations from the predictions of Gaussianity or statistical isotropy, or further explore the well-known anomalies via number of new and independent statistical assessments in order to quantify how strongly they violate the two predictions.

Throughout the work, the analysis was heavily based on realistic Monte-Carlo simulations of Gaussian random fields (GRF), fitted to the technical specifications of the WMAP satellite, and to the requirements of the generic simplest cosmological scenario: i.e. to the best-fit, reconstructed power spectrum of the ΛCDM model. These simulations defined a reference null hypothesis, and were created along with a number of statistical tools, estimators and other utilities, and were entirely implemented in a self-developed, dedicated numerical, object-oriented code, that has been developed over the last few years.

We have addressed number of problems some of which were published in a refereed cosmological journals.

In particular we performed a blind real-space test of Gaussianity and the statistical isotropy via local measurements one-point temperature distributions, and addressed the statistical significance in a number of different ways (chapter 3). We tested the significance of local deviations from Gaussianity (such as the cold spot), and global hemispherical asymmetry anomalies in the CMBR power distribution. This work has been published in Journal of Cosmology and Astroparticle Physics (Lew, 2008b), and the details are specifically summarized in section 3.8.

We developed and implemented a new efficient method for measuring the significance of the hemispherical power asymmetry, and working within a dipolar modulation model, for the first time, we explicitly estimated the relevant model parameters as a function of scale (multipole bin) (chapter 4). These results are currently accepted for publication in Journal of Cosmology and Astroparticle Physics and are also available as Lew (2008a), and the details of this analysis are summarized in section 4.9.

We also pursued a Gaussianity tests of the CMBR WMAP maps using Minkowski functionals (chapter 5). This work was motivated by the fact that the previous analyses were not optimized on testing gaussianity exclusively within a particular angular scale, and therefore were dominated by the cumulative effects up to the effective scale of the either: resolution defined by the pixel size of a low-resolution map, or by an applied low-ℓ-pass filter. In order to complement some of these tests we performed an analysis in a selected multipole bins, testing thereby the corresponding scales with larger signal-to-noise ratio, and detected a strong signatures of the residual galactic foregrounds contamination. The details of these tests are outlined in section 5.6. The work presented in chapter 5 is now in preparation for subsequent publication.

We have also pursued (chapter 8) a statistical analysis of one of the cosmological scenarios with compact space topology: i.e. the one with multiply-connected topology of Poincaré dodecahedral shape. These models were motivated by tentative indication from the recent measurements of curvature of space, hinting on its slightly positive values, as well as by the apparent suppression of power at the currently observable horizon scales. We tested the model via the method of identified circles. We found results consistent with our previous works on this subject, however we and managed to refute the previously found correlation signals as statistical unimportant. This work therefore complements the earlier one, in which the statistical significance analysis was not performed. This work has been published in Astronomy and Astrophysics (Lew & Roukema, 2008) and the main results are summarized in section 8.8.

Furthermore, we studied the well-known large-scale anomalies of the CMBR: i.e. the alignments of the low-ℓ multipoles, issues of so called m-preference (chapter 7) and the low order multipoles power spectrum shape anomalies (chapter 6). We introduced and performed a new real-space based tests, corresponding to their, previously studied, spherical harmonic space counterparts. We extended the search for the inter-multipole alignments,

quantified their significance, and proposed a number of possibly useful generalizations. We also discussed the stability and robustness of the methods assumed in the literature, and performed a comparison statistics using a number of currently available renditions of the CMBR full sky foregrounds cleaned maps. We also estimated the significance of the anomalous, at first glance, shape of the low order multipoles in the CMBR angular power spectrum. The analysis yielded consistent results with the previous similar analysis. The results and discussion are detailed in sections 6.4 and 6.5 respectively.

Outlook There is a number of future paths that will be of great interest from both: the scientific point of view, but also from the point of view of the numerical programming. Both of these always provide lots of fun and a large margin for optimizations.

Apart from few updates, which were already discussed in the concluding sanctions of the previous chapters, the practical constraints of the levels of the primordial non-Gaussianity will require generation of high resolution non-Gaussian simulations, which serve as testing ground for various statistical estimators. Also, utilization of the future, accurate CMBR polarization observations will increase the signal-to-noise ratio of maps of the primordial gravitational potential, reconstructed from temperature and polarization observations. Since the polarization radiation transfer function is out-of-phase with the temperature radiation transfer function, it will necessarily complement the primordial gravitational potential signals. These challenges open up a qualitatively new window on the aspects of tests of primordial non-Gaussianity, whose importance is continuously increasing in the recent years in the context of constraining consistent inflationary models.

Finally, as far as the constraints on the inflationary physics are concerned, it will be very interesting to address the issue of the significance of glitches, found in the CMBR power spectrum, and investigate these in light of the previously proposed inflationary models that can explain these features via modifications in the shape of the inflaton potential, especially via the expected glitch-like features that alter the primordial power spectrum at both the horizon scales and the sub-horizon scales.

Bibliography

Abramo, L. R., Bernui, A., Ferreira, I. S., Villela, T., & Wuensche, C. A. 2006, Phys. Rev. D, 74, 063506, [astro-ph/0604346]

Abramovitz, M. & Stegun, I. A. 1972, Handbook of Mathematical Functions

Acquaviva, V., Bartolo, N., Matarrese, S., & Riotto, A. 2003, Nuclear Physics B, 667, 119, [arXiv:astro-ph/0209156]

Adams, J. & Cresswell, B. 2003, in Astronomy, Cosmology and Fundamental Physics, ed. P. A. Shaver, L. Dilella, & A. Giménez, 415–+

Adams, J., Cresswell, B., & Easther, R. 2001, Phys. Rev. D, 64, 123514, [arXiv:astro-ph/0102236]

Akofor, E., Balachandran, A. P., Jo, S. G., Joseph, A., & Qureshi, B. A. 2007, ArXiv e-prints, 710, [0710.5897]

Alabidi, L. & Lidsey, J. E. 2008, ArXiv e-prints, 807, [0807.2181]

Alabidi, L. & Lyth, D. H. 2006, Journal of Cosmology and Astro-Particle Physics, 8, 13, [astro-ph/0603539]

Arkani-Hamed, N., Creminelli, P., Mukohyama, S., & Zaldarriaga, M. 2004, Journal of Cosmology and Astro-Particle Physics, 4, 1, [arXiv:hep-th/0312100]

Armendariz-Picon, C. & Pekowsky, L. 2008, ArXiv e-prints, 807, [0807.2687]

Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31, [astro-ph/0510447]

Astier, P. et al. 2006, Astron. Astrophys., 447, 31, [astro-ph/0510447]

Aurich, R. 1999a, ApJ, 524, 497

Aurich, R. 1999b, ArXiv Astrophysics e-prints, 3032

Aurich, R., Lustig, S., Steiner, F., & Then, H. 2004, Classical and Quantum Gravity, 21, 4901, [astro-ph/0403597]

Bacon, D. J., Massey, R. J., Refregier, A. R., & Ellis, R. S. 2003, MNRAS, 344, 673, [arXiv:astro-ph/0203134]

Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625, [arXiv:astro-ph/0003008]

Bahcall, N. A., Dong, F., Bode, P., et al. 2003, ApJ, 585, 182, [arXiv:astro-ph/0205490]

Barnaby, N. & Cline, J. M. 2008, Journal of Cosmology and Astro-Particle Physics, 6, 30, [arXiv:0802.3218]

Barrow, J. D., Ferreira, P. G., & Silk, J. 1997, Physical Review Letters, 78, 3610, [arXiv:astro-ph/9701063]

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291, [arXiv:astro-ph/9912508]

Bartolo, N., Komatsu, E., Matarrese, S., & Riotto, A. 2004, Phys. Rep., 402, 103, [astro-ph/0406398]

Bartolo, N. & Liddle, A. R. 2002, Phys. Rev. D, 65, 121301, [arXiv:astro-ph/0203076]

Bartolo, N., Matarrese, S., & Riotto, A. 2004, Phys. Rev., D69, 043503, [hep-ph/0309033]

Bechtold, J., Crotts, A. P. S., Duncan, R. C., & Fang, Y. 1994, ApJ, 437, L83, [arXiv:astro-ph/9409007]

Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003a, ApJS, 148, 1, [astro-ph/0302207]

Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003b, ApJS, 148, 97, [astro-ph/0302208]

Bennett, C. L., Kogut, A., Hinshaw, G., et al. 1994, ApJ, 436, 423

Bennett, C. L., Smoot, G. F., Hinshaw, G., et al. 1992, ApJ, 396, L7

Bernardeau, F., Brunier, T., & Uzan, J.-P. 2006, ArXiv Astrophysics e-prints, [astro-ph/0604200]

Bernui, A. & Hipólito-Ricaldi, W. S. 2008, MNRAS, 929, [0807.1076]

Bernui, A., Mota, B., Rebouças, M. J., & Tavakol, R. 2007a, International Journal of Modern Physics D, 16, 411, [0706.0575]

Bernui, A., Mota, B., Rebouças, M. J., & Tavakol, R. 2007b, A&A, 464, 479, [astro-ph/0511666]

Bielewicz, P., Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2005, ApJ, 635, 750

Bielewicz, P., Górski, K. M., & Banday, A. J. 2004, MNRAS, 355, 1283, [astro-ph/0405007]

Boggess, N. W. 1992, JRASC, 86, 282

Bridges, M., McEwen, J. D., Lasenby, A. N., & Hobson, M. P. 2007, MNRAS, 377, 1473, [arXiv:astro-ph/0605325]

Brown, I. & Crittenden, R. 2005, Phys. Rev. D, 72, 063002

Cabella, P., Hansen, F., Marinucci, D., Pagano, D., & Vittorio, N. 2004, Phys. Rev. D, 69, 063007, [astro-ph/0401307]

