UNIVERSITY OF HELSINKI REPORT SERIES IN PHYSICS

HU-P-D165

Measuring the Cosmic Microwave Background
—
Topics along the analysis pipeline

Reijo Keskitalo

Helsinki Institute of Physics and
Division of Elementary Particle Physics
Department of Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

ACADEMIC DISSERTATION
To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in the Small Auditorium (E204) of Physicum, Gustaf Hällströmin katu 2a, on Tuesday, June 30, 2009 at 12 o'clock.

Helsinki 2009

ISBN 978-952-10-4241-6 (printed version)
ISSN 0356-0961
ISBN 978-952-10-4242-3 (pdf-version)
http://ethesis.helsinki.fi
Yliopistopaino
Helsinki 2009

Preface

This thesis is based on work carried out in the University of Helsinki Department of Physics and Helsinki Institute of Physics during the years 2006–2009. It was made possible by the irreplaceable financial support of the Jenny and Antti Wihuri Foundation to which I wish to extend my grateful regards. I also acknowledge the support of the Academy of Finland and the Väisälä foundation.

First and foremost, I must thank my supervisor, Dr. Hannu Kurki-Suonio. Receiving education and guidance from such an intelligent, educated person has been a privilege. I only wish some part of his talent for minute details has transferred over these years. Second, I want to thank Dr. Torsti Poutanen, whom I consider my second supervisor, for setting an example of an open-minded and practical scientist. From him I have learned the mentality of focusing to the relevant and adopting fearlessly new methods and tools.

Our team at the university would be much lesser if there wasn't for Dr. Elina Keihänen. Without her commitment to CMB map-making and her Madam code, this thesis would likely be of a very different subject. Coincidentally, she has taught me a great deal about scientific programming through the code and our discussions. Thank you, Elina. Similarly, our team would not be complete without Anna-Stiina Sirviö. Watching Ansku absorb cosmological knowledge at an alarming rate has been a source of inspiration to all of us. It has been a pleasure sharing the office with her and Elina.

My thanks go to professors Keijo Kajantie and Kari Enqvist for their inspiring examples. Keijo's course on many body physics and Kari's course on cosmology have both had a substantial impact on my view of the world.

Many of my friends and colleagues at the University of Helsinki not mentioned yet have made my time there enjoyable. I want to thank, in no particular order, Aleksi, Jussi, Vesa, Sami, Tomi, Jens, Heikki and everyone else at the Department of Physics and the Institute of Physics.

Last, but certainly far from the least, I thank, from the bottom of my heart, my beloved wife, Petra, who has expressed the patience of an angel while I have pursued this enterprise.

June, 2009
Reijo Keskitalo

R. Keskitalo: Measuring the Cosmic Microwave Background – Topics along the analysis pipeline. University of Helsinki, 2009, 58 p. University of Helsinki, Report Series in Physics, HU-P-D165, ISSN 0356-0961, ISBN 978-952-10-4241-6 (printed version), ISBN 978-952-10-4242-3 (pdf version).
INSPEC classification: A9575P, A9870V, A9880L
Keywords: Cosmology, early universe, cosmic microwave background radiation, CMB theory, data analysis, isocurvature, parameter estimation, Markov chain Monte Carlo

Abstract

The first quarter of the 20^{th} century witnessed a rebirth of cosmology, study of our Universe, as a field of scientific research with testable theoretical predictions. The amount of available cosmological data grew slowly from a few galaxy redshift measurements, rotation curves and local light element abundances into the first detection of the cosmic microwave background (CMB) in 1965. By the turn of the century the amount of data exploded incorporating fields of new, exciting cosmological observables such as lensing, Lyman alpha forests, type Ia supernovae, baryon acoustic oscillations and Sunyaev-Zeldovich regions to name a few.

CMB, the ubiquitous afterglow of the Big Bang, carries with it a wealth of cosmological information. Unfortunately, that information, delicate intensity variations, turned out hard to extract from the overall temperature. Since the first detection, it took nearly 30 years before first evidence of fluctuations on the microwave background were presented. At present, high precision cosmology is solidly based on precise measurements of the CMB anisotropy making it possible to pinpoint cosmological parameters to one-in-a-hundred level precision. The progress has made it possible to build and test models of the Universe that differ in the way the cosmos evolved some fraction of the first second since the Big Bang.

This thesis is concerned with the high precision CMB observations. It presents three selected topics along a CMB experiment analysis pipeline. Map-making and residual noise estimation are studied using an approach called *destriping*. The studied approximate methods are invaluable for the large datasets of any modern CMB experiment and will undoubtedly become even more so when the next generation of experiments reach the operational stage.

We begin with a brief overview of cosmological observations and describe the general relativistic perturbation theory. Next we discuss the map-making problem of a CMB experiment and the characterization of residual noise present in the maps. In the end, the use of modern cosmological data is presented in the study of an extended cosmological model, the correlated isocurvature fluctuations. Current available data is shown to indicate that future experiments are certainly needed to provide more information on these extra degrees of freedom. Any solid evidence of the isocurvature modes would have a considerable impact due to their power in model selection.

List of included publications

This PhD thesis contains an introductory part describing the field of CMB measurements and data analysis and three research articles:

I H. Kurki-Suonio et al.,
Destriping CMB Temperature and Polarization Maps,
Submitted to Astron. Astrophys. [arXiv:0904.3623]

II R. Keskitalo et al.,
Residual Noise Covariance for PLANCK *low resolution data analysis*,
Submitted to Astron. Astrophys. [arXiv:0906.0175]

III R. Keskitalo, H. Kurki-Suonio, V. Muhonen and J. Väliviita,
Hints of Isocurvature Perturbations in the Cosmic Microwave Background?,
J. Cosmol. Astropart. Phys. **0709**, 008 (2006) [arXiv:astro-ph/0611917]

Author's contribution

Paper I: This paper is a detailed study of the destriping principle. The linearity of destriping is exploited to track the propagation of various signals from time ordered data into maps. I simulated the test data, ran and modified the destriping code and processed the outputs using tools both available and written by myself. I put together an early draft version of the paper and assisted in the subsequent analysis.

Paper II: This paper presents three residual noise covariance matrices for three different map-makers. It is a collaboration paper between all the code developers. I wrote and tested the code corresponding to the generalized destriping principle. I assumed the responsibility for the progress, participated in choosing the test cases and validation procedures, analyzed the outputs from all the teams, produced the plots and wrote most of the paper.

Paper III: This paper contains a parameter estimation study using the WMAP 3$^{\text{rd}}$ year and other cosmological data. It was a continuation of the earlier work by the rest of the authors. I reprocessed their old Markov chains using a self-modified version of COSMOMC, then ran independent Markov chains for analysis using different datasets, processed the results, explored parametrization priors and produced the plots.

Contents

Preface i
Abstract ii
List of included publications iii
Author's contribution iii

1 Background **1**
1.1 Our Universe 1
1.2 Cosmic Microwave Background (CMB) 6
1.3 Other cosmological data 9
1.4 Cosmological perturbation theory 12
1.5 Planck Surveyor mission 15
1.6 Notational conventions 20

2 Destriping approach to map-making **21**
2.1 Preliminaries 21
2.2 Maximum likelihood map-making 22
2.3 Destriping 23

3 Residual noise characterization **28**
3.1 Preliminaries 28
3.2 To noise covariance from its inverse 29
3.3 Structure of the noise covariance 30
3.4 Downgrading maps and matrices 30
3.5 Making use of residual noise covariance 33

4 Isocurvature models **36**
4.1 Cosmological parameter estimation 36
4.2 Isocurvature degrees of freedom 39
4.3 Results and implications for cosmology 42

5 Conclusions **44**

List of Figures

1.1 Evolution of a cosmological scale 5
1.2 WMAP 5yr best fit 7
1.3 Angular power spectra 9
1.4 Frequency coverage 15
1.5 Noise model 16
1.6 Cycloidal scanning 18
1.7 PLANCK analysis pipeline 19

2.1 Component powers in spectrum domain 26

3.1 Noise covariance patterns 31
3.2 Noise covariance patterns II 32

4.1 Parametrization effects 41
4.2 Posterior likelihood for isocurvature fraction 43

Chapter 1

Background

1.1 Our Universe

1.1.1 Big Bang cosmology

Our Universe is expanding. Expanding in the sense that the average separation between two distant observers increases. This has been an experimental fact since the late 1920's [1, 2]. By that time it was already well known from the redshift measurements of other galaxies that they were receding from us in all directions. The real breakthrough, however, was to estimate the distances of the galaxies using particular variable stars called Cepheids whose absolute brightness is known to high accuracy. Comparison to the redshift data revealed a relation between a galaxy's radial velocity, v, and distance, d:

$$v = H_0\, d \tag{1.1}$$

This proportionality factor is called the *Hubble constant*, H_0. By an invocation of the Cosmological Principle[1], the relation is interpreted as expansion of the Universe

General Theory of Relativity [3] had already been applied to cosmology in 1917 [4]. Only at that time Einstein considered the Universe to be static. The theory itself is well equipped to describe the expansion of an isotropic, spatially homogeneous Universe. The description is known as the Friedmann-Lemaître-Robertson-Walker solution [5, 6, 7, 8, 9]. In this setting the line element of space-time can be written in spherical coordinates as:

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = -c^2\mathrm{d}t^2 + a(t)^2\left(\frac{\mathrm{d}r^2}{1-Kr^2} + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2\right), \tag{1.2}$$

where $g_{\mu\nu}$ is the metric, c is the speed of light, $a(t)$ is the scale factor and K is the curvature constant of the Universe. The isotropic, homogeneous approximation reduces

[1] Cosmological Principle states that at the very largest scales, the Universe is homogeneous and isotropic. It follows that our position in the Universe is not special.

the Einstein equation to the Friedmann equations that in one of their many forms read:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} \tag{1.3}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right). \tag{1.4}$$

Here, G is the Newton's gravitational constant, ρ and p are the density and the pressure of the Universe and a dot over a quantity denotes the derivative with respect to the cosmic time. In this context the scale factor can be interpreted as a dimensionless measure of separation between two observers. Changes in the separation are driven by the energy density and the pressure that are due to various components of the cosmic fluid, e.g. radiation ($p = \frac{1}{3}\rho$), non-relativistic matter ($p = 0$) and later to be defined scalar fields. All the components differ, not just by their equations of state, $p(\rho)$, but also by the way their density connects with the evolution of the scale factor.

The two Friedmann equations are often accompanied with the cosmic fluid energy-momentum continuity equation that also follows from the Einstein equation:

$$\dot{\rho} = -3\left(\rho + \frac{p}{c^2}\right)\frac{\dot{a}}{a}. \tag{1.5}$$

Along with the equations of state, the continuity equation tells us that matter and radiation densities are proportional to a^{-3} and a^{-4} respectively. Furthermore, we can use these results and the first Friedmann equation (1.3) to find that, when either matter or radiation dominates, the density of the Universe, the scale factor evolves as $t^{2/3}$ and $t^{1/2}$.

An unavoidable[2] consequence of the cosmological expansion is, that the Universe must have been smaller and denser. This line of thinking leads one to consider the initial conditions that are popularly referred to as the *Big Bang*. If no ad hoc model is devised to prevent backward extrapolation of the current matter distribution to the primordial fireball, one must conclude that the Universe used to be extremely dense and hot. This model has been immensely successful in explaining the observed abundances of light elements such as hydrogen and helium [11].