Cabella, P., Hansen, F. K., Liguori, M., et al. 2006, MNRAS, 369, 819

Cabella, P., Liguori, M., Hansen, F. K., et al. 2005, MNRAS, 358, 684

Cabella, P. & Marinucci, D. 2008, ArXiv e-prints, 807, [0807.1816]

Caillerie, S., Lachièze-Rey, M., Luminet, J.-P., et al. 2007, A&A, 476, 691, [0705.0217]

Cayón, L., Jin, J., & Treaster, A. 2005, MNRAS, 362, 826

Cayón, L., Martínez-González, E., Argüeso, F., Banday, A. J., & Górski, K. M. 2003, MNRAS, 339, 1189

Chen, G., Mukherjee, P., Kahniashvili, T., Ratra, B., & Wang, Y. 2004, ApJ, 611, 655, [arXiv:astro-ph/0403695]

Chen, G. & Szapudi, I. 2006a, ArXiv Astrophysics e-prints, [astro-ph/0606394]

Chen, G. & Szapudi, I. 2006b, ApJ, 647, L87, [astro-ph/0606394]

Chen, X., Easther, R., & Lim, E. A. 2007a, Journal of Cosmology and Astro-Particle Physics, 6, 23, [arXiv:astro-ph/0611645]

Chen, X., Huang, M.-x., Kachru, S., & Shiu, G. 2007b, Journal of Cosmology and Astro-Particle Physics, 1, 2, [arXiv:hep-th/0605045]

Chiang, L.-Y., Naselsky, P. D., Verkhodanov, O. V., & Way, M. J. 2003, ApJ, 590, L65

Cline, J. M., Crotty, P., & Lesgourgues, J. 2003, JCAP, 0309, 010

Coc, A. 2008, in American Institute of Physics Conference Series, Vol. 1016, American Institute of Physics Conference Series, 119–126

Cole, S. et al. 2005, Mon. Not. Roy. Astron. Soc., 362, 505, [astro-ph/0501174]

Colley, W. N. & Gott, J. R. 2003, MNRAS, 344, 686

Contaldi, C. R., Peloso, M., Kofman, L., & Linde, A. 2003, Journal of Cosmology and Astro-Particle Physics, 7, 2, [astro-ph/0303636]

Cooray, A. & Hu, W. 2001, ApJ, 554, 56, [arXiv:astro-ph/0012087]

Copi, C. J., Huterer, D., Schwarz, D. J., & Starkman, G. D. 2006a, MNRAS, 367, 79

Copi, C. J., Huterer, D., Schwarz, D. J., & Starkman, G. D. 2006b, ArXiv Astrophysics e-prints, [astro-ph/0605135]

Copi, C. J., Huterer, D., Schwarz, D. J., & Starkman, G. D. 2008, ArXiv e-prints, 808, [0808.3767]

Copi, C. J., Huterer, D., & Starkman, G. D. 2004, Phys. Rev. D, 70, 043515, [astro-ph/0310511]

Cornish, N., Spergel, D., & Starkman, G. 1997, ArXiv Astrophysics e-prints, 8225

Cornish, N. J., Spergel, D., & Starkman, G. 1998a, Phys. Rev. D, 57, 5982

Cornish, N. J. & Spergel, D. N. 1999, arXiv:astro-ph/9906401

Cornish, N. J., Spergel, D. N., & Starkman, G. D. 1996, ArXiv General Relativity and Quantum Cosmology e-prints, [gr-qc/9602039]

Cornish, N. J., Spergel, D. N., & Starkman, G. D. 1998b, Classical and Quantum Gravity, 15, 2657

Cornish, N. J., Spergel, D. N., Starkman, G. D., & Komatsu, E. 2004, Physical Review Letters, 92, 201302, [astro-ph/0310233]

Covi, L., Hamann, J., Melchiorri, A., Slosar, A., & Sorbera, I. 2006, Phys. Rev. D, 74, 083509, [astro-ph/0606452]

Cristiani, S., D'Odorico, S., D'Odorico, V., et al. 1997, MNRAS, 285, 209, [arXiv:astro-ph/9610006]

Cruz, M., Cayón, L., Martínez-González, E., Vielva, P., & Jin, J. 2007, ApJ, 655, 11, [astro-ph/0603859]

Cruz, M., Martínez-González, E., Vielva, P., & Cayón, L. 2005, MNRAS, 356, 29

Cruz, M., Martinez-Gonzalez, E., Vielva, P., et al. 2008, ArXiv e-prints, 804, [0804.2904]

Cumberbatch, D., Ichikawa, K., Kawasaki, M., et al. 2007, Phys. Rev. D, 76, 123005, [arXiv:0708.0095]

Curto, A., Aumont, J., Macías-Pérez, J. F., et al. 2007, A&A, 474, 23, [astro-ph/0612148]

Curto, A., Macias-Perez, J. F., Martinez-Gonzalez, E., et al. 2008a, ArXiv e-prints, 804, [0804.0136]

Curto, A., Martinez-Gonzalez, E., Mukherjee, P., et al. 2008b, ArXiv e-prints, 807, [0807.0231]

Cyburt, R. H., Fields, B. D., & Olive, K. A. 2003, Physics Letters B, 567, 227, [arXiv:astro-ph/0302431]

Cyburt, R. H., Fields, B. D., & Olive, K. A. 2008, ArXiv e-prints, 808, [0808.2818]

de Oliveira-Costa, A. & Smoot, G. F. 1995, ApJ, 448, 477

de Oliveira-Costa, A., Smoot, G. F., & Starobinsky, A. A. 1996, ApJ, 468, 457

de Oliveira-Costa, A. & Tegmark, M. 2006, Phys. Rev. D, 74, 023005, [astro-ph/0603369]

de Oliveira-Costa, A., Tegmark, M., Zaldarriaga, M., & Hamilton, A. 2004, Phys. Rev. D, 69, 063516, [astro-ph/0307282]

de Troia, G., Ade, P. A. R., Bock, J. J., et al. 2007, New Astronomy Review, 51, 250, [0705.1615]

Delabrouille, J., Cardoso, J. ., Le Jeune, M., et al. 2008, ArXiv e-prints, 807, [0807.0773]

Destri, C., de Vega, H. J., & Sanchez, N. G. 2008, ArXiv e-prints, 804, [0804.2387]

Diego, J. M., Vielva, P., Martinez-Gonalez, E., & Silk, J. 2004, ArXiv Astrophysics e-prints, [astro-ph/0403561]

Dineen, P., Rocha, G., & Coles, P. 2005, MNRAS, 358, 1285, [astro-ph/0404356]

Donoghue, E. P. & Donoghue, J. F. 2005, Phys. Rev. D, 71, 043002, [astro-ph/0411237]

Dunkley, J., Komatsu, E., Nolta, M. R., et al. 2008, ArXiv e-prints, 803, [0803.0586]

Durrer, R., Juszkiewicz, R., Kunz, M., & Uzan, J.-P. 2000, Physical Review D, 62, 021301

Dvali, G., Gruzinov, A., & Zaldarriaga, M. 2004, Phys. Rev. D, 69, 023505, [arXiv:astro-ph/0303591]

Dvorkin, C., Peiris, H. V., & Hu, W. 2008, Phys. Rev. D, 77, 063008, [arXiv:0711.2321]

Efstathiou, G. 2003a, MNRAS, 343, L95, [arXiv:astro-ph/0303127]

Efstathiou, G. 2003b, MNRAS, 346, L26, [astro-ph/0306431]

Efstathiou, G. 2004, MNRAS, 348, 885, [astro-ph/0310207]

Eisenstein, D. J. et al. 2005, Astrophys. J., 633, 560, [astro-ph/0501171]

Enqvist, K. & Takahashi, T. 2008, ArXiv e-prints, 807, [0807.3069]

Erickcek, A. L., Kamionkowski, M., & Carroll, S. M. 2008, ArXiv e-prints, 806, [0806.0377]

Eriksen, H. K., Banday, A. J., Górski, K. M., Hansen, F. K., & Lilje, P. B. 2007, ApJ, 660, L81, [astro-ph/0701089]

Eriksen, H. K., Dickinson, C., Jewell, J. B., et al. 2008, ApJ, 672, L87, [0709.1037]

Eriksen, H. K., Hansen, F. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004a, ApJ, 605, 14

Eriksen, H. K., Novikov, D. I., Lilje, P. B., Banday, A. J., & Górski, K. M. 2004b, ApJ, 612, 64

Ferreira, P. G., Magueijo, J., & Gorski, K. M. 1998, ApJ, 503, L1+

Finkbeiner, D. P. 2003, ApJS, 146, 407, [arXiv:astro-ph/0301558]

Finkbeiner, D. P., Davis, M., & Schlegel, D. J. 1999, ApJ, 524, 867, [arXiv:astro-ph/9905128]

Gangui, A., Martin, J., & Sakellariadou, M. 2002, Phys. Rev. D, 66, 083502, [arXiv:astro-ph/0205202]

Garriga, J. & Mukhanov, V. F. 1999, Physics Letters B, 458, 219, [arXiv:hep-th/9904176]

Gausmann, E., Lehoucq, R., Luminet, J.-P., Uzan, J.-P., & Weeks, J. 2001, Classical and Quantum Gravity, 18, 5155 (arXiv:gr-qc/0106033)

Gaztañaga, E. & Wagg, J. 2003, Phys. Rev. D, 68, 021302

Gaztañaga, E., Wagg, J., Multamäki, T., Montaña, A., & Hughes, D. H. 2003, MNRAS, 346, 47

Gaztanaga, E., Cabre, A., Castander, F., Crocce, M., & Fosalba, P. 2008, ArXiv e-prints, 807, [0807.2448]

Germani, C. & Liguori, M. 2008, General Relativity and Gravitation, 100, [0706.0025]

Giannantonio, T., Scranton, R., Crittenden, R. G., et al. 2008, Phys. Rev. D, 77, 123520, [arXiv:0801.4380]

Gnedin, N. Y. & Hamilton, A. J. S. 2002, MNRAS, 334, 107, [arXiv:astro-ph/0111194]

Gold, B., Bennett, C. L., Hill, R. S., et al. 2008, ArXiv e-prints, 803, [0803.0715]

Golovnev, A., Mukhanov, V., & Vanchurin, V. 2008, Journal of Cosmology and Astro-Particle Physics, 6, 9, [arXiv:0802.2068]

Gomero, G. I. & Rebouças, M. J. 2003, Physics Letters A, 311, 319, [gr-qc/0202094]

Gomero, G. I., Reboucas, M. J., & Tavakol, R. 2002a, ArXiv General Relativity and Quantum Cosmology e-prints, [gr-qc/0210016]

Gomero, G. I., Rebouças, M. J., & Teixeira, A. F. F. 2000, Phys. Lett A, 275, 355 (arXiv:gr-qc/9909078)

Gomero, G. I., Teixeira, A. F. F., Rebouças, M. J., & Bernui, A. 2002b, Int. J. Mod. Phys. D, 11, 869 (arXiv:gr-qc/9811038)

Gordon, C. 2007, ApJ, 656, 636, [arXiv:astro-ph/0607423]

Gordon, C., Hu, W., Huterer, D., & Crawford, T. 2005, Phys. Rev. D, 72, 103002, [astro-ph/0509301]

Gordon, C. & Lewis, A. 2003, Phys. Rev. D, 67, 123513, [arXiv:astro-ph/0212248]

Gordon, C. & Trotta, R. 2007, MNRAS, 382, 1859, [arXiv:0706.3014]

Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759, [astro-ph/0409513]

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Hajian, A. & Souradeep, T. 2003, ArXiv Astrophysics e-prints, [astro-ph/0301590]

Hajian, A. & Souradeep, T. 2005, ArXiv Astrophysics e-prints, [astro-ph/0501001]

Hajian, A. & Souradeep, T. 2006, Phys. Rev. D, 74, 123521, [astro-ph/0607153]

Hansen, F. K., Banday, A. J., & Górski, K. M. 2004a, MNRAS, 354, 641, [astro-ph/0404206]

Hansen, F. K., Cabella, P., Marinucci, D., & Vittorio, N. 2004b, ApJ, 607, L67, [astro-ph/0402396]

Hansen, F. K., Górski, K. M., & Hivon, E. 2002, MNRAS, 336, 1304, [arXiv:astro-ph/0207464]

Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399, [astro-ph/0608064]

Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1

Hikage, C., Komatsu, E., & Matsubara, T. 2006, ArXiv Astrophysics e-prints, [astro-ph/0607284]

Hinshaw, G., Barnes, C., Bennett, C. L., et al. 2003a, ApJS, 148, 63, [astro-ph/0302222]

Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2007, ApJS, 170, 288, [astro-ph/0603451]

Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2006, ArXiv Astrophysics e-prints, [astro-ph/0603451]

Hinshaw, G., Spergel, D. N., Verde, L., et al. 2003b, ApJS, 148, 135, [astro-ph/0302217]

Hinshaw, G., Weiland, J. L., Hill, R. S., et al. 2008, ArXiv e-prints, 803, [0803.0732]

Hirabayashi, H., Fomalont, E. B., Horiuchi, S., et al. 2000, PASJ, 52, 997

Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2, [astro-ph/0105302]

Ho, S., Hirata, C. M., Padmanabhan, N., Seljak, U., & Bahcall, N. 2008, ArXiv e-prints, 801, [0801.0642]

Hodges, H. M., Blumenthal, G. R., Kofman, L. A., & Primack, J. R. 1990, Nuclear Physics B, 335, 197

Hoekstra, H., Mellier, Y., van Waerbeke, L., et al. 2006, ApJ, 647, 116, [arXiv:astro-ph/0511089]

Holman, R. & Tolley, A. J. 2008, Journal of Cosmology and Astro-Particle Physics, 5, 1, [arXiv:0710.1302]

Huang, M.-X., Shiu, G., & Underwood, B. 2008, Phys. Rev. D, 77, 023511, [arXiv:0709.3299]

Huterer, D. 2006, New Astronomy Review, 50, 868, [astro-ph/0608318]

Inoue, K. T. 1999, Classical and Quantum Gravity, 16, 3071

Inoue, K. T. & Silk, J. 2006, ApJ, 648, 23, [astro-ph/0602478]

Inoue, K. T. & Silk, J. 2007, ApJ, 664, 650, [astro-ph/0612347]

Jaffe, T. R., Banday, A. J., Eriksen, H. K., Górski, K. M., & Hansen, F. K. 2005, ApJ, 629, L1, [astro-ph/0503213]

Jaffe, T. R., Hervik, S., Banday, A. J., & Górski, K. M. 2006, ApJ, 644, 701, [arXiv:astro-ph/0512433]

Jannuzi, B. T., Bahcall, J. N., Bergeron, J., et al. 1998, ApJS, 118, 1, [arXiv:astro-ph/9805148]

Jedamzik, K., Choi, K.-Y., Roszkowski, L., & Ruiz de Austri, R. 2006, Journal of Cosmology and Astro-Particle Physics, 7, 7, [arXiv:hep-ph/0512044]

Jing, Y. & Fang, L. 1994, ArXiv Astrophysics e-prints, 9072

Jittoh, T., Kohri, K., Koike, M., et al. 2007, Phys. Rev. D, 76, 125023, [arXiv:0704.2914]

Kahniashvili, T., Maravin, Y., & Kosowsky, A. 2008, ArXiv e-prints, 806, [0806.1876]

Kaiser, N., Wilson, G., & Luppino, G. A. 2000, ArXiv Astrophysics e-prints, [astro-ph/0003338]

Kawasaki, M., Kohri, K., & Moroi, T. 2005, Phys. Rev. D, 71, 083502, [arXiv:astro-ph/0408426]

Kecskemeti, S., Maiden, J., Shiu, G., & Underwood, B. 2006, ArXiv High Energy Physics - Theory e-prints, [hep-th/0605189]

Kenney, J. F. & Keeping, E. S. 1951, Mathematics of Statistics Pt. 2, 2nd ed.

Key, J. S., Cornish, N. J., Spergel, D. N., & Starkman, G. D. 2007, Phys. Rev. D, 75, 084034, [astro-ph/0604616]

Kim, J., Naselsky, P., & Christensen, P. R. 2008, ArXiv e-prints, 803, [0803.1394]

Kim, S. A. & Liddle, A. R. 2006, ArXiv Astrophysics e-prints, [astro-ph/0605604]

Kinney, W. H., Kolb, E. W., Melchiorri, A., & Riotto, A. 2004, Phys. Rev. D, 69, 103516

Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2008, ArXiv e-prints, 803, [0803.0547]

Komatsu, E., Kogut, A., Nolta, M., et al. 2003, The Astrophysical Journal, 148, 119

Koyama, K., Mizuno, S., Vernizzi, F., & Wands, D. 2007, Journal of Cosmology and Astro-Particle Physics, 11, 24, [arXiv:0708.4321]

Koyama, K. & Wands, D. 2007, Journal of Cosmology and Astro-Particle Physics, 4, 8, [arXiv:hep-th/0703040]

Kuehr, H., Witzel, A., Pauliny-Toth, I. I. K., & Nauber, U. 1981, A&AS, 45, 367

Kunz, M., Aghanim, N., Cayon, L., et al. 2006, Phys. Rev. D, 73, 023511, [astro-ph/0510164]

Kunz, M., Aghanim, N., Riazuelo, A., & Forni, O. 2008, Phys. Rev. D, 77, 023525, [0704.3076]

Kusakabe, M., Kajino, T., Boyd, R. N., Yoshida, T., & Mathews, G. J. 2008, ApJ, 680, 846, [arXiv:0711.3858]

Land, K. & Magueijo, J. 2005a, MNRAS, 357, 994

Land, K. & Magueijo, J. 2005b, Physical Review Letters, 95, 071301, [astro-ph/0502237]

Land, K. & Magueijo, J. 2005c, Phys. Rev. D, 72, 101302, [astro-ph/0507289]

Land, K. & Magueijo, J. 2005d, MNRAS, 362, 838

Land, K. & Magueijo, J. 2006, MNRAS, 367, 1714

Land, K. & Magueijo, J. 2007, MNRAS, 378, 153, [astro-ph/0611518]

Landt, H., Padovani, P., Perlman, E. S., et al. 2001, MNRAS, 323, 757, [astro-ph/0012356]

Lehoucq, R., Luminet, J.-P., & Uzan, J.-P. 1999, A&A, 344, 735, [gr-qc/9604050]

Levin, J., Scannapieco, E., & Silk, J. 1998, Phys. Rev. D, 58, 103516

Lew, B. 2008a, (accepted for publication in JCAP), ArXiv e-prints, 808, [0808.2867]

Lew, B. 2008b, Journal of Cosmology and Astroparticle Physics, 2008, 017 (42pp), [0803.1409]

Lew, B. S. & Roukema, B. F. 2008, A&A, 482, 747, [0801.1358]

Lewis, A. & Bridle, S. 2002, Phys. Rev. D, 66, 103511, [arXiv:astro-ph/0205436]

Lewis, A. & Challinor, A. 2006, Phys. Rep., 429, 1, [arXiv:astro-ph/0601594]

Linde, A. 1994, Phys. Rev. D, 49, 748, [arXiv:astro-ph/9307002]

Linde, A. 2003, Journal of Cosmology and Astro-Particle Physics, 5, 2, [arXiv:astro-ph/0303245]

Linde, A. D. 1982, Physics Letters B, 108, 389

Lu, L., Sargent, W. L. W., Womble, D. S., & Takada-Hidai, M. 1996, ApJ, 472, 509, [arXiv:astro-ph/9606033]

Luminet, J. & Roukema, B. F. 1999, ArXiv Astrophysics e-prints, 1364

Luminet, J.-P., Weeks, J. R., Riazuelo, A., Lehoucq, R., & Uzan, J.-P. 2003, Nature, 425, 593, [astro-ph/0310253]