Under the extreme temperature and density of the primordial Universe, photon distribution is necessarily at thermal equilibrium. Expansion of the Universe cools down the plasma and finally allows neutral matter to form. This process, known as recombination, decouples photons from the rest of the energy content of the Universe. The decoupled photons are left to traverse the Universe unimpeded, only to cool down with the expansion [12]. This residual radiation from the violent birth of our Universe is called the cosmic microwave background (CMB). It was first discovered in 1965 [13]. Existence of the CMB is deemed to be one of the strongest evidence for the big bang cosmology.

[2] If no complicated models are devised to account for the observations. A well-known example is the steady state theory [10] that avoided the singular initial condition by continuous creation of matter.

1.1.2 Concept: horizon

Cosmologists define the *horizon* as the surface of a volume of 3–dimensional space that has causal connection over cosmological time scales (most commonly relative to the age of the Universe). By defining the time-dependent Hubble parameter as a function of the scale factor:

$$H = \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{\dot{a}}{a}, \tag{1.6}$$

we define two cosmological quantities, the Hubble time and the Hubble length as

$$t_{\mathrm{H}} = H^{-1} \quad \text{and} \quad l_{\mathrm{H}} = c\, t_{\mathrm{H}} = cH^{-1}. \tag{1.7}$$

Hubble time is the time constant of cosmological expansion. If the Hubble parameter remained nearly constant we could solve equation (1.6),

$$H\mathrm{d}t = \frac{\mathrm{d}a}{a} \quad \Rightarrow \quad a(t) \approx a_0\, e^{t/t_{\mathrm{H}}}, \tag{1.8}$$

and the Universe would expand one e-fold in one Hubble time. Thus t_{H} defines a cosmological timescale. Hubble length is the physical distance light could travel during one Hubble time in a static Universe. Distances shorter than this are in causal connection over cosmological time scales.

1.1.3 Inflation

The Big Bang paradigm alone does not explain cosmological isotropy. Structure developed in regions that could not have been causally connected in an ordinary matter and radiation filled Universe, e.g. northern and southern celestial spheres, indicates that some process in the primordial Universe connects the two regions even though the standard Big Bang cosmology assumes them to be causally disconnected. A very profound example is the cosmological background radiation (discussed in detail below). It is an isotropic radiation field emanating from the early Universe, 13 billion years before present. If the standard Big Bang scenario was correct, the horizon, a causally connected patch of the CMB field, would span roughly just a few degrees of our view. Yet, the radiation field is almost exactly the same on the opposite side of our CMB sky, a hundred horizon distances away.

Another problem is that the Big Bang does not explain why our Universe's geometry is extremely close to flat. If we define *critical density* as the density of a flat ($K = 0$ in equation (1.3)) Universe,

$$\rho_{\mathrm{crit}} \equiv \frac{3H^2}{8\pi G}, \tag{1.9}$$

and a density parameter $\Omega \equiv \rho/\rho_{\mathrm{crit}}$, we can write equation (1.3) as a quantity directly describing deviation from flat geometry,

$$\Omega - 1 \equiv \Omega_{\mathrm{K}} = \frac{Kc^2}{a^2H^2}. \tag{1.10}$$

Now, if we start from the current estimate, $|\Omega_{\mathrm{K}}| \lesssim 0.01$ [14], we may ask how do those limits translate to initial conditions in the early Universe. If the Universe is filled with

ordinary matter or radiation, the scale factor evolves as $a \propto t^r$, where r is 2/3 or 1/2 for matter and radiation respectively. Thus we have

$$|\Omega_{\rm K}| = |\Omega_{\rm K}|\Big|_{t=t_0} \left(\frac{t}{t_0}\right)^{2r}, \tag{1.11}$$

and any deviation from flat geometry will promptly escalate. In fact, to conform with the observations, $|\Omega_{\rm K}|$ would need to have been less than 10^{-20} in the single second old Universe. If not utterly unimaginable, the cosmologists still consider the required initial conditions as awkwardly fine-tuned.

A third problem of the Big Bang would rise if the speculated grand unified theory (GUT) phase transition took place in our Universe. Most proposals of the GUT include production of topological defects, such as magnetic monopoles or cosmic strings. Despite our efforts, none have been observed. The unobserved defects are commonly known as relics.

To address these problems, the inflationary scenario [15, 16] was proposed as a refinement of the Big Bang model. In cosmological inflation the Universe undergoes a period of rapid, exponential expansion that disconnects causally connected regions and drives its geometry flat. While the disassociation of causal regions is intuitive, the solution to the flat geometry problem follows from $a \propto e^{Ct}$, as equation (1.11) implies

$$|\Omega_{\rm K}| \propto e^{-2Ct}, \tag{1.12}$$

exponential decay of curvature during the inflationary expansion.

In order to solve the problems associated with the Big Bang model, inflation needs to take place in the very early Universe, some fraction of a second after Big Bang. Inflation erases all information about the initial conditions before it. Required amount of expansion is so large (a factor of $\sim e^{60}$), that the seeds of structure around us must be created during or after inflation. Instead, to dissolve all relics and not create new ones, the GUT symmetry must be broken before inflation and the consequent stages may not include reheating into energies above the GUT energy scale $\gtrsim 10^{14}\,\mathrm{GeV}$.

Inflationary models differ in the way the expansion is generated but typically it is driven by a scalar field, ϕ, slowly approaching its potential minimum. The equation of motion for such a field is

$$\nabla^\mu \nabla_\mu \phi - V'(\phi) = 0 \tag{1.13}$$

and assuming the spatially homogeneous FLRW geometry provides a simplification:

$$\ddot{\phi} + 3H\dot{\phi} + V' = 0, \tag{1.14}$$

where we have written the covariant derivative in terms of the FLRW connection coefficients, $\Gamma^\alpha_{\beta\gamma}$, as

$$\nabla^\mu(\nabla_\mu\phi) = \nabla^\mu(\partial_\mu\phi) = \partial_\mu\partial^\mu\phi + \Gamma^\mu_{\mu\alpha}\partial^\alpha\phi. \tag{1.15}$$

We can read the energy density and the pressure for such a field from the energy-momentum tensor [17]:

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi) \quad \text{and} \quad p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{1.16}$$

Plugging these formulae into the second Friedmann equation (1.4):

$$\frac{\ddot{a}}{a} = -\frac{8\pi G}{3c^2}\left(\dot{\phi}^2 - V(\phi)\right) \tag{1.17}$$

makes it evident that when the potential energy dominates over the kinetic energy, the expansion accelerates.

Apart from solving the problems of the Big Bang cosmology, the inflationary scenario conveniently explains the origin of structure in the Universe. As the expansion is being driven by a field dominating the energy density of the Universe, quantum fluctuations in the very same field are extended into macroscopic scales and begin to evolve classically. These seeds of inhomogeneity then evolve into the rich multi-scale structures that surround us.

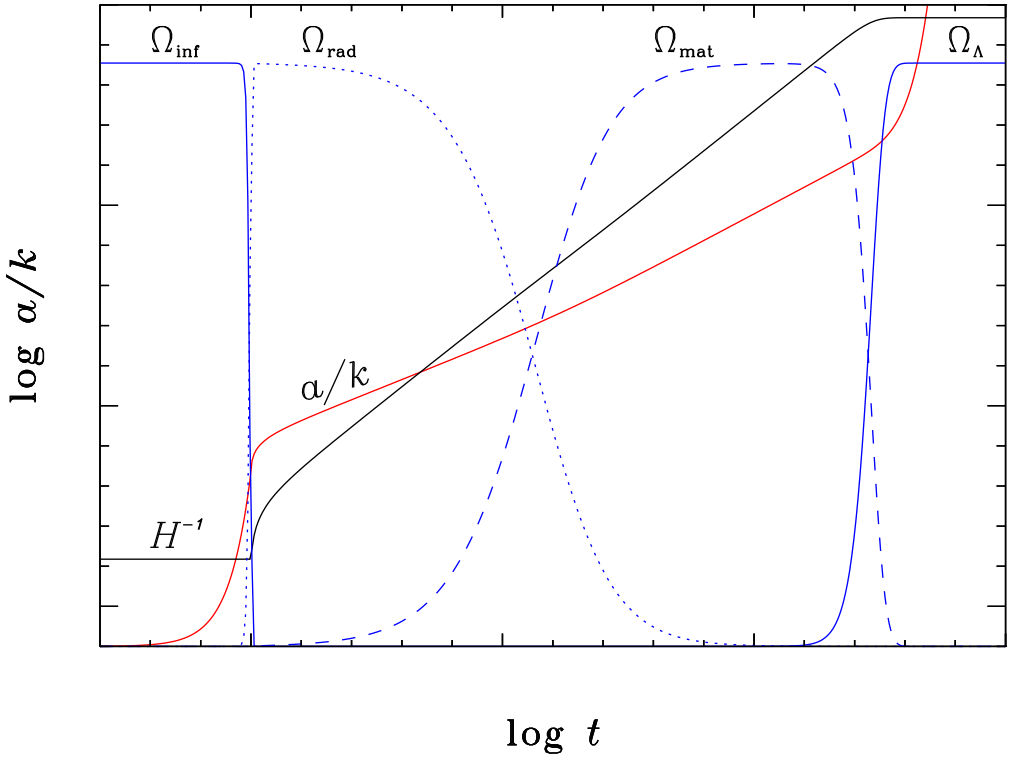


Figure 1.1: Evolution of a cosmological scale from the inflation until late times. The solid red curve is the chosen cosmological scale and the solid black curve is the Hubble horizon. The blue curves are the relative energy densities of the inflaton, radiation, matter and dark energy. The CMB photons are set free in recombination some time after the matter-radiation equality.

Figure 1.1 shows the evolution of a cosmological scale, for example a distance of two galaxy groups. It shows the relative densities of the inflaton field, radiation, matter and a possible cosmological constant, Λ. The figure depicts how the scale is driven

outside the horizon by inflationary expansion, then returns into the horizon in the radiation dominated era and finally exits the horizon again somewhere in the future. Primordial structure at this scale is evolved by dynamics of the baryon-photon plasma before recombination that takes place after matter-radiation equality at the center of the plot. Structure at the smallest scales is damped by photon diffusion while the largest scales are enhanced by clustering dark matter and propagating sound waves. These dynamics are discussed further in the following sections.

1.2 Cosmic Microwave Background (CMB)

Before the recombination the photons are strongly coupled through electromagnetic interaction to baryonic matter[3]. The coupling allows for some very interesting dynamics in the photon-baryon fluid. The primordial density fluctuations evolve first by gravitational amplification of the density differences. Then photon pressure rises to cancel the gravitational pull and leads to suppression of the same densities. Baryons are forced to follow the oscillation but they drag its equilibrium towards higher over-densities. The pattern is modified further by clustering dark matter and the evolving ratio of energy densities leading from radiation dominated into matter dominated era. Furthermore, free streaming photons smooth the energy density contrasts at scales smaller than the mean free path. These dynamics leave an imprint on the decoupled photon distribution of the CMB. The pre-recombination dynamics are often referred as *acoustic oscillations*.