Lyth, D. H. & Stewart, E. D. 1996, Phys. Rev. D, 53, 1784, [arXiv:hep-ph/9510204]

Lyth, D. H., Ungarelli, C., & Wands, D. 2003, Phys. Rev. D, 67, 023503, [arXiv:astro-ph/0208055]

Lyth, D. H. & Wands, D. 2002, Physics Letters B, 524, 5, [arXiv:hep-ph/0110002]

Magueijo, J. & Medeiros, J. 2004, MNRAS, 351, L1

Magueijo, J. & Sorkin, R. D. 2007, MNRAS, 377, L39

Martinez-Gonzalez, E. 2008, ArXiv e-prints, 805, [0805.4157]

McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2006a, MNRAS, 371, L50, [astro-ph/0604305]

McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2006b, MNRAS, 369, 1858, [astro-ph/0510349]

Mollerach, S. 1990, Phys. Rev. D, 42, 313

Mota, B., Gomero, G. I., Rebouças, M. J., & Tavakol, R. 2004, Classical and Quantum Gravity, 21, 3361, [arXiv:astro-ph/0309371]

Mukherjee, P. & Wang, Y. 2004, ApJ, 613, 51

Naselsky, P., Chiang, L.-Y., Olesen, P., & Novikov, I. 2005, Phys. Rev. D, 72, 063512

Naselsky, P. & Kim, J. 2008, ArXiv e-prints, 804, [0804.3467]

Naselsky, P. D., Christensen, P. R., Coles, P., et al. 2007, ArXiv e-prints, 712, [0712.1118]

Nesseris, S. & Perivolaropoulos, L. 2004, Phys. Rev. D, 70, 043531, [arXiv:astro-ph/0401556]

Niarchou, A. & Jaffe, A. 2007, Physical Review Letters, 99, 081302

Niarchou, A. & Jaffe, A. H. 2006, in American Institute of Physics Conference Series, Vol. 848, Recent Advances in Astronomy and Astrophysics, ed. N. Solomos, 774–778

Page, L. et al. 2007, Astrophys. J. Suppl., 170, 335, [astro-ph/0603450]

Park, C.-G. 2004, MNRAS, 349, 313, [astro-ph/0307469]

Park, C.-G., Park, C., & Gott, J. R. I. 2006, ArXiv Astrophysics e-prints, [astro-ph/0608129]

Peiris, H. V., Komatsu, E., Verde, L., et al. 2003, ApJS, 148, 213, [astro-ph/0302225]

Perlman, E. S., Padovani, P., Giommi, P., et al. 1998, AJ, 115, 1253, [astro-ph/9801024]

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1998, ArXiv Astrophysics e-prints, 12133

Phillips, N. G. & Kogut, A. 2006, ApJ, 645, 820, [astro-ph/0404400]

Pietrobon, D., Amblard, A., Balbi, A., et al. 2008, ArXiv e-prints, 809, [0809.0010]

Primack, J. R. 2004, in IAU Symposium, Vol. 220, Dark Matter in Galaxies, ed. S. Ryder, D. Pisano, M. Walker, & K. Freeman, 53–+

Rakic, A., Rasanen, S., & Schwarz, D. J. 2006, ArXiv Astrophysics e-prints, [astro-ph/0609188]

Rakic, A. & Schwarz, D. J. 2007, ArXiv Astrophysics e-prints, [astro-ph/0703266]

Rassat, A., Land, K., Lahav, O., & Abdalla, F. B. 2006, ArXiv Astrophysics e-prints, [astro-ph/0610911]

Reichardt, C. L. & ACBAR collaboration. 2007, in Bulletin of the American Astronomical Society, Vol. 38, Bulletin of the American Astronomical Society, 938–+

Reichardt, C. L., Ade, P. A. R., Bock, J. J., et al. 2008, ArXiv e-prints, 801, [0801.1491]

Renka, R. J. 1997, ACM Trans. Math. Softw., 23, 435

Riazuelo, A., Caillerie, S., Lachièze-Rey, M., Lehoucq, R., & Luminet, J.-P. 2006, ArXiv Astrophysics e-prints, [astro-ph/0601433]

Riazuelo, A., Uzan, J.-P., Lehoucq, R., & Weeks, J. 2004a, Phys. Rev. D, 69, 103514, [astro-ph/0212223]

Riazuelo, A., Weeks, J., Uzan, J., Lehoucq, R., & Luminet, J. 2004b, Phys. Rev. D, 69, 103518 (arXiv:astro-ph/0311314)

Riess, A. G. et al. 2004, Astrophys. J., 607, 665, [astro-ph/0402512]

Rigopoulos, G. I., Shellard, E. P. S., & van Tent, B. J. W. 2007, Phys. Rev. D, 76, 083512, [arXiv:astro-ph/0511041]

Roukema, B. F. 2000, MNRAS, 312, 712, [astro-ph/9910272]

Roukema, B. F. & Blanlœil, V. 1998a, Classical and Quantum Gravity, 15, 2645+

Roukema, B. F. & Blanlœil, V. 1998b, Classical and Quantum Gravity, 15, 2645 (arXiv:astro-ph/9802083)

Roukema, B. F., Lew, B., Cechowska, M., Marecki, A., & Bajtlik, S. 2004, A&A, 423, 821, [astro-ph/0402608]

Rudnick, L., Brown, S., & Williams, L. R. 2007, ApJ, 671, 40, [arXiv:0704.0908]

Sachs, R. K. & Wolfe, A. M. 1967, ApJ, 147, 73

Samal, P. K., Saha, R., Jain, P., & Ralston, J. P. 2008, MNRAS, 385, 1718, [0708.2816]

Sasaki, M. 2008, Progress of Theoretical Physics, 120, 159, [arXiv:0805.0974]

Savage, R., Battye, R. A., Carreira, P., et al. 2004, MNRAS, 349, 973

Schmalzing, J. & Gorski, K. M. 1998, MON.NOT.R.ASTRON.SOC., 297, 355

Schwarz, D. J., Starkman, G. D., Huterer, D., & Copi, C. J. 2004, Physical Review Letters, 93, 221301

Seljak, U., Slosar, A., & McDonald, P. 2006, Journal of Cosmology and Astro-Particle Physics, 10, 14, [astro-ph/0604335]

Shafieloo, A. & Souradeep, T. 2007, ArXiv e-prints, 709, [0709.1944]

Shandarin, S. F. 2002, MNRAS, 331, 865

Sievers, J. L., Bond, J. R., Cartwright, J. K., et al. 2003, ApJ, 591, 599, [arXiv:astro-ph/0205387]

Silverstein, E. & Tong, D. 2004, Phys. Rev. D, 70, 103505, [arXiv:hep-th/0310221]

Sinha, R. & Souradeep, T. 2006, Phys. Rev. D, 74, 043518, [arXiv:astro-ph/0511808]

Smith, K. M. & Huterer, D. 2008, ArXiv e-prints, 805, [0805.2751]

Souradeep, T., Hajian, A., & Basak, S. 2006, ArXiv Astrophysics e-prints, [astro-ph/0607577]

Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS, 170, 377, [astro-ph/0603449]

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175, [astro-ph/0302209]

Starkman, G. D. 1998, Classical and Quantum Gravity, 15, 2529

Stevens, D., Scott, D., & Silk, J. 1993, Physical Review Letters, 71, 20

Stickel, M., Meisenheimer, K., & Kuehr, H. 1994, A&AS, 105, 211

Taylor, J. F., Ashdown, M. A. J., & Hobson, M. P. 2007, ArXiv e-prints, 708, [0708.2989]

Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. J. 2003, Phys. Rev. D, 68, 123523, [astro-ph/0302496]

Tegmark, M., Eisenstein, D., Strauss, M., et al. 2006, ArXiv Astrophysics e-prints, [astro-ph/0608632]

Tegmark, M. et al. 2006, Phys. Rev., D74, 123507, [astro-ph/0608632]

Teräsranta, H., Urpo, S., Wiren, S., & Valtonen, M. 2001, A&A, 368, 431

Uzan, J., Lehoucq, R., & Luminet, J. 1999a, ArXiv Astrophysics e-prints, 3155

Uzan, J.-P. 1998, Phys. Rev. D, 58, 087301+

Uzan, J.-P., Kirchner, U., & Ellis, G. F. R. 2003, MNRAS, 344, L65, [arXiv:astro-ph/0302597]

Uzan, J.-P., Lehoucq, R., & Luminet, J.-P. 1999b, A&A, 351, 766 (arXiv:astro-ph/9903155)

Vale, C. 2005, ArXiv Astrophysics e-prints, [astro-ph/0509039]

Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, ApJ, 609, 22

Wang, L. & Kamionkowski, M. 2000, Phys. Rev. D, 61, 063504, [astro-ph/9907431]

Weatherley, S. J., Warren, S. J., Croom, S. M., et al. 2003, MNRAS, 342, 9L (arXiv:astro-ph/0304290)

Weeks, J. & Gundermann, J. 2006, ArXiv Astrophysics e-prints, [astro-ph/0611640]

Weeks, J., Lehoucq, R., & Uzan, J. 2003, Classical and Quantum Gravity, 20, 1529 (arXiv:astro-ph/0209389)

Weeks, J. R. 2003, Modern Physics Letters A, 18, 2099, [astro-ph/0212006]

Weymann, R. J., Jannuzi, B. T., Lu, L., et al. 1998, ApJ, 506, 1, [arXiv:astro-ph/9806123]

Wiaux, Y., Vielva, P., Martinez-Gonzalez, E., & Vandergheynst, P. 2006, Physical Review Letters, 96, 151303

Wood-Vasey, W. M., Miknaitis, G., Stubbs, C. W., et al. 2007, ApJ, 666, 694, [astro-ph/0701041]

Wu, J. H. P., Balbi, A., Borrill, J., et al. 2001, Physical Review Letters, 87, 251303

Yadav, A. P. S., Komatsu, E., & Wandelt, B. D. 2007, ApJ, 664, 680, [astro-ph/0701921]

Yadav, A. P. S., Komatsu, E., Wandelt, B. D., et al. 2008, ApJ, 678, 578, [0711.4933]

Yadav, A. P. S. & Wandelt, B. D. 2008, Physical Review Letters, 100, 181301, [0712.1148]

Appendix

A-1 WMAP simulations

A-1.1 Signal and noise pseudo C_ℓ tests.

In order to assess the confidence levels we have performed Monte-Carlo simulations of the signal ($\ell_{max} \leq 1024$ and $C_{0,1} = 0$) and inhomogeneous but uncorrelated Gaussian noise maps for all DAs, according to the best fit to the ΛCDM model power spectrum, extracted from observations (Hinshaw et al., 2007). The simulations include the WMAP window functions for each DA. The $N_{\rm sim}(= 10^4)$ full sky simulations were generated at the *HEALPIX* resolution $n_s = 512$, for each DA and preprocessed in the same way as the data.

In Fig. A-1 an example of simulated, and recovered pseudo-C_ℓ is compared with the pseudo-C_ℓ of the WMAP data from channel Q1 as a simple consistency check. Similar simulations were performed for the remaining DAs.

As a simple consistency test, we present in Fig. A-2 statistics of the full sky WMAP third year data as compared with the simulations. The data are well consistent with the simulations.

Tests of modulated-simulations' power spectrum

As mentioned in Section 3.5.4 the modulation (Eq. 3.3) alters the underlying power spectrum of the modulated simulations and could possibly mislead us in the interpretation of the high rejection confidence level thresholds, reported in Table 3.2 for the modulated simulations, since the additional power could be merely a result of the inconsistency on grounds of the intrinsic power spectrum mismatch, rather than the regional variance distribution analysis, and violation of the statistical isotropy.

In this section we quantify this effect. We generate a set of 10 WMAP V3 simulations and modulate their CMB component using modulations $A_{1024}(\hat{\mathbf{n}} = (225^\circ, -27^\circ)) = \{0.114, 0.2\}$. We next extract their Kp03 cut sky pseudo-power spectra up to $\ell_{\rm max} = 700$ i.e. up to the scales where the noise component already strongly dominates over any possible CMB modulation signals. We derive the contribution to the total variance of the map

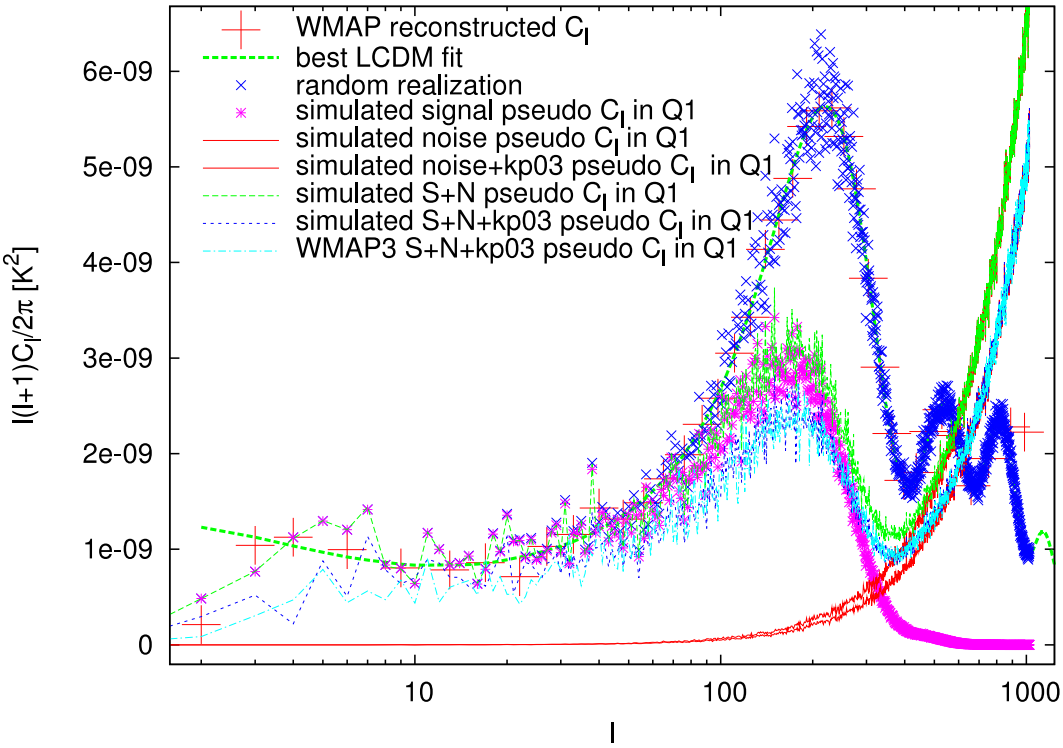


Fig A-1: Consistency check between the simulations and WMAP three-year observations. The pseudo-C_ℓ power spectra of the WMAP (light blue/dot-dashed curve) and its simulation (the underlying blue/dashed line) in channel Q1 are well consistent both in Kp03 sky cut regime and in the noise dominated regime. Reconstructed power spectrum of the Hinshaw et al. (2007) is plotted as big red crosses.

per multipole according to: $\sigma_\ell = \frac{2l+1}{4\pi} C_\ell$ ($\ell \in \{2, \ldots, 700\}$). We measure the degree of the consistency of the modulated simulations' power spectra with the non-modulated simulations' power spectra, using the unbiased estimator of the full covariance matrix $\Sigma_{ll'} = \mathrm{Cov}(\sigma_\ell, \sigma_{\ell'})$ derived from 3000, analogously prepared σ_ℓ vectors extracted from the GRF WMAP V3 simulations, and using the corresponding Monte-Carlo probed χ^2 values distribution from 223 additional simulations.

We find that the average rejection thresholds for the modulations $A_{1024} = 0.114$ and $A_{1024} = 0.2$ are 54% and 59% respectively with the standard deviation $\sim 30\%$ in both cases. We therefore conclude that our results given in Table 3.2 cannot result from simply alteration of the power spectra after the modulation field has been introduced. Similar results are obtained if the off-diagonal terms of the covariance matrix are neglected, which indicates that the low detection threshold does not result from the limited number of simu-

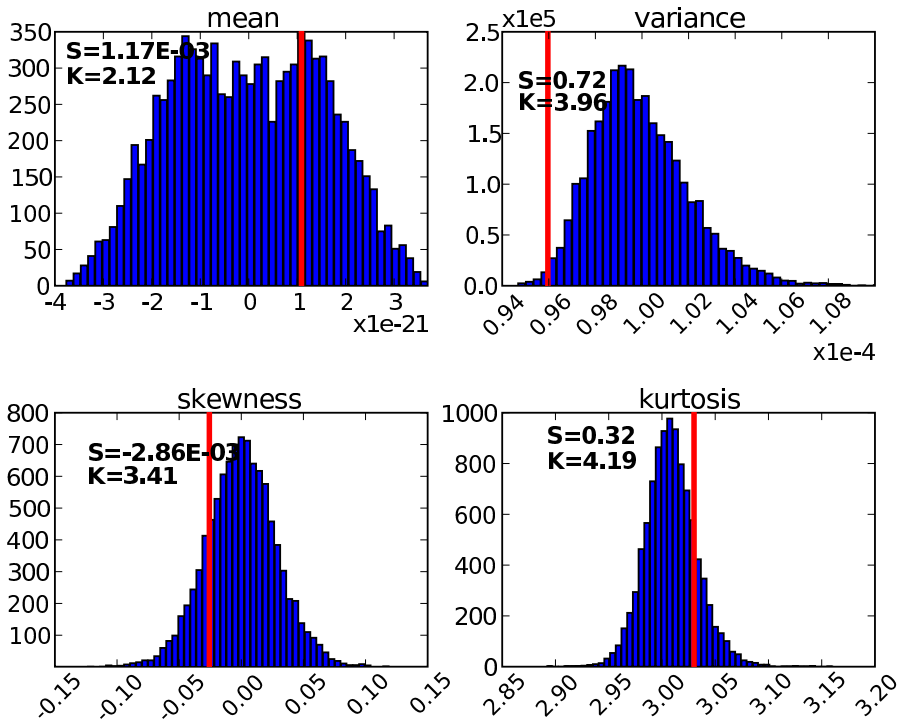


Fig A-2: The distributions of means, variances, skewness and kurtosis of $N_{\rm sim} = 10^4$, full sky, INC simulated data realizations, calculated outside the Kp03 sky mask. Skewness and kurtosis values of the distributions are also given. Vertical bars indicate the values derived from the WMAP three-year data. The quantile probabilities of the mean, variance, S, and K values of the WMAP data are $\{0.57, 0.015, 0.34, 0.35\}$ respectively – well consistent with Gaussian, random simulations. Low probability of the total variance results primarily from the low quadrupole (octupole) of the WMAP data as compared with the best fit ΛCDM model. Note that the distribution of the means of the simulations represents only a numerical noise since during preprocessing all maps were shifted to zero the mean outside the Kp03 sky mask ($< T > \rightarrow 0$) and therefore it does not carry any important cosmological information. The spectral analysis yields a similar results.

lations and a possible non-convergence of the covariance matrix.

A-2 Assessing statistical significance

In this appendix we give details on out statistical approaches. For fast reference, we summarize the principal symbols used in Table: A-1.