We observe the CMB photons from all of the directions of the sky. Their temperature fluctuations form a two-dimensional field upon which the acoustic oscillations induce a correlation that only depends on the angle of separation between two lines of sight. If the field is Gaussian, all statistical information is conveniently expressed in a form of an *angular power spectrum*: let the CMB temperature to be field $s(\theta, \phi)$ and expand it in terms of the spherical harmonic functions as

$$s(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^{\mathrm{T}} Y_m^{\ell}(\theta, \phi). \tag{1.18}$$

Recall that for the Y_m^{ℓ}, ℓ determines the angular scale and m the orientation or pattern. Therefore a measure of correlation of the temperature field, sensitive only to the angular scale, is the *angular power spectrum*,

$$C_{\ell}^{\mathrm{TT}} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}^{\mathrm{T}}|^2, \quad \text{where} \quad \langle C_{\ell}^{\mathrm{TT}} \rangle = \langle a_{\ell m}^{\mathrm{T}*} a_{\ell m}^{\mathrm{T}} \rangle. \tag{1.19}$$

The latter expression is an ensemble average of the former. The factors C_{ℓ}^{TT} are related to correlations on the sky, $\langle s(\theta, \phi) s(\theta', \phi') \rangle$, in a manner analogous to how Fourier power spectral density of time ordered signal relates to the autocorrelation function of that

[3] Interaction is mediated by electrons through Thomson scattering. Cosmologists loosely refer to all ordinary matter (hadrons and leptons) as baryonic since the only significant contribution to the cosmological energy density of ordinary matter is on the account of baryons.

same signal: the two are each others Fourier transforms. Furthermore, we can write the angular correlation function, $\xi(\alpha)$, in terms of the $C_\ell^{\rm TT}$ as

$$\xi(\alpha) = \langle s(\mathbf{x}) s(\mathbf{x}') \rangle = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell^{\rm TT} P_\ell(\alpha), \tag{1.20}$$

where $P_\ell(\alpha)$ are the Legendre polynomials, $\mathbf{x}$ is a unit vector pointing in the direction (θ, ϕ) and α is the angle between $\mathbf{x}$ and $\mathbf{x}'$.

From the observations we know that the relative fluctuations in CMB temperature are of the order of $1 : 10^5$. This makes it possible to use linear perturbation theory to evolve [18] the primordial perturbations by perturbing the Friedmann equations (1.3) and estimate their imprint on the CMB photon distribution. Customarily, the initial perturbations are characterized by a primordial power spectrum, $\mathcal{P}_\mathcal{R}(k)$, for the curvature perturbation, $\mathcal{R}$, from the inflation. For typical inflationary models the primordial power spectrum is well approximated by a power law:

$$\mathcal{P}_\mathcal{R}(k) = A \left(\frac{k}{k_{\rm p}} \right)^{n-1}, \quad \text{where} \quad \frac{2\pi^2}{k^3} \mathcal{P}_\mathcal{R}(k) \delta(\mathbf{k} - \mathbf{k}') = \langle \mathcal{R}(\mathbf{k}) \mathcal{R}(\mathbf{k}') \rangle, \tag{1.21}$$

where A, n and $k_{\rm p}$ are constants defining the power spectrum and $\mathbf{k}$ is the Fourier wave vector of amplitude k. The primordial curvature perturbation with adiabatic initial conditions are enough to define completely the initial perturbations for all the energy species. Fig. 1.2 presents a sample angular power spectrum evolved from the adiabatic initial conditions fixed by the WMAP best fit cosmological parameters[4].

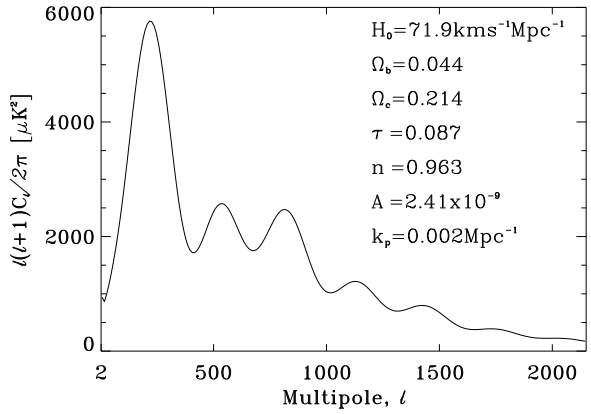


Figure 1.2: Angular power spectrum evolved from the primordial curvature perturbations assuming the WMAP 5 year data release best fit values. τ is the optical depth to reionization of the intergalactic medium.

In addition to the photon temperature anisotropy the photons exhibit also a slight polarization anisotropy that they acquire through Thomson scattering only moments prior to total decoupling from the baryon plasma. The polarization anisotropy requires quadrupole density differences between photon and baryon fluids [19, 20] that can only be supported during a narrow window near recombination. This feature follows from

[4] http://lambda.gsfc.nasa.gov/

the multipole expansion of the differential Thomson cross section:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{3}{8\pi}|\boldsymbol{\epsilon}\cdot\boldsymbol{\epsilon}'|^2\sigma_{\mathrm{T}}, \tag{1.22}$$

where $\boldsymbol{\epsilon}$ denotes photon polarization and σ_{T} is the Thomson cross section.

Given that the Thomson scattering is the only cosmological process that polarizes CMB photons, the polarization is linear. Incoming photons can thus be characterized by three Stokes parameters I, Q and U. Polarization fraction p and angle ψ relate to the Stokes parameters as

$$p = \frac{\sqrt{Q^2+U^2}}{I} \quad \text{and} \quad 2\psi = \tan^{-1}\frac{Q}{U} \tag{1.23}$$

and I is the total photon intensity, a. k. a. temperature. CMB polarization experiments are usually designed to ignore circular polarization parametrized by the Stokes parameter V, since all circular polarization is expected to originate from secondary sources.

The polarization parameters Q and U form spin$\pm$2 fields on the sky as $(Q \pm iU)$. They too can be expanded in terms of spherical base functions, but to support their spin-2 nature, we need to use spin-weighted spherical harmonics [21]:

$$(Q \pm iU)(\theta,\phi) = \sum_{\ell=2}^{\infty}\sum_{m=-\ell}^{\ell} {}_{\pm2}a_{\ell m}\,{}_{\pm2}Y_m^{\ell}(\theta,\phi). \tag{1.24}$$

Of these two fields, we can extract a curl-free and a source-free component just like the electromagnetic field is separated into E and B. Their expansions,

$$E(\theta,\phi) = \sum_{\ell=2}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell m}^{\mathrm{E}}Y_{\ell m}^{\mathrm{E}}(\theta,\phi) \quad \text{(and similarly for B)} \tag{1.25}$$

are related to the combinations $(Q \pm iU)$ by

$$a_{\ell m}^{\mathrm{E}} \equiv -({}_2a_{\ell m} + {}_{-2}a_{\ell m})/2, \quad Y_{\ell m}^{\mathrm{E}}(\theta,\phi) \equiv -\left({}_2Y_m^{\ell}(\theta,\phi) + {}_{-2}Y_m^{\ell}(\theta,\phi)\right)/2 \tag{1.26}$$

and

$$a_{\ell m}^{\mathrm{B}} \equiv -({}_2a_{\ell m} - {}_{-2}a_{\ell m})/2i, \quad Y_{\ell m}^{\mathrm{B}}(\theta,\phi) \equiv -\left({}_2Y_m^{\ell}(\theta,\phi) - {}_{-2}Y_m^{\ell}(\theta,\phi)\right)/i2. \tag{1.27}$$

Again, under the assumption that the fluctuations are Gaussian and isotropic, all statistical information can be compressed into angular power spectra. Polarized CMB information is often conveyed as TT, TE, EE, BB, TB and EB angular power spectra,

$$C_\ell^{\mathrm{XY}} = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} a_{\ell m}^{\mathrm{X}*}a_{\ell m}^{\mathrm{Y}}, \quad \text{where} \quad \langle C_\ell^{\mathrm{XY}}\rangle = \langle a_{\ell m}^{\mathrm{X}*}a_{\ell m}^{\mathrm{Y}}\rangle. \tag{1.28}$$

Note that although individual $a_{\ell m}^{\mathrm{X}}$ are complex, being expansion coefficients of real fields they satisfy $a_{\ell m}^{\mathrm{X}} = a_{\ell -m}^{\mathrm{X}*}$ and as a result, even the cross spectra are real. A sample of angular power spectra that correspond to WMAP 5 year best fit values is displayed in Figure 1.3.

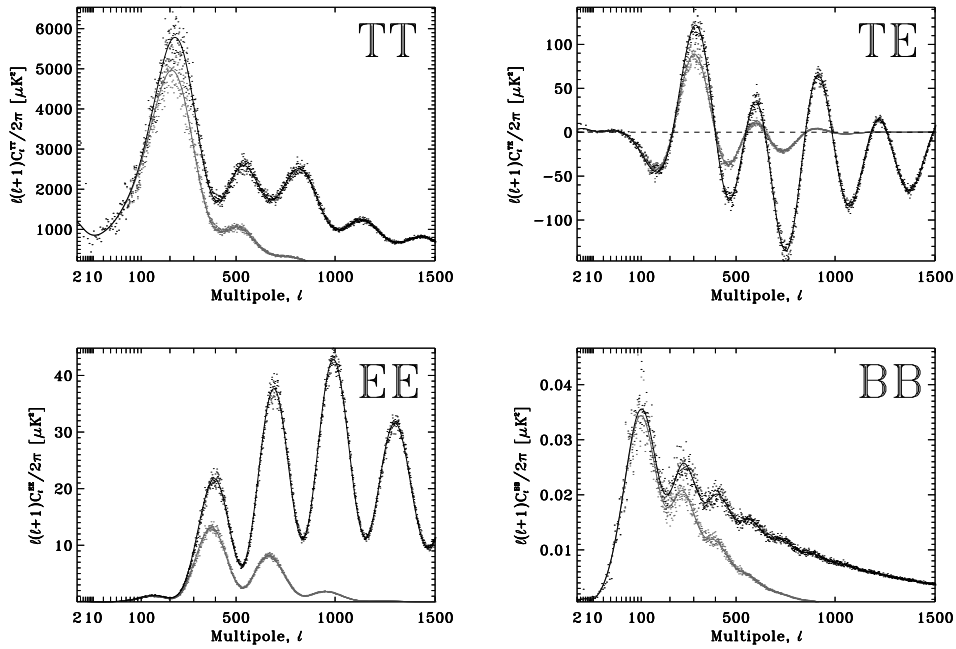


Figure 1.3: Simulated CMB angular power spectra corresponding to the WMAP 5 year best fit parameters with an allowed 10% primordial tensor perturbation responsible for the BB-mode. The solid black line is the CAMB[6] [22] simulated spectrum and the dots represent single realization. Their grey counterparts include beam smoothing corresponding to the PLANCK 70 GHz channel. Not shown here are the TB and EB power spectra that would appear through cross-polar leakage in the instrument.

1.3 Other cosmological data

As the cosmological models become more complex, the CMB data alone is unable to constrain all model parameters. This is due to parameter degeneracies, i.e. the ability of different parameter combinations to produce similar acoustic structure of the CMB photons. A notorious example is that the CMB data alone does not require dark energy at all. It is necessitated by supernova observations, the Hubble constant [23] and measurements of the large scale structure. In the following we list the most important cosmological datasets and constraints other than the CMB observations.

1.3.1 Hubble constant

The first steps to precision observational cosmology were taken in the late 1920's when the galaxy redshift and distance data were used to estimate the Hubble constant. To this end the distances of extra-galactic objects need to be determined by employing

[6] `http://camb.info`

knowledge of astrophysics. Frequently this means identifying so-called *standard candles* that possess an observable quantity related to their intrinsic luminosity. Two commonly encountered standard candles are the Cepheid variables and the type Ia supernovae. When neither are available one can resort to statistical properties of the galaxies such as the absolute magnitude of the brightest star.

Varying level of astrophysical assumptions in the H_0 estimates has kept the field controversial over decades. Even in the 1990's different approaches led to very diverge estimates between 40km/s/Mpc [24] and what has become widely accepted 72km/s/Mpc [25] (for a review of methodology cf. [26]).

1.3.2 Big bang nucleosynthesis (BBN)

The light elements, ^{2}H, ^{3}He, ^{4}He and ^{7}Li, are produced in a very short period of time when the temperature decreases sufficiently and collisions start to build elements from the free nucleons (100 keV range). Fractions of these isotopes all depend on the baryon-photon ratio. Although of considerable historic importance, the BBN limits only provide a weak prior on baryon density. For a review of the matter cf. [27]. Abundances of the light elements were forged much before the CMB, making them an interesting complementary dataset.

1.3.3 Large scale structure (LSS)

The large scale distribution of matter is measured by the distribution of galaxies. The most notable datasets are due to the 2 Degree Field Galaxy Redshift Survey[7] (2dFGRS) [28] and the Sloan Digital Sky Survey[8] (SDSS) [29]. Both measure the redshifts of galaxies, translating that into distance and characterizing the 3D distribution by appropriate statistical measures.

In the past the most frequently used statistical measure was the matter power spectrum, $P(k)$, but a number of systematic difficulties has driven cosmologists into developing more robust statistics to aid parameter estimation from galaxy distributions. New studies make increasingly use of the baryon acoustic oscillations (BAO) [30]. These oscillations are remnants of the photon acoustic oscillations of the CMB. Using the BAO one, instead of fitting a spectrum with unknown calibration and selection effects, fits the characteristic scale of the acoustic oscillations. In order to understand this quantity we must start from the *sound horizon*, r_s. We define it to be the distance perturbations in the primordial plasma could travel before recombination. The sound horizon determines the relative distance of the acoustic peaks in the angular power spectrum, as the wave modes, k, that correspond to that distance and its harmonics, impart the strongest correlations on the angular power spectrum.

When we measure the matter power spectrum from the distribution of galaxies, we find an acoustic structure imposed on top of the dominant form of the spectrum. Again, the distance of the acoustic peaks is determined by the sound horizon and we have thus

[7] `http://www2.aao.gov.au/~TDFgg/`
[8] `http://www.sdss.org/`

two angular size measurements of the same comoving size, r_s, at two very different redshifts. The way that the angular size has evolved is depends on a combination of cosmological parameters.

1.3.4 Weak Gravitational Lensing

Gravitational deflection of light by nearby clusters distorts our view of distant objects. The reader may recall an image of a notably distorted, arc-shaped galaxy. Weak lensing refers to the more subtle statistical effects that matter distribution induces to shapes and apparent sizes of distant galaxies. The same deflection field also distorts our view of the CMB and needs to be considered prior to comparing theoretical and measured spectra. Statistical measures of weak lensing are most sensitive to matter density fluctuation amplitude, σ_8, but can be studied as a source of rich cosmological data [31, 32]. For a review of the matter, see e.g. [33].

1.3.5 Type Ia supernovae

Type Ia supernovae rose to the spotlight at the end of the millennium [34]. It is believed that the event is caused by an exploding white dwarf that, after accumulating critical mass from a binary, becomes unstable at the advent of carbon fusion. The type Ia have a characteristic luminosity curve that make it possible to compute their intrinsic luminosity, $\mathcal{L}$. Intrinsic luminosity on the other hand is translated into *luminosity distance* by comparison to the observed flux, $\mathcal{F}$:

$$d_L = \left(\frac{\mathcal{L}}{4\pi\mathcal{F}}\right)^{1/2}. \tag{1.29}$$

With redshifts, the luminosity distances provide an independent measurement of the Universe's expansion history between present and distant past when the Universe was billions of years younger. The data show to high degree of confidence the transition from decelerated expansion of the matter dominated era into accelerated expansion of the dark energy dominated era [35].

1.3.6 Lyman-α forest (Lyα)

Despite their obvious differences, both the CMB and the LSS data probe the same density fluctuations. Both possess a very profound limitation in characterizing these fluctuations at the smallest observable scales. For the CMB, photon diffusion washes out the information and for the LSS the non-linear structure evolution makes it extremely hard to compare theory to observations. A way to get around the LSS limit is to observe the small scales *before* non-linear evolution confuses them. One probe of this large redshift structure are the hydrogen Lyα absorption lines in quasar spectra [36]. They probe the neutral hydrogen distribution in the intergalactic space and can resolve linear matter power spectrum at $z = 3$. Extending the range of observable scales is especially useful for the constraints of the primordial power spectrum beyond amplitude and tilt [37, 38].

1.4 Cosmological perturbation theory

Linear perturbation theory is an invaluable tool to many fields of science. We present here a short review of the perturbation theory applied to cosmology [39, 40, 41, 18]. Our focus is to outline how primordial perturbations are characterized and how their evolution can be followed into the angular power spectrum of the CMB today.

The Einstein equation, in one of its many forms,

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{1.30}$$

connects the geometry of the Universe, described by the Einstein tensor $\mathbf{G}$, to the energy content of the Universe, described by the stress-energy tensor $\mathbf{T}$. This 4×4 differential equation (the derivatives are hidden in $\mathbf{G}$) is symmetric, leaving 10 degrees of freedom.

When we model the Universe, the zeroth order (background) of equation (1.30) is assumed homogeneous and isotropic, defining the FLRW universe. Furthermore it suffices to most of our purposes to limit our considerations to spatially flat backgrounds ($K = 0$). With these assumptions the unperturbed line element reads

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = a(\eta)^2 \left(\mathrm{d}\eta^2 - \delta_{ij}\mathrm{d}x^i \mathrm{d}x^j\right), \quad \mathrm{d}\eta \equiv \frac{\mathrm{d}t}{a(t)}, \tag{1.31}$$

where we have defined the conformal time, η, that leads to a particularly simple form of the metric, $g_{\mu\nu} = a(\eta)^2 \eta_{\mu\nu}$, $\eta_{\mu\nu}$ being the Minkowski metric of Special Relativity.

In order to refine this description we assume that the real Universe can be represented as a small perturbation around our idealized (zeroth order) background. For the geometrical side this means adding a small perturbation to the metric:

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu} \equiv a(\eta)^2 \left(\eta_{\mu\nu} + h_{\mu\nu}\right). \tag{1.32}$$

From here on, the over-barred quantities refer to the background solution. The perturbation, $h_{\mu\nu}$, represents a deviation from a spatially flat background. It is not unique, but depends on our choice of the mapping between background and perturbed coordinates. The arbitrariness associated with the choice of these coordinates is referred to in cosmology as *gauge freedom*. Imposing additional constraints to $h_{\mu\nu}$ is known as choosing a gauge. The general perturbation contains more degrees of freedom than the physical set-up. Thus a cosmologist can limit the number of unphysical degrees of freedom by useful choices of the gauge.

We call a coordinate transformation between two perturbed space-times a gauge transformation. Let $\widetilde{x}^\mu$ and $\widehat{x}^\mu$ be the coordinates of a background point $\bar{x}^\mu$ in the two gauges. Difference between the two coordinate systems is $\widetilde{x}^\mu(\bar{x}^\mu) - \widehat{x}^\mu(\bar{x}^\mu) = \xi^\mu$. For coordinate systems that are sufficiently close to each other, ξ^μ defines a gauge transformation between the two systems and can be considered a tensor in the background space-time:

$$\widetilde{x}^\mu(\bar{x}^\mu) = X^\mu_\nu \widehat{x}^\nu(\bar{x}^\nu), \quad \text{where} \quad X^\mu_\nu = \frac{\partial \widetilde{x}^\mu}{\partial \widehat{x}^\nu} = \delta^\mu_\nu + \partial_\nu \xi^\mu. \tag{1.33}$$

In the following we shall see how the factors $\partial_\nu \xi^\mu$ are useful in choosing convenient gauges to work in.

A general perturbation can be decomposed into scalar, vector and tensor perturbations based on the individual part's transformation properties in spatial rotations of the background space time. We parametrize it with the scalars A and D, 3-vector $\mathbf{B}$ and a traceless 3×3 matrix $\mathbf{E}$:

$$h_{\mu\nu} = \begin{bmatrix} -2A & -\mathbf{B}^{\mathrm{T}} \\ -\mathbf{B} & -2D \cdot \mathbb{1}_{3\times 3} + 2\mathbf{E} \end{bmatrix} \tag{1.34}$$

$$\equiv \begin{bmatrix} -2A & (\nabla B^{\mathrm{s}})^{\mathrm{T}} \\ \nabla B^{\mathrm{s}} & -2D\mathbb{1}_{3\times 3} + 2\left(\nabla\nabla^{\mathrm{T}} - \mathbb{1}_{3\times 3}\frac{1}{3}\nabla^2\right) E^{\mathrm{s}} \end{bmatrix} \tag{1.35}$$

$$- \begin{bmatrix} 0 & (\mathbf{B}^{\mathrm{v}})^{\mathrm{T}} \\ \mathbf{B}^{\mathrm{v}} & \nabla\mathbf{E}^{\mathrm{v}} + (\nabla\mathbf{E}^{\mathrm{v}})^{\mathrm{T}} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 2E^{\mathrm{t}}_{ij} \end{bmatrix}, \tag{1.36}$$

where $\nabla \cdot \mathbf{B}^v$, $\nabla \cdot \mathbf{E}^v$, $\delta^{ik}(\partial_k E^t_{ij})$ and $\mathrm{Tr}\,\mathbf{E}^t$ all vanish. In the final form we have extracted the scalar (s), vector (v) and tensor (t) quantities from $\mathbf{B} = -\nabla B^s + \mathbf{B}^v$ and $\mathbf{E} = \left(\nabla\nabla^{\mathrm{T}} - \mathbb{1}_{3\times 3}\frac{1}{3}\nabla^2\right) E^s + \nabla\mathbf{E}^v + (\nabla\mathbf{E}^v)^{\mathrm{T}} + E^t_{ij}$.

In the linear perturbation theory the scalar, vector and tensor perturbations evolve independently. Moreover, the available observations can all be modelled by just the scalar perturbations. The future CMB polarization measurements such as PLANCK are hoped to detect the primordial B-mode polarization signal [42] stemming from the inflationary gravity waves that require tensor perturbations.

Limiting our attention to the scalar perturbations A, B, D and E we perform the gauge transformation (1.33) to a perturbed metric. Now the perturbation reads

$$h_{\mu\nu} = \begin{bmatrix} -2(A - \partial_0\xi^0 - \mathcal{H}\xi^0) & \nabla(B + \partial_0\xi + \xi^0)^{\mathrm{T}} \\ \nabla(B + \partial_0\xi + \xi^0) & -2(D - \frac{1}{3}\nabla^2\xi + \mathcal{H}\xi^0) + 2\left(\nabla\nabla^{\mathrm{T}} - \frac{1}{3}\nabla^2\right)(E + \xi) \end{bmatrix}. \tag{1.37}$$

We have chosen $\boldsymbol{\xi} = (\xi^0, \nabla\xi)$ in order not to introduce new vector perturbations.