A-2.1 Individual region analysis

In the case of individual regions statistics for every ith region ($i \in \{1, .., N_{reg}(r, m)\}$) of every *multi-mask* $m \in \{1..N_m\}$ and for every MOD ($X \in \{\mathrm{m}, \sigma, \mathrm{S}, \mathrm{K}\}$) and every data-set ($d \in \{\mathrm{Q}, \mathrm{V}, \mathrm{W}, \mathrm{INC}\}$) we define a parameter vector $\mathbf{p} = \{r, m, d\}$ and independently calculate the tail probability $P\big(X_{\mathbf{p},i}\big)$ of occurrence of $X_{\mathbf{p},i}$ value of the data in the $N_{\text{sim}} = 10^4$ simulations, probing the corresponding probability distribution functions (PDF). The quantities $\mathrm{m}, \sigma, \mathrm{S}, \mathrm{K}$ correspond to the first four moments of distribution respectively. Hence we define $P\big(X_{\mathbf{p},i}\big)$ as:

$$P\big(X_{\mathbf{p},i}\big) \equiv P_q\big(|X^{sim}_{\mathbf{p},i} - Q_{2_{\mathbf{p},i}}| > |X^{data}_{\mathbf{p},i} - Q_{2_{\mathbf{p},i}}|\big) \tag{A-1}$$

where $Q_{2_{\mathbf{p},i}}$ is the second quartile of the corresponding PDF, and P_q is the quantile tail probability.

In principle, considering N simulations allows us to probe a region of the underlying PDF corresponding to Gaussian number of "sigmas"

$$\pm\, n_\sigma^{\mathrm{MC}} = \sqrt{2}\mathrm{erf}^{-1}(1 - 2/\mathrm{N}) = |{\mathrm{cdf}_{\mathrm{G}}}^{-1}(1/\mathrm{N})| \tag{A-2}$$

where ${\mathrm{cdf}_{\mathrm{G}}}^{-1}$ is the inverse Gaussian cumulative distribution function (CDF). For $N = N_{\text{sim}} = 10^4$, as it is the case for individual region statistics, $n_\sigma^{\mathrm{MC}} \approx 3.72$ corresponding to a probability of exceeding of 0.02%.

To derive the probability from Eq. A-1, we use linearly interpolated quantile probability within the MC probed PDF range:

$$\begin{aligned} P\big(X_{\mathbf{p},i}\big) \;\; &= \;\; Q_{lin}\big(X_{\mathbf{p},i}\big) \;\; \text{for} \\ &\qquad X_{\mathbf{p},i} \in \big(\min(X^{\mathrm{sim}}_{\mathbf{p},i}), \max(X^{\mathrm{sim}}_{\mathbf{p},i})\big) \end{aligned} \tag{A-3}$$

where Q_{lin} is the linearly interpolated quantile probability, and $\min(X^{\mathrm{sim}}_{\mathbf{p},i})$ and $\max(X^{\mathrm{sim}}_{\mathbf{p},i})$ are the minimal and maximal $X_{\mathbf{p},i}$ values obtained from a sample of N_{sim} simulations that probe the underlying $X_{\mathbf{p},i}$ PDF. Outside the probed range ($X_{\mathbf{p},i} \in \big(-\infty, \min(X^{\mathrm{sim}}_{\mathbf{p},i})\big] \cup$

$\big[\max(X^{\rm sim}_{\mathbf{p},i}),\infty\big)$) we extrapolate using a Gaussian distribution of the form

$$P\big(X_{\mathbf{p},i}\big) = \Big(1-\mathrm{erf}\big(\tfrac{\mathrm{n}_\sigma}{\sqrt{2}}\big)\Big)$$
$$n_\sigma = \begin{cases} n_\sigma^{\rm MC}\big(1+\frac{\min(X^{\rm sim}_{\mathbf{p},i})-X_{\mathbf{p},i}}{|\min(X^{\rm sim}_{\mathbf{p},i})-Q_{2_{\mathbf{p},i}}|}\big) & ; \quad X_{\mathbf{p},i} \le \min(X^{\rm sim}_{\mathbf{p},i}) \\ \\ n_\sigma^{\rm MC}\big(1+\frac{X_{\mathbf{p},i}-\max(X^{\rm sim}_{\mathbf{p},i})}{|\max(X^{\rm sim}_{\mathbf{p},i})-Q_{2_{\mathbf{p},i}}|}\big) & ; \quad X_{\mathbf{p},i} \ge \max(X^{\rm sim}_{\mathbf{p},i}) \end{cases} \tag{A-4}$$

Note that the extrapolation form is connected to the MC probed PDF range by the matching condition $P(X_{\mathbf{p},i}(n_\sigma^{\rm MC})) = P(\min(X^{\rm sim}_{\mathbf{p},i})) = P(\max(X^{\rm sim}_{\mathbf{p},i})) = 2/N$ – i.e. the probability of finding a $X_{\mathbf{p},i}$ value anywhere in range $(-\infty,\min(X^{\rm sim}_{\mathbf{p},i})] \cup [\max(X^{\rm sim}_{\mathbf{p},i}),\infty)$. An example of this extrapolation is shown in Fig. A-3. This extrapolation is obviously not validated for MODs higher than the mean, but we use it as a guide for very low probability events. Note that all our results with a lower significance (roughly outside 3σ CL) are obtained modulo this approximation.

A-2.2 Multi-region joint analysis

In the multi-region analysis, we account for all correlations between regions of a given *multi-mask* using an unbiased estimator of the full covariance matrix. Using the same parameter vector $\mathbf{p} = \{r, m, d\}$, we define a one column vector for each MOD ($\mathbf{X_p} \in \{\mathrm{m_p}, \sigma_\mathbf{p}, \mathrm{S_p}, \mathrm{K_p}\}$) of size $N_{reg}(r,m)$ such that $\mathbf{X_p} = \big(X_{r,m,d,i=1},...,X_{r,m,d,i=N_{reg}(r,m)}\big)^T$ contain $X_{\mathbf{p},i}$ values for each region of a given *multi-mask* m of given pixelization scheme r and data-set $d \in \{\mathrm{Q,V,W,INC}\}$. Introducing a parameter vector $\mathbf{q} = \{X, r, m, d\}$ we define a corresponding $\chi^2_\mathbf{q}$ value as:

$$\chi^2_\mathbf{q} = (\mathbf{X}^{\rm data}_\mathbf{p} - \langle \mathbf{X}^{\rm sim}_\mathbf{p} \rangle)^T \mathbf{C}^{-\mathbf{1}}_\mathbf{q} (\mathbf{X}^{\rm data}_\mathbf{p} - \langle \mathbf{X}^{\rm sim}_\mathbf{p} \rangle) \tag{A-5}$$

where the $\langle\rangle$ is the average $\mathbf{X}$ from $N_{\mathrm{sim(C_q)}}$ simulations and $\mathbf{C}^{-\mathbf{1}}_\mathbf{q}$ is an unbiased estimator of the inverse covariance matrix (Hartlap et al., 2007) calculated from $N_{\mathrm{sim(C_q)}}$ simulations and is given by:

$$(\mathbf{C}^{-\mathbf{1}}_\mathbf{q})_{ii'} = \frac{N_{\mathrm{sim(C_q)}} - N_{reg}(r,m) - 2}{N_{\mathrm{sim(C_q)}} - 1} (\tilde{\mathbf{C}}^{-1}_\mathbf{q})_{ii'} \tag{A-6}$$

and $\tilde{\mathbf{C}}^{-1}_\mathbf{p}$ is the inverse covariance matrix.

In Fig: A-4 we show the convergence $\chi^2_\mathbf{q}$ with the number of simulations used to calculate the covariance matrix, $N_{\mathrm{sim(C_q)}}$, for all six types of pixelization schemes. The biasing of the $\chi^2_\mathbf{q}$ values with regards to the ideal $\chi^2_\mathbf{q}(N_{\mathrm{sim(C_q)}} = \infty)$ value is caused by the limited

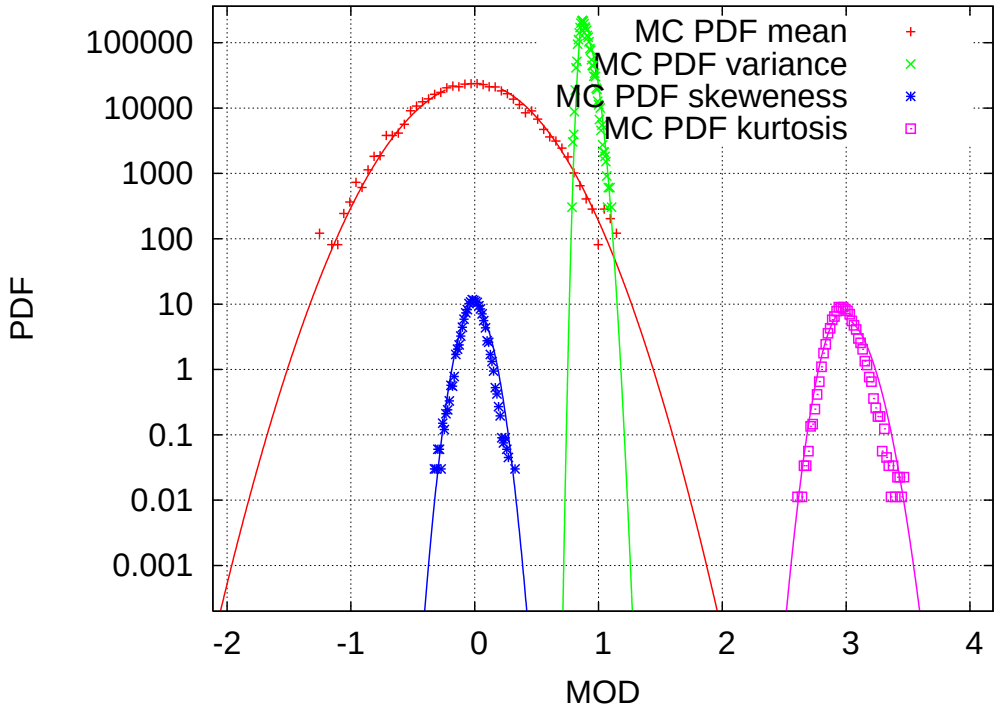


Fig A-3: Example of extrapolation (solid lines) used to compute the distribution of 10^4 MODs from one of the regions in one of the *multi-masks*. Mean and variance MODs values were multiplied by a factor of 10^4 for the sake of clarity. Note that the Gaussian extrapolation (used only outside MC probed PDF region) also interpolates the data quite well in case of the means (red crosses).

number of available simulations. As expected, for a given number of simulations, this bias decreases with the effective number of DOF and as such with the *multi-mask* resolution. We account for this biasing by using simulations to probe the underlying PDF of the $\chi^2_{\mathbf{q}}$ values, instead of using theoretical χ^2 distributions. In fact, it would not be valid to use theoretical χ^2 PDFs, since the distributions of individual MODs (except for the mean) are not Gaussian. As such, we will probe the underlying distributions using $N_{\mathrm{simPDF}(\chi^2_{\mathbf{q}})} = N_{\mathrm{sim}} - N_{\mathrm{sim}(\mathrm{C}_{\mathbf{q}})}$ simulations, for each MOD and for each *multi-mask*.

In the case of the joint multi-region statistics we use $N = N_{\mathrm{simPDF}(\chi^2_{\mathbf{q}})} = 10^3$ simulations corresponding to $n_\sigma^{\mathrm{MC}} \approx 3.09$, which gives a probability of exceeding and 0.2%. The remaining $N_{\mathrm{sim}(\mathrm{C}_{\mathbf{q}})} = 9 \cdot 10^3$ simulations are used for covariance matrix calculation.

Given a $(\chi^2_{\mathbf{q}})^{\mathrm{data}}$ value we define a corresponding joint probability as:

$$P\big(\chi^2_{\mathbf{q}}\big) \equiv 1 - P_q\big((\chi^2_{\mathbf{q}})^{\mathrm{sim}} > (\chi^2_{\mathbf{q}})^{\mathrm{data}}\big) \tag{A-7}$$

We use the same formulas for inter/extrapolation as described in Sect. A-2.1 by replacing

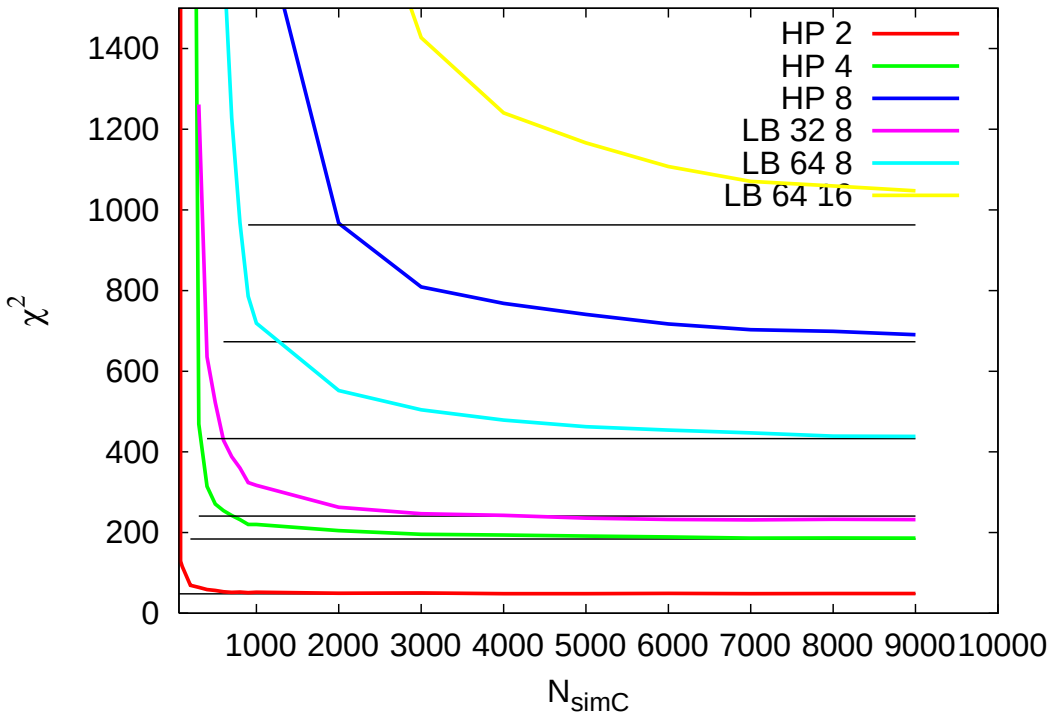


Fig A-4: Convergence of $\chi^2_{\mathbf{q}}$ values as a function of the number of simulation used for the covariance matrix calculation for a given MOD. Each color corresponds to a type of pixelization schemes. The number of regions increases from bottom (48, for *HP* 2) to top (1024 for *LB* 64 16) (see. Table 3.1). Horizontal lines indicate the effective numbers of degrees of freedom for that *multi-mask* (i.e. number of unmasked regions due to Kp03 sky mask). These χ^2 corresponds to the the first MOD only (mean) but other MODs exhibit similar dependence.

$X_{\mathbf{p},i}$ with $\chi^2_{\mathbf{q}}$.

We note that, in fact, in case of the multi-region analysis it doesn't make any difference whether the derived χ^2 values are statistically debiased or not, since exactly the same biasing affects the simulated χ^2 values used to probe the underlying PDF. We also find that the convergence of such derived probabilities is actually much better than one could infer when looking at the worst case of *LB* 64 16 presented in Fig. A-4. We estimate that the derived probabilities converge to their true values within $\sim 10\%$ of that value in this worst case of *LB* 64 16 pixelization scheme, while in case of the *HP* 2 the convergence of the derived probabilities is $\lesssim 2\%$.

The statistical debasing used for the χ^2 values matters however in our third-stage analysis, i.e. in case of the all-*multi-masks* analysis.

A-2.3 All *Multi-masks* analysis

There is no a unique way to generalize from the results of the multi-region analysis. Although the most straightforward way would be to compute the inverse covariance matrix between all the MODs and for all regions of all *multi-masks*, this turns out to be computationally not feasible.

Note the fact that the $\chi^2_{\mathbf{q}}$ values are partially correlated with each other, since they result from a different sampling of the same data-set. However the degree of correlation strongly depends on the particular *multi-mask* properties and resolutions. In particular the inter-*multi-mask* correlations are largest in the lowest resolution *multi-masks*. The smaller correlations between various *multi-masks*, the more additional information the *multi-mask* analysis explores about the data-set. Large inter-*multi-mask* correlations indicate that not much new information is gained making it unnecessary to process large number of *multi-masks*.

In the following, in order the integrate all the information probed by different *multi-masks* we calculate the cumulative probability of rejecting the GRF hypothesis using the median distribution $\overline{\varphi}(\chi^2_{\mathbf{q}\mathrm{MC}})$ of all processed $\chi^2_{\mathbf{q}\mathrm{MC}}$ distributions and the median $\overline{\chi^2_{\mathbf{q}}}$ value of the data. Therefore we calculate the distribution and $\chi^2_{\mathbf{q}}$ value as:

$$\begin{aligned} \overline{\varphi}(\chi^2_{\mathbf{q}\mathrm{MC}}) &= \langle \varphi(\chi^2_{\mathbf{q}\mathrm{MC}}/\mathrm{DOF}_{\mathrm{eff}})\rangle_m \\ \overline{\chi^2_{\mathbf{q}}} &= \langle \chi^2_{\mathbf{q}}/\mathrm{DOF}_{\mathrm{eff}}\rangle_m \end{aligned} \tag{A-8}$$

where $\langle\rangle_{\mathrm{m}}$ is the average over all *multi-masks*, $\mathrm{DOF}_{\mathrm{eff}} = \mathrm{DOF}(r, m)$ is the effective number of degrees of freedom which is a function of resolution r, of the pixelization scheme and of a particular *multi-mask* due to interplay with the sky cut.

We define the joint cumulative probability of rejecting the GRF hypothesis of the WMAP CMB data via inconsistency with our simulations as a function of $\mathbf{q}$ analogically as in Eq. A-7, and we use the same extrapolation and interpolation formula as in case of multi-region analysis.

Table A-1: List of principal acronyms used in this section and in the main body of section 3, briefly summarized for quick reference. $\star$ - indicates "unless specified otherwise".

Symbol	explanation
MOD	moment of distribution
data	upper index to indicate a measurement on data
sim	upper index to indicate a measurement on simulations
N_m	total number of *multi-masks* in a given pixelization scheme ($N_m = 100^\star$)
m	*multi-mask* index number $m \in \{1, \ldots, N_m\}$
r	pixelization scheme resolution parameter $r \in \{48, 192, 256, 512, 768, 1024\}$ (Tab. 3.1)
X	MOD: $X \in \{\mathrm{m}, \sigma, \mathrm{S}, \mathrm{K}\}$, mean, variance, skewness, kurtosis respectively
d	data-set: $d \in \{\mathrm{Q}, \mathrm{V}, \mathrm{W}, \mathrm{INC}\}$
$N_{reg}(r, m)$	number of regions in m'th *multi-mask* of the pixelization scheme r
i	region index within a *multi-mask*
$\mathbf{p}$	parameter vector: $\mathbf{p} = \{r, m, d\}$
$\mathbf{q}$	parameter vector: $\mathbf{q} = \{X, r, m, d\}$
$X_{\mathbf{p},i}$	value of a MOD for i'th region of m'th *multi-mask* of r'th pixelization scheme in a data-set d
N_{sim}	total number of simulations; number of simulations used in the single-region analysis
$P\big(X_{\mathbf{p},i}\big)$	probability corresponding to $X_{\mathbf{p},i}$ derived from N_{sim} simulations (Eq. A-1)
n_σ^{MC}	defined in Eq. A-2
$\mathbf{X_p}$	vector of MOD values for for a given value of parameter $\mathbf{p}$
$N_{\mathrm{sim}(\mathrm{C_q})}$	number of GRF simulations used to derive the covariance matrix in multi-region analysis ($N_{\mathrm{sim}(\mathrm{C_q})} = 9\,000^\star$)
$N_{\mathrm{simPDF}(\chi^2_\mathbf{q})}$	number of GRF simulations used to probe the distribution of $\chi^2_\mathbf{q}$ values in multi-region analysis ($N_{\mathrm{simPDF}(\chi^2_\mathbf{q})} = 1\,000^\star$)
$\chi^2_\mathbf{q}$	χ^2 value for the corresponding $\mathbf{X}_\mathbf{p}^{\mathrm{data}}$ (Eq. A-5)
$P\big(\chi^2_\mathbf{q}\big)$	probability corresponding to $\chi^2_\mathbf{q}$ (Eq. A-7)

A-3 Noise simulations tests

Difference maps obtained from observations in nearly the same frequency, and with nearly the same beams profile provide a good opportunity to measure the statistical properties of the instrumental noise.