The two remaining degrees of freedom in the gauge transformation allow us to choose $\xi = -E$ and $\xi^0 = \partial_0 E - B$, rendering the perturbation diagonal. This choice of $\boldsymbol{\xi}$ is called the *Newtonian gauge*. It is customary to call the remaining degrees of freedom $A = \Phi$ and $D = \Psi$:

$$\mathrm{d}s^2 = a(\eta)^2\left[-(1 + 2\Phi)\,\mathrm{d}\eta^2 + (1 - 2\Psi)\,\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j\right], \qquad \Phi, \Psi \ll 1. \tag{1.38}$$

Newtonian in this context refers to the fact that in the limit of sub-horizon scales and non-relativistic matter the perturbations can be identified as a single Newtonian gravitational potential, $\phi = \Phi = \Psi$, causing acceleration of a massive test particle as

$$\mathbf{a} = \nabla\phi. \tag{1.39}$$

So far we have only discussed perturbations in the geometry of the Universe i.e. the LHS of the Einstein equation (1.30). Accordingly there must be perturbations in the energy content of the Universe. Perturbing the stress-energy tensor will, however, lead us astray from the ideal fluid description and we need to identify additional components in the perturbed tensor. First, we may not wish to work in the perturbed fluid rest frame,

allowing for a possible 3-velocity component, $\mathbf{v}$. Second, the fluid can now support anisotropic stress: $\bar{p}\,\mathbf{\Pi} = \delta\mathbf{T}^{(3)} - \mathbb{1}_{3\times3}\,\frac{1}{3}\mathrm{Tr}\,\delta\mathbf{T}^{(3)}$, defined here as the traceless part of the spatial stress-energy tensor perturbation. We then have

$$\mathbf{T} = \bar{\mathbf{T}} + \delta\mathbf{T} \quad \Rightarrow \quad \begin{aligned} \rho &= \bar{\rho} + \delta\rho \\ p &= \bar{p} + \delta p \\ \mathbf{v} &= \bar{\mathbf{v}} + \delta\mathbf{v} \\ \mathbf{\Pi} &= \bar{\mathbf{\Pi}} + \delta\mathbf{\Pi} \end{aligned} \quad \Rightarrow \quad \begin{aligned} \delta &\equiv \frac{\delta\rho}{\bar{\rho}} \\ v &\equiv \delta v\,, \\ \Pi &\equiv \delta\Pi \end{aligned} \tag{1.40}$$

where, due to the isotropy of the background model, we can choose a frame where $\bar{\mathbf{v}} = 0$, and consider only the scalar modes of the velocity and shear perturbations ($\delta\mathbf{v} = -\nabla v$ and $\delta\mathbf{\Pi} = (\nabla\nabla^{\mathrm{T}} - \frac{1}{3}\nabla^2)\Pi$).

These perturbations are gauge dependent. If we make a gauge transformation, $\widetilde{x}^\mu = x^\mu + \xi^\mu$, with $\xi = (\xi^0, -\nabla\xi)$, they change:

$$\begin{aligned} \widetilde{\delta} &= \delta - \frac{\bar{\rho}'}{\bar{\rho}}\xi^0 & \widetilde{v} &= v + \partial_0\xi \\ \delta\widetilde{p} &= \delta p - \bar{\rho}'\xi^0 & \widetilde{\Pi} &= \Pi. \end{aligned} \tag{1.41}$$

Another useful gauge can be defined by choosing the transformation $\boldsymbol{\xi}$ to eliminate v and B. By equations (1.37) and (1.41) this means

$$B + \partial_0\xi + \xi^0 = 0 = v + \partial_0\xi \quad \Rightarrow \quad \begin{aligned} \partial_0\xi &= -v \\ \xi^0 &= v - B. \end{aligned} \tag{1.42}$$

This gauge condition leads to the *comoving gauge*. It is of particular interest since it has become a standard to parametrize the primordial perturbations from inflation by their comoving curvature perturbation. Recall that we assume the background to be spatially flat and homogeneous. Then any curvature on a constant η hypersurface is due to the perturbations. If we compute the Ricci curvature scalar of such a hypersurface we find

$$R = \frac{4}{a^2}\nabla^2\left(D + \tfrac{1}{3}\nabla^2 E\right) \equiv -\frac{4}{a^2}\nabla^2\mathcal{R}, \tag{1.43}$$

where we have defined the *curvature perturbation*, $\mathcal{R}$, which is directly related to the scalar curvature. The symbol $\mathcal{R}$ is taken to refer to the curvature perturbation *in the comoving gauge*.

Adding the scalar perturbations to both sides of the Einstein equation (1.30) provides us with the perturbation field equations, that in the conformal Newtonian gauge read:

$$\begin{aligned} \nabla^2\Psi - 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) &= 4\pi G a^2\delta\rho \\ \Psi' + \mathcal{H}\Phi &= 4\pi G a^2(\bar{\rho} + \bar{p})v \\ \Psi'' + \mathcal{H}(\Phi' + 2\Psi') + (2\mathcal{H}' + \mathcal{H}^2)\Phi - \tfrac{1}{3}\nabla^2(\Phi - \Psi) &= 4\pi G a^2\delta p \\ (\partial_i\partial_j - \tfrac{1}{3}\delta^i_j\nabla^2)(\Phi - \Psi) &= 8\pi G a^2\bar{p}(\partial_i\partial_j - \tfrac{1}{3}\delta^i_j\nabla^2))\Pi, \end{aligned} \tag{1.44}$$

where $\mathcal{H}$ is the *conformal Hubble parameter*, $\mathcal{H} = aH$. These need to be combined with our knowledge that the energy content of the Universe consists of several species of energy: photons, neutrinos, dark and baryonic matter, and (possibly) dark energy. Thus

we split ρ and p into components and couple them to each other based on their interactions. A necessary tool in doing so is the overall continuity of the energy-momentum tensor, Eq. (1.45), or

$$\nabla_\mu T^{\mu\nu} = 0. \tag{1.45}$$

The perturbed Einstein equations are written in terms of perturbations to individual species of energy and coupling between different species is entered as collision terms in the relativistic Boltzmann equations. It then turns out that to evolve these field equations it is very convenient to move to Fourier space and expand the angular dependence of the photon distribution in terms of Legendre polynomials. After setting the initial conditions at a time well before cosmological scales enter the horizon, the perturbations can be directly integrated into the CMB multipole expansion at present. Details of such calculation are unfortunately outside the scope of this thesis and we refer the reader to the papers [e.g. 18, 43, 44, 45] and textbooks [39, 41] of the field.

1.5 PLANCK Surveyor mission

Much of the material covered briefly in this section is explained in detail in the PLANCK Bluebook [46] and the references therein.

PLANCK is a satellite mission designed to produce full sky intensity and polarization maps in 9 different frequency bands between 30 and 857 GHz (0.35–10 mm). Combination of multi-level cryogenic stages, bolometer technology of the high frequency instrument (HFI) and HEMT radiometers in the low frequency instrument (LFI) grant PLANCK an unprecedented frequency coverage. That multi-frequency information is used to discern several galactic and extra-galactic foregrounds to produce pristine CMB maps [47]. Figure 1.4 shows the PLANCK frequency bands alongside with the COBE [48] and WMAP [49] bands.

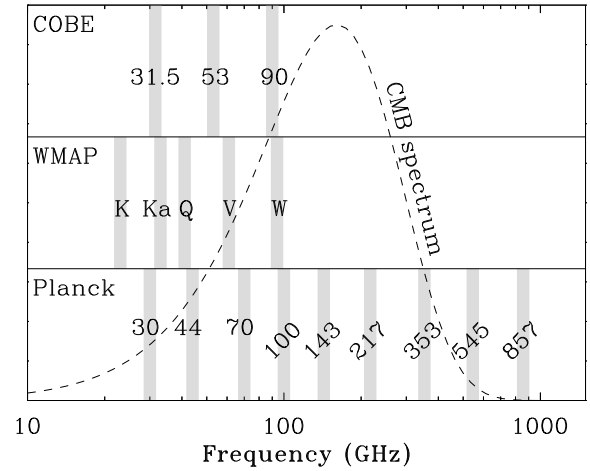


Figure 1.4: Frequency bands of the three full sky CMB surveys. Shown here is also the CMB intensity in arbitrary thermodynamical units.

1.5.1 Instrument and noise features

In contrast to its predecessors, PLANCK is an actively cooled, single focal plane instrument. Excess heat is first dissipated by passive V-groove radiators leading to the ambient 50 K temperature. The LFI front-end is then cooled down to 20 K using a novel hydrogen sorption cooler. The HFI is cooled down further first to 4 K by another Joule-Thomson device, and all the way down to 0.1 K by a ^{3}He/^{4}He dilution cooler. The 4 K stage provides the LFI the reference loads that relieve the LFI from using differencing assemblies like the COBE and WMAP. While cooling allows PLANCK to reach extremely low noise equivalent temperatures (NET) in its detectors, it also exposes the correlated $1/f$ behaviour of the noise samples [50, 51]. The feature is enhanced by the lack of real differencing assemblies immersed in the same thermal environment.

Let $\widetilde{\mathbf{n}}$ be the discrete Fourier transform of an instrument noise sample vector of length N:

$$\widetilde{n}_k = \mathcal{F}\{\mathbf{n}\}_k \equiv \sum_{s=0}^{N-1} n_s\, e^{-i2\pi ks/N}. \tag{1.46}$$

The PLANCK instrument noise is assumed Gaussian and has a zero mean. The two independent noise components, correlated and white, can be modelled by a power spectral density (PSD):

$$P(f) = \langle \widetilde{n}_f^* \widetilde{n}_f \rangle = \frac{\sigma^2}{f_\mathrm{s}} \cdot \frac{f^\alpha + f_\mathrm{knee}^\alpha}{f^\alpha + f_\mathrm{min}^\alpha}, \quad f = k\frac{f_\mathrm{s}}{N}, \tag{1.47}$$

where σ^2 is the white noise sample variance, f_s is the detector sampling frequency and slope, knee and minimum frequencies $(\alpha, f_\mathrm{knee}, f_\mathrm{min})$ are the noise model parameters. A sample noise power spectrum is depicted in Fig. 1.5.

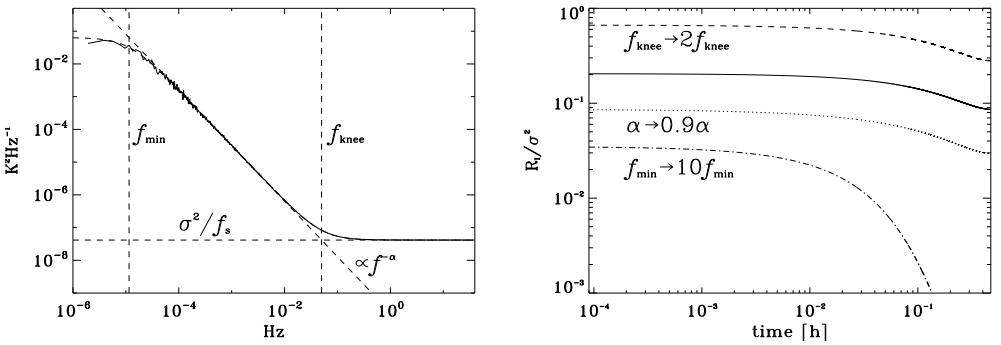


Figure 1.5: *Left:* parametrized and simulated instrument noise power spectrum. *Right:* noise autocorrelation computed for the model on the left, but f_min multiplied by 10 (solid) and effect of α, f_min and f_knee.

A non-white noise power spectrum implies a correlation between noise samples from the detector. According to the Wiener-Khinchin Theorem, the time domain noise autocorrelation function is the inverse Fourier transform of the noise power spectral density:

$$R_l \equiv \langle n_s n_{s+l} \rangle = f_\mathrm{s}\, \mathcal{F}^{-1}\{P(f)\}_l = f_\mathrm{s} \frac{1}{N} \sum_f P(f) e^{i2\pi l f}, \tag{1.48}$$

where the sum over f ranges from $-f_\mathrm{s}/2$ to $f_\mathrm{s}/2$ and N is the length of the noise vector. Right panel in Fig. 1.5 depicts the autocorrelation function normalized with the white noise variance and the effects of noise model parameters. One may conclude that the only parameter relevant for the correlation length is the minimum frequency that sets the scale for correlation. Others simply control the level of correlation with same correlation length.

1.5.2 Scanning strategy

PLANCK will be positioned in the Lagrange point 2 (L2 for short) of the Sun-Earth gravitational system. It is located 1.5 million kilometers from Earth, leaving Earth between the satellite and Sun. The satellite spin axis is directed away from the Sun and set to precess at an amplitude of 7.5 degrees with a period of seven months. As a result the spin axis tracks a cycloid pattern around the ecliptic on the celestial sphere once per year. The PLANCK focal plane axis is directed at a boresight angle of 85° away from the spin axis and follows nearly great circles between the ecliptic poles at a rate of 1 rpm. PLANCK completes a full sky survey roughly once in every seven months of operation. These parameters have been chosen to [52]

- adequately sample all sky pixels, needs both sufficient hit count and scanning directions
- observe the same regions of the sky over various time scales
- avoid unnecessary strong gradients in pixel hit counts

Unfortunately, all of these constraints cannot be satisfied at the same time. Figure 1.6 contains an illustration of the PLANCK scanning strategy and an example of a simulated hit distribution corresponding to that scanning strategy. It shows two deep fields of high hit count at the ecliptic poles. Outline of both fields is characterized by a strong gradient in pixel hit counts. In Paper **I** we discuss some aspects of these gradients in conjunction with map-making.

1.5.3 PLANCK analysis pipeline

The PLANCK data analysis is organized into five separate levels, four of which are shown as a diagram in Figure 1.7. Not shown in the plot is the vital simulation level (Level S) that has been a crucial tool in pipeline development and will provide necessary support for data analysis through Monte Carlo studies.

The PLANCK satellite will relay its observations daily to the mission operations centre (MOC) in Darmstadt. The low and high frequency instrument consortia have established two separate data processing centres (DPCs) that receive the instrument data

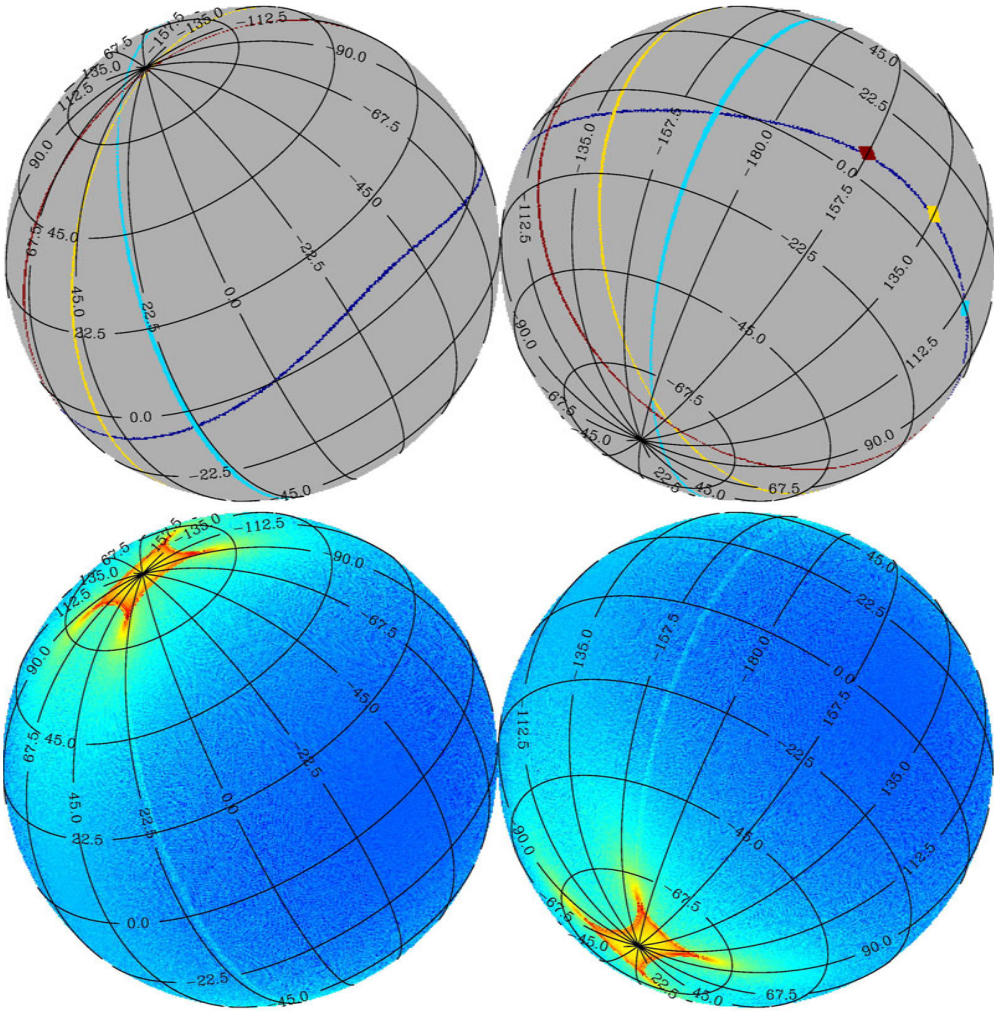


Figure 1.6: *Top:* Visualization of the cycloidal PLANCK scanning strategy for one year of observations. The satellite spin axis precesses around the equator along the dark blue line. The three scanning circles represent focal plane center pointing for one day of observations one month apart from each other. The corresponding spin axis pointings are marked by diamond shapes on the right side of top row. The PLANCK focal plane center opens at an angle $\theta_b = 85^\circ$ from the satellite spin axis. *Bottom:* Hits per pixel for one year of simulated observations using the 12 70 GHz detectors. The characteristic distribution of hits at the ecliptic poles is due to the cycloidal scanning strategy. The color scale extends logarithmically from 3,500 (blue) to 340,000 hits (red) per $7' \times 7'$ pixel.

and satellite telemetry from MOC. Level 1 constitutes of storing that data and processing the telemetry into satellite attitude information. Level 1 also provides feedback to the MOC about the satellite.

Level 2 receives the raw uncalibrated time-ordered information (TOI) from Level 1 and calibrates and compacts the data into intensity time lines. Most importantly, Level 2 processes those calibrated time lines into frequency maps of the microwave sky. Paper **I** deals with this map-making stage.

Level 3 is the component separation stage where the full frequency information from both of the instruments is combined to separate the foreground components, both galactic and extra-galactic, and the CMB into separate component maps. The DPC:s share data at this level but implement component separation independently for a necessary consistency check.

Level 4 stands for all subsequent scientific exploitation of the data. The CMB component map is analyzed into angular power spectra and cosmological parameter estimates. Paper **II** studies the residual noise covariance from the map-making stage. The noise covariance matrix developed in Paper **II** is an input to the power spectrum estimation. Paper **III** presents a study of an expanded phenomenological model hosting additional inflationary degrees of freedom. It provides the cosmological parameter estimates using a dataset similar to the PLANCK data.

Additional statistical analysis such as the non-gaussianity studies take place on level 4. A body of non-CMB science is also performed on the component maps, e.g. the study of Sunyaev-Zeldovich effect that is a signal of ionized intergalactic media.

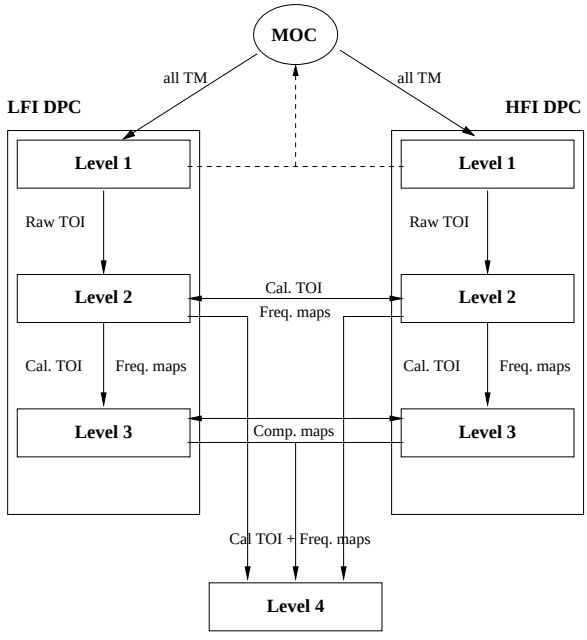


Figure 1.7: PLANCK data analysis pipeline organized into four levels. Details: see text. The diagram is reproduced from [46].

1.6 Notational conventions

In the remainder of this thesis we employ the natural unit system defined by setting

$$c = k_{\rm B} = \hbar = 1 \tag{1.49}$$

if not explicitly stated otherwise.

1.6.1 CMB analysis

The field of CMB experiments is still rather new and for that reason many frequently encountered quantities still lack an agreed symbol. An alert reader may notice that not even the papers **I** and **II** conform to the same notation. Table 1.1 lists some frequently used quantities in the CMB analysis with their chosen symbols.

Table 1.1: Frequently encountered symbols

Symbol	Definition
$\mathbf{P}$	detector pointing matrix
$\mathbf{F}$	offset projection matrix
$\mathbf{a}$	baseline offset vector
$\mathbf{m}$	$3N_{\rm pix}$ map vector
$\hat{\mathbf{m}}$	map estimate
$\mathbf{n}$	noise vector
$\mathbf{n}_{\rm w}$	white noise vector
$\mathbf{d}$	TOD vector, $\mathbf{d} = \mathbf{Pm} + \mathbf{n}$
$\mathbf{N}$	noise covariance, map domain
$\mathcal{N}$	noise covariance, time domain
$\mathcal{N}_{\rm c}$	correlated noise covariance, time domain
$\mathbf{N}_{\rm w}$	white noise covariance, map domain
$\mathcal{N}_{\rm w}$	white noise covariance, time domain
$\mathbf{M}_{\rm p}$	3×3 observation matrix for pixel p
$\mathcal{N}_{\rm a}$	prior baseline offset covariance

Throughout the thesis, maps are presented in Hierarchical Equal Area iso-Latitude Pixelization (HEALPix) [53]. The resolution of a HEALPix map is defined by its resolution parameter, r, or its $\mathrm{N_{side}}{=}2^r$. The number of pixels in a HEALPix map is $\mathrm{N_{pix}} = 12 \cdot \mathrm{N_{side}^2}$.

Chapter 2

Destriping approach to map-making

The satellite scans the microwave sky sampling the temperature and polarization producing sizable time ordered data vectors. Before we can compare cosmological models to these observations we need to reduce the data size in some manner. An obvious reduction operation is to make a map of the observations. However, care must be taken in order not to destroy any cosmological information present in the TOD. Failure to model and treat the instrument noise properly will lead to greater uncertainty of the pixel temperatures and thus degradation of the cosmological information.

2.1 Preliminaries

The calibrated output of a detector is a vector of observations. The magnitude of an observation (sample) depends on the sky temperature, its polarization and instrument noise. This time ordered data vector (TOD for short), $\mathbf{d}$, can be decomposed as

$$\mathbf{d} = \mathbf{s} + \mathbf{n} = \mathbf{s} + \mathbf{n}_{\mathrm{w}} + \mathbf{n}_{\mathrm{c}}, \tag{2.1}$$

where $\mathbf{s}$ denotes the sky signal and the noise vector $\mathbf{n}$ is broken down into an uncorrelated white $\mathbf{n}_{\mathrm{w}}$ and a correlated $\mathbf{n}_{\mathrm{c}}$ part.

The instrument response to sky temperature and polarization can be formulated as

$$s_t = (1+\epsilon)\mathbf{I}_p + (1-\epsilon)\left[\mathbf{Q}_p \cos(2\chi_t) + \mathbf{U}_p \sin(2\chi_t)\right], \tag{2.2}$$

where $\mathbf{I}$,$\mathbf{Q}$, and $\mathbf{U}$ are the pixelized, beam-smoothed sky maps of the corresponding Stokes components, χ is the detector polarization angle, and t and p label samples and pixels respectively. ϵ is the small cross polar leakage factor that for the LFI and this thesis can be approximated as zero. We ignore beam asymmetry and idealize that the detector response is a function of only the targeted pixel of the beam-smoothed sky map. For a more precise treatment, the detector beam should be convolved with the

full sky. Depending on the selected map resolution and frequency channel, a detector beam is 2–5 times wider than an average pixel.

Following the idealized sky response model we can group the three Stokes maps into a single map vector $\mathbf{m} = [\mathbf{I}, \mathbf{Q}, \mathbf{U}]$ and establish a linear transformation between the sky TOD $\mathbf{s}$ and the map $\mathbf{m}$:

$$\mathbf{s} = \mathbf{P}\mathbf{m}. \tag{2.3}$$

The transformation is defined by the pointing matrix $\mathbf{P}$ that picks the correct contributions from the map for each of the samples in the TOD. In our case, $\mathbf{P}$ is an extremely sparse $n_{\mathrm{samples}} \times 3\mathrm{N}_{\mathrm{pix}}$ matrix hosting only three non-zero elements per row. In a more precise treatment, such as the deconvolution map-making [54], the rows would be replaced by maps of the beam sensitivity in proper orientations.

2.2 Maximum likelihood map-making

We formulate map-making as an attempt to recover an estimate, $\widetilde{\mathbf{m}}$, of the true sky map, $\mathbf{m}$, from the TOD, $\mathbf{d}$:

$$\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{n}, \quad \widetilde{\mathbf{m}} = \mathbf{L}\mathbf{d}, \tag{2.4}$$

and wish to choose the linear operator, $\mathbf{L}$, so that it minimizes the difference between $\mathbf{m}$ and $\widetilde{\mathbf{m}}$:

$$\langle (\widetilde{\mathbf{m}} - \mathbf{m}) (\widetilde{\mathbf{m}} - \mathbf{m})^{\mathrm{T}} \rangle. \tag{2.5}$$

Both the map, $\mathbf{m}$, and the noise, $\mathbf{n}$, are assumed to be Gaussian, zero mean vectors.

Let us start by assuming that we know the noise covariance matrix of the detector noise, allowing us to rotate the noise vector into an uncorrelated Gaussian with unit variance:

$$\mathcal{N} \equiv \langle \mathbf{n}\mathbf{n}^{\mathrm{T}} \rangle \quad \Rightarrow \quad \mathbf{n} = \mathcal{N}^{1/2}\mathbf{n}', \tag{2.6}$$

where $\langle \mathbf{n}'\mathbf{n}'^{\mathrm{T}} \rangle = \mathbb{1}$ and $\langle \mathbf{n} \rangle = 0$. If noise is stationary, then $\mathcal{N}$ is circulant. For a real experiment, stationarity over the whole mission may be unjustified, but even in that case $\mathcal{N}$ can be treated as a block diagonal matrix where each block corresponds to a stationary interval. It is also band diagonal, the width being determined by the correlated noise minimum frequency (cf. Figure 1.5).

Multiplying the expression for $\mathbf{d}$ by $\mathcal{N}^{-1/2}$, yields

$$\mathcal{N}^{-1/2}\mathbf{d} = \mathcal{N}^{-1/2}\mathbf{P}\mathbf{m} + \mathbf{n}'. \tag{2.7}$$

Since $\mathbf{n}'$ lacks all statistical structure, the best we can do is to simply neglect it, leading to a solvable estimate of $\widetilde{\mathbf{m}}$:

$$\widetilde{\mathbf{m}} = \left(\mathbf{P}^{\mathrm{T}}\mathcal{N}^{-1}\mathbf{P}\right)^{-1} \mathbf{P}^{\mathrm{T}}\mathcal{N}^{-1}\mathbf{d}, \tag{2.8}$$

where we have multiplied by $\mathbf{P}^{\mathrm{T}}\mathcal{N}^{-1/2}$ to make the prefactor of $\widetilde{\mathbf{m}}$ invertible. We identify $\left(\mathbf{P}^{\mathrm{T}}\mathcal{N}^{-1}\mathbf{P}\right)^{-1} \mathbf{P}^{\mathrm{T}}\mathcal{N}^{-1}$ as our linear map-making operator in $\widetilde{\mathbf{m}} = \mathbf{L}\mathbf{d}$, and note that this estimate, $\widetilde{\mathbf{m}}$, leads to *minimum variance* of the difference $\widetilde{\mathbf{m}} - \mathbf{m}$. A

map-making operator defined in this manner was applied already to the COBE data [55] and belongs to the equivalence class of lossless map-making methods listed in [56]. In fact, it is the information content of the *noise-weighted map*, $\mathbf{P}^{\mathrm{T}}\mathcal{N}^{-1}\mathbf{d}$, that determines the possible loss of information.

2.3 Destriping

Solving minimum variance maps from (2.8) for a PLANCK sized dataset is exhaustive in terms of both memory and processing power. Nevertheless the method has been successfully applied to the BOOMERanG [57] data and simulated LFI data within the PLANCK Working Group 3 [58, 59, 60, 61]. It is desirable to have in hand a tool that can, under suitable conditions, perform nearly optimally using significantly fewer resources. The $1/f$ form of the correlated noise spectrum ensures that most of the correlated noise power manifests at low frequencies. That being the case, the correlated part of the noise is effectively modelled by a series of constant offsets of fixed length called *baselines*. In this approximation the correlated noise part is written as

$$\mathbf{n}_{\mathrm{c}} = \mathbf{F}\mathbf{a}, \tag{2.9}$$

where $\mathbf{a}$ is a vector of the baseline amplitudes and $\mathbf{F}$ is a matrix that projects the amplitudes onto the TOD. The length of the baseline offset is usually chosen to be in the range of 1 second to 1 hour, corresponding roughly to 10^2 to 10^5 samples. Now the time domain noise covariance becomes [62]

$$\mathcal{N} = \mathcal{N}_{\mathrm{w}} + \mathcal{N}_{\mathrm{c}} = \mathcal{N}_{\mathrm{w}} + \mathbf{F}\mathcal{N}_{\mathrm{a}}\mathbf{F}^{\mathrm{T}}. \tag{2.10}$$

Like $\mathcal{N}$, the baseline correlation matrix, $\mathcal{N}_{\mathrm{a}}$, is approximately band diagonal, the width of the diagonal being inversely proportional to baseline length.

In our derivation of the minimum variance map we took the underlying map, $\mathbf{m}$, as fixed and removed noise statistically. Standard destriping [63, 64, 62, 65] takes additionally also the baseline amplitudes as fixed. Therefore one solves for $\widetilde{\mathbf{m}}$ and $\widetilde{\mathbf{a}}$ and only considers variations in the white noise levels. The decorrelation procedure looks like:

$$\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{F}\mathbf{a} + \mathbf{n}_{\mathrm{w}} \quad \Rightarrow \quad \mathcal{N}_{\mathrm{w}}^{-1/2}\mathbf{d} = \mathcal{N}_{\mathrm{w}}^{-1/2}\mathbf{P}\mathbf{m} + \mathcal{N}_{\mathrm{w}}^{-1/2}\mathbf{F}\mathbf{a} + \mathbf{n}' \tag{2.11}$$

Neglecting $\mathbf{n}'$ allows us to solve for $\widetilde{\mathbf{m}}$,

$$\widetilde{\mathbf{m}} = \left(\mathbf{P}^{\mathrm{T}}\mathcal{N}_{\mathrm{w}}^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathcal{N}_{\mathrm{w}}^{-1}\left(\mathbf{d} - \mathbf{F}\widetilde{\mathbf{a}}\right), \tag{2.12}$$

and substituting $\widetilde{\mathbf{m}}$ into Eq. (2.11) leads to an equation for $\widetilde{\mathbf{a}}$:

$$\mathbf{F}^{\mathrm{T}}\mathcal{N}_{\mathrm{w}}^{-1}\mathbf{Z}\mathbf{F}\widetilde{\mathbf{a}} = \mathbf{F}^{\mathrm{T}}\mathcal{N}_{\mathrm{w}}^{-1}\mathbf{Z}\mathbf{d}, \quad \text{where} \quad \mathbf{Z} \equiv \mathbb{1} - \mathbf{P}\left(\mathbf{P}^{\mathrm{T}}\mathcal{N}_{\mathrm{w}}^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathcal{N}_{\mathrm{w}}^{-1}. \tag{2.13}$$

After solving the baselines from equation (2.13), one then subtracts them from the original TOD ending up with essentially sky signal and simple white noise. Some residual correlated noise will remain but is approximated as white for the purpose of problem

reduction. Solving the minimum variance map (2.8) from such a TOD is computationally much less intensive than the complete case since the band diagonal $\mathcal{N}$ is replaced by a diagonal $\mathcal{N}_{\rm w}$.

In practice the baselines are most conveniently solved through conjugate gradient iteration [66]. In fact, the $n_{\rm baselines} \times n_{\rm baselines}$ matrix $\mathbf{F}^{\rm T}\mathcal{N}_{\rm w}^{-1}\mathbf{Z}\mathbf{F}$ is not even invertible.

2.3.1 Destriping with prior information

The fundamental difference between the destriping and optimal approaches is that the destriping only employs information about the instrument noise equivalent temperatures, whereas the optimal method makes use of the full noise power spectrum. A middle ground of the two approaches is a map-making algorithm that does model the correlated noise by constant baseline offsets but employs the noise power spectrum to constrain the baseline amplitudes. Here, we call this approach *generalized destriping*. In this case, the noise decorrelation procedure includes a baseline correlation matrix, $\mathcal{N}_{\rm a}$, that can be computed from the correlated instrument noise power spectrum. Basically this matrix contains the correlations of averages of noise samples.

We formulate this algorithm by substituting the destriping approximation for noise covariance, Eq. (2.10) into the minimum variance map equation (2.8):

$$\widetilde{\mathbf{m}} = \left(\mathbf{P}^{\rm T}\left(\mathcal{N}_{\rm w} + \mathbf{F}\mathcal{N}_{\rm a}\mathbf{F}^{\rm T}\right)^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\rm T}\left(\mathcal{N}_{\rm w} + \mathbf{F}\mathcal{N}_{\rm a}\mathbf{F}^{\rm T}\right)^{-1}\mathbf{d}. \tag{2.14}$$

Unfortunately, this form of the generalized destriping bears no computational advantages over the optimal method since all relevant matrix products are still enormous. Instead, one can follow the standard destriping steps and first solve for the baseline offsets, now employing the full knowledge of the noise properties. This time the baseline amplitudes are solved from

$$\left(\mathbf{F}^{\rm T}\mathcal{N}_{\rm w}^{-1}\mathbf{Z}\mathbf{F} + \mathcal{N}_{\rm a}^{-1}\right)\widetilde{\mathbf{a}} = \mathbf{F}^{\rm T}\mathcal{N}_{\rm w}^{-1}\mathbf{Z}\mathbf{d}. \tag{2.15}$$

If no *a priori* information on the baseline correlation is available, one is forced to assign infinite variances to the baseline amplitudes. In such a case, the term $\mathcal{N}_{\rm a}^{-1} \propto \sigma_a^{-2}$ vanishes and we recover the standard destriping equation (2.13)

If the baseline vector, $\mathbf{a}$, is stationary or can be treated as stationary pieces, then $\mathcal{N}_{\rm a}^{-1}$ is approximately circulant and band diagonal, allowing a very efficient computation of $\mathcal{N}_{\rm a}^{-1}\widetilde{\mathbf{a}}$ using Fourier techniques. Thus the additional noise information has induced only minimal delay to each conjugate gradient iteration. In effect, the noise filter ($\mathcal{N}_{\rm a}^{-1}$) constrains the solved baseline amplitudes and allows solving for shorter baselines than standard destriping.

2.3.2 Breakdown into components

Being a linear operation, the destriping allows individual study of all the elements present in the TOD: we can make separate maps from all the independent components

present in the time ordered signal. Paper **I** features a scrupulous analysis of all components encountered in destriping. Furthermore, the relative significance of these individual components varies over three different domains: frequency, time and map domain. Note that when applying the destriping operation to the full TOD, one unavoidably applies it to the sky signal and white noise, to neither of which it is advantageous to do so.

In the map domain the components are hierarchically as follows:

- sky signal
  - binned sky signal
  - pixelization noise baselines
  - pixelization noise baseline errors
- instrument noise
  - white noise
    - white noise
    - reference baselines
    - baseline errors
  - correlated noise
    - correlated noise
      - part modelable by baseline offsets
      - part unmodelled by baseline offsets
    - reference baselines
    - baseline errors

By reference baselines we mean the naïve average of samples over the baseline. The pixelization noise is the variation in samples drawn from the same pixel, stemming from subpixel signal gradients. When sufficient information is available and subpixel structure of the sky signal does not interfere, a destriper solves nearly zero baseline amplitudes for the sky signal and white noise and removes the part of the correlated noise modelable by baseline offsets.

Figure 2.1 features three top level components (the CMB, white and correlated noise) in frequency domain. Power spectral densities are plotted before and after subtracting 15 s baseline offsets solved by a standard destriper. A number of notable features are present in the figure: destriping alters dramatically the noise spectra but leaves CMB almost intact and correlated noise is removed more effectively than white noise. Also, the CMB spectrum exhibits harmonics of the satellite spin rate, 1 rpm, and repointing frequency $1\,\mathrm{h}^{-1} \approx 3 \times 10^{-4}\,\mathrm{Hz}$. Baseline effects only show at frequencies lower than $1/15\,\mathrm{s} \approx 7 \times 10^{-2}\,\mathrm{Hz}$. Furthermore, the rise in low frequency power of the destriped noise spectra is due to the difference between solved and reference baselines (baseline error).

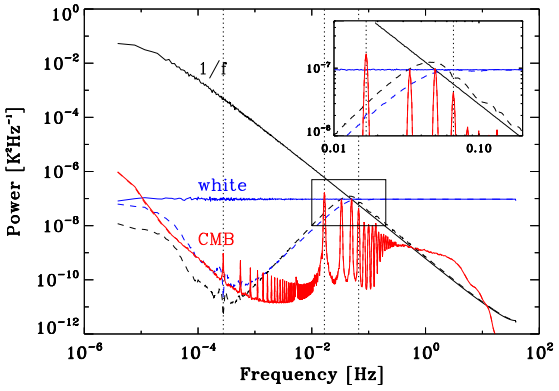


Figure 2.1: Component powers in spectrum domain. The solid lines represent the power spectral densities of three TOD components. The dashed lines of the corresponding colours are the same TOD:s after destriping with 15 s baselines. The three vertical dotted lines mark (from left to right) the repointing, satellite spin and baseline frequencies.

Binning these signals into a map introduces additional phenomena. If we write the destriped TOD as $\mathbf{d}' = \mathbf{d} - \mathbf{Fa}$ and compute the covariance we get

$$\langle \mathbf{d}'\mathbf{d}'^{\mathrm{T}} \rangle = \langle \mathbf{d}\mathbf{d}^{\mathrm{T}} \rangle - \langle \mathbf{d}\mathbf{a}^{\mathrm{T}} \rangle \mathbf{F} - \mathbf{F} \langle \mathbf{a}\mathbf{y}^{\mathrm{T}} \rangle + \mathbf{F} \langle \mathbf{a}\mathbf{a}^{\mathrm{T}} \rangle \mathbf{F}^{\mathrm{T}}, \tag{2.16}$$

where all signals are assumed to have zero mean. For a sample y_t, on a baseline a_t, we find

$$\langle d_t'^{\,2} \rangle = \sigma_{d'}^2 = \sigma_d^2 - 2\langle d_t a_t \rangle + \sigma_a^2 \tag{2.17}$$

and we see that unless there is a strong positive correlation between the TOD samples and their baselines, destriping only adds to the variance in time domain. That correlation is dictated by the fraction of the signal that can be represented by the baseline offsets and is strongest for the $1/f$ noise and weakest for the CMB signal.

It is not given that these cancellations make it to the pixel temperatures. For a simple case, consider a non-polarized measurement with a single detector. Now the pixel temperature from binning is the simple average over samples hitting that pixel,

$$m_p = \frac{1}{n_{\mathrm{hits}}} \sum_{t \in p} d_t, \tag{2.18}$$

resulting in a pixel variance,

$$\begin{aligned} \langle m_p^2 \rangle &= \frac{1}{n_{\mathrm{hits}}^2} \sum_{t,s \in p} \langle d_t' d_s' \rangle \\ &= \frac{1}{n_{\mathrm{hits}}^2} \sum_{t,s \in p} \Big(\langle d_t d_s \rangle - \langle a_t d_s \rangle - \langle d_t a_s \rangle + \langle a_t a_s \rangle \Big), \end{aligned} \tag{2.19}$$

which is strongly influenced by sample-baseline correlation. For white noise, terms with t and s on different baselines vanish. If we approximate the solved baselines by reference baselines, we have $\langle d_t a_t \rangle = \sigma_{\mathrm{a}}^2 = \sigma_{\mathrm{w}}^2 / n_{\mathrm{base}}$, leaving

$$\sigma_p^2 \approx \frac{1}{n_{\mathrm{hits}}^2} \sum_{t \in p} \left(\sigma_{\mathrm{w}}^2 - b\sigma_{\mathrm{a}}^2 \right) = \frac{\sigma_{\mathrm{w}}^2 - b\sigma_{\mathrm{a}}^2}{n_{\mathrm{hits}}} \approx \left(1 - \frac{b}{n_{\mathrm{base}}}\right) \frac{\sigma_{\mathrm{w}}^2}{n_{\mathrm{hits}}}, \tag{2.20}$$

where we include the factor $b \sim 1$ to account for multiple samples on a single baseline falling into same pixel. Table 2.1 lists sample and pixel standard deviations for the CMB and noise and shows how sample-baseline correlations alter them. Signal components listed in Table 2.1 are the same as in Figure 2.1, showing the components in three different domains: time, frequency, and map. Notable is that while subtracting reference baselines, in accordance to equation (2.20), reduces pixel variance, baseline misestimation negates the effect for white noise.

Table 2.1: Component powers before and after destriping with 15 s baselines

	Standard deviation [μK]						
	Time domain[a]				Map domain[b]		
Component	σ_d	$\sigma_{d'}$	σ_a	σ_a(ref)	σ_p(bin)	σ_p(ds)	σ_p(ref)
CMB	99.94	99.93	0.462	24.88	99.85	99.83	99.82
$1/f$	1 766	153.6	1 862	1 862	487.1	4.163	3.852
white	2 700	2 699	79.63	79.55	55.19	55.24	55.16

[a] d and d' refer to the time line samples before and after subtracting the baselines, a. "ref" stands for reference baselines (naive averages) instead of actual solved baselines.

[b] We report pixel standand deviations of an N_{side}=512 Stokes I map. The abbreviations "bin", "ds" and "ref" refer to binned map, destriped map and destriped with reference baselines respectively.

Chapter 3

Residual noise characterization

The second of the selected topics deals with map-making residuals. For any high precision observation the quality of science deliverables is only as good as the understanding of non-scientific components still present in the data. That knowledge ultimately limits all attempts to further refine the data products. The PLANCK frequency maps act as inputs to component separation and power spectrum estimation that both rely on a solid understanding of the residuals.

3.1 Preliminaries

No map-making method is perfect in the sense that the estimated map, $\widetilde{\mathbf{m}}$, always contains some degree of residuals. The residuals are defined to be the difference between the beam-smoothed underlying sky map and the estimate:

$$\triangle \mathbf{m} = \widetilde{\mathbf{m}} - \mathbf{m}. \tag{3.1}$$

According to Paper **I**, there are numerous components to this difference. Even if our noise model is perfect, we can only remove the statistically most probable noise realization from the time ordered data. To which extent that realization is defined, depends on aspects of the survey, such as the scanning strategy, focal plane geometry and instrument performance. Also, since the destriping principle employs a coarse grained noise model (baselines), that clearly limits the accuracy of our noise description. The part of noise that escape our noise model is dubbed *unmodelled* noise. For the destriping approximation the term includes all non-white features in the power spectral density above the baseline frequency $f_{\rm base} = 1/t_{\rm base}$.

In addition to the noise effects, the map-making algorithm may induce distortion in the sky signal. These distortions only become significant when map-making is applied at a low resolution where the sky signal has considerable subpixel structure. Such will not be the case for a high precision CMB experiment. Therefore this chapter deals with noise related residuals.

Statistical description of the residuals is established in the form of a pixel-pixel noise

covariance matrix

$$\mathbf{N} = \langle \triangle\mathbf{m} \triangle\mathbf{m}^{\mathrm{T}} \rangle. \tag{3.2}$$

Inserting here the minimum variance map estimate (2.8) allows one to write for this matrix

$$\mathbf{N} = \left(\mathbf{P}^{\mathrm{T}} \mathcal{N}^{-1} \mathbf{P} \right)^{-1}. \tag{3.3}$$

For generalized destriping it suffices to insert here the approximated noise covariance,

$$\mathbf{N} = \left(\mathbf{P}^{\mathrm{T}} \left(\mathcal{N}_{\mathrm{w}} + \mathbf{F} \mathcal{N}_{\mathrm{a}} \mathbf{F}^{\mathrm{T}} \right)^{-1} \mathbf{P} \right)^{-1}, \tag{3.4}$$

and to make full use of the reduced problem dimension we furthermore employ the Woodbury formula[1] to write

$$\mathbf{N}^{-1} = \mathbf{P}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{P} - \mathbf{P}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{F} \left(\mathbf{F}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{F} + \mathcal{N}_{\mathrm{a}}^{-1} \right)^{-1} \mathbf{F}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{P}, \tag{3.5}$$

effectively reducing the problem size from number of samples into number of baselines.

Standard destriping is the limit of generalized destriping when no *a priori* information of the baseline correlation are available. Thus we let $\sigma_a^2 \to \infty