We have performed a reduced χ^2 tests, directly in pixel domain, of the difference maps obtained from different channels of the WMAP data (Q12, V12, QV and Q1V1) at the Healpix resolution $n_s = 512$, and compared with results of the same tests performed at a low Healpix resolution $n_s = 4$.

Since the covariance matrix of the noise realizations is well diagonal a single variate Gaussian statistics was assumed, and reduced χ^2 distributions used. The Q12 and V12 yielded a well consistency with the simulations at both resolutions. The QV and Q1V1 difference data however, turns out to be more troublesome. Whereas there is a good consistency at high resolution, the low resolution reduced χ^2 tests show significant discrepancy yielding a "probability of rejecting" $P = 0.999963$ in case of QV map and $P = 0.998$ in case of Q1V1 map. This result is also discussed in Sect. 3.6.5 in light of the anomalous dipole component of the V band map.

We also performed a single-region, joint multi-region, and all *multi-masks* analyses on the Q12 difference map, using a subset of 10 selected *multi-masks* of the *HP* 2 pixelization scheme. Since the low resolution analysis yields a quick convergence (Fig. A-4 in Sect. A-2.2) a small number of 500 simulations were generated and half of them used for covariance matrix calculation, and the other half was used for probing distributions of the χ^2 values.

We found strong anomalies in the distribution of means (of which joint probability is well extrapolated using Eq. A-4 outside the MC probed range (Fig. A-3)). The variance of the scrambled Δ maps show that the rms amplitude of the differences is limited to the 1.7μK at these scales which is consistent with the limits to the residual systematical uncertainties in Q1 and Q2 channels of the WMAP (Hinshaw et al., 2003a) at these scales. The constraint includes not only the systematical uncertainties but also possible differences due to uncorrelated white noise used in our simulations, which in principle in the regional statistics do not average out in the same way as the pre-whitened 1/f pink noise of the WMAP data.

A difference Δ map of the variances can also serve as a rough estimate of the level of

local systematical effects. Anomalies in the scrambled map are indeed found, with strongest deviations concentrated in parts of regions adjacent to the Galactic Center, with extreme values $< 3\mu$K. However a close orientation of the regions to the Galactic Center is more likely a hint on the residual foregrounds contamination, due to slight differences in the effective frequency of the Q1 and Q2 differential assemblies, as well as in the beam profiles, rather than a manifestation of a systematical effects. Due to this leakage the limits to the aforementioned systematical effects at the level of 1.7μK should be considered as an upper limits.

A-4 Symmetries of the spherical harmonics

In chapter 6 we argued that the zigzag shape, apparent in the reconstructed angular CMB power spectrum would necessarily induce a point symmetries, alternatively anti-symmetries in the sky due to the symmetries in realizations of Gaussian random fields synthesized only out of a single multipole. This feature is easily seen by noticing that the synthesis of a temperature field $T(\pi-\theta,\phi+\pi)$ yields:

$$\begin{aligned}
T(-\hat{\mathbf{n}}) &= T(\pi-\theta,\phi+\pi)\\
&= a_{00}N_{00}e^{\mathbf{i}(0)(\phi+\pi)}P_0^0(\cos(\pi-\theta)) + a_{1-1}N_{1-1}e^{\mathbf{i}(-1)(\phi+\pi)}P_1^{-1}(\cos(\pi-\theta))\\
&+ a_{10}N_{10}e^{\mathbf{i}(0)(\phi+\pi)}P_1^0(\cos(\pi-\theta)) + a_{11}N_{11}e^{\mathbf{i}(1)(\phi+\pi)}P_1^1(\cos(\pi-\theta)) + \cdots\\
&= (-1)a_{00}N_{00}e^{\mathbf{i}(0)(\phi)}P_0^0(\cos(-\theta)) + (-1)a_{1-1}N_{1-1}e^{\mathbf{i}(-1)(\phi)}P_1^{-1}(\cos(-\theta))\\
&+ (-1)a_{10}N_{10}e^{\mathbf{i}(0)(\phi)}P_1^0(\cos(-\theta)) + (-1)a_{11}N_{11}e^{\mathbf{i}(1)(\phi)}P_1^1(\cos(-\theta)) + \cdots\\
&= (-1)(-1)^0a_{00}N_{00}e^{\mathbf{i}(0)(\phi)}P_0^0(\cos(\theta)) + (-1)(-1)^0a_{1-1}N_{1-1}e^{\mathbf{i}(-1)(\phi)}P_1^{-1}(\cos(\theta))\\
&+ (-1)(-1)^1a_{10}N_{10}e^{\mathbf{i}(0)(\phi)}P_1^0(\cos(\theta)) + (-1)(-1)^2a_{11}N_{11}e^{\mathbf{i}(1)(\phi)}P_1^1(\cos(\theta)) + \cdots\\
&= \begin{cases} T(\hat{\mathbf{n}}) = T(\theta,\phi) & \text{for} \quad \ell\in\{0,2,4,\cdots\}\wedge a_{\ell+1,m}=0\\ -T(\hat{\mathbf{n}}) = -T(\theta,\phi) & \text{for} \quad \ell\in\{1,3,5,\cdots\}\wedge a_{\ell-1,m}=0\end{cases}
\end{aligned} \tag{A-9}$$

where we introduced $N_{\ell m}=\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}$ to denote the normalization coefficients of the spherical harmonics, and where the $P_{\ell m}(\cos\theta)$ denote the associated Legendre polynomials.

Similarly, the plane symmetry is realized only if the only non-vanishing $a_{\ell m}$ coefficients are those for which the $\ell+m$ is even.

$$\begin{aligned} T(\pi-\theta,\phi) &= (-1)^0 a_{00} N_{00} e^{\mathbf{i}(0)(\phi)} P_0^0(\cos(\theta)) + (-1)^0 a_{1-1} N_{1-1} e^{\mathbf{i}(-1)(\phi)} P_1^{-1}(\cos(\theta)) \\ &+ (-1)^1 a_{10} N_{10} e^{\mathbf{i}(0)(\phi)} P_1^0(\cos(\theta)) + (-1)^2 a_{11} N_{11} e^{\mathbf{i}(1)(\phi)} P_1^1(\cos(\theta)) + \cdots \\ &= T(\theta,\phi) \Leftrightarrow \bigwedge_{\ell+m} (\ell+m+1)/2 \in \mathbb{Z},\ a_{\ell m} = 0 \end{aligned} \tag{A-10}$$

where $\mathbb{Z}$ denote the set of integer numbers. However note that zeroing $a_{\ell m}$ coefficients that yield odd $\ell + m$ removes power from both: even and odd multipoles.

A-5 Minkowski Functionals

Following the derivation of Schmalzing & Gorski (1998), for a temperature field $f \equiv T(\hat{\mathbf{n}})$, defined over the surface of the sphere, and for the field temperature threshold ν, we calculate the jth Minkowski functional v_j, as:

$$v_j(f,\nu) = \frac{1}{4\pi} \int_{\mathbb{S}^2} \mathcal{I}_j(f,\nu) da \tag{A-11}$$

where

$$\begin{aligned} \mathcal{I}_0(f,\nu) &= \Theta(f-\nu) \\ \mathcal{I}_1(f,\nu) &= \tfrac{1}{4}\delta(f-\nu)\sqrt{f_{;\phi}^2 + f_{;\theta}^2} \\ \mathcal{I}_2(f,\nu) &= \tfrac{1}{2\pi}\delta(f-\nu)\frac{2f_{;\phi}f_{;\theta}f_{;\theta\phi} - f_{;\phi}^2 f_{;\theta\theta} - f_{;\theta}^2 f_{;\phi\phi}}{f_{;\phi}^2 + f_{;\theta}^2} \end{aligned} \tag{A-12}$$

where $\Theta(f-\nu)$ is the Heaviside step function, and $\delta(f-\nu) = \frac{1}{\Delta_\nu} w(f-\nu)$ with

$$w(f-\nu) = \begin{cases} 0, & f \notin [\nu - \Delta_\nu/2, \nu + \Delta_\nu/2), \\ 1, & f \in [\nu - \Delta_\nu/2, \nu + \Delta_\nu/2) \end{cases} \tag{A-13}$$

where Δ_ν defines the distance between the neighbouring thresholds. The ";" denotes the covariant derivative with respect to the spherical coordinates θ or ϕ. For the sphere, these derivatives yield: $f_{;\phi} = \frac{1}{\sin\theta} f_{,\phi}$ and $f_{;\theta} = f_{,\theta}$, where "," denote the partial derivatives.

The derivatives in longitudinal direction are calculated utilizing the iso-latitude, ring pixels orientation of the Healpix pixelization scheme, while in the latitudinal direction the closest matched pixel is used. In order to account for the edge effects from the KQ75 cut-sky, we neglect the regions that are directly adjacent to the sky mask. This allows to utilize as much of the map sky as possible - in contrast to the methods based on the harmonic space based derivative calculations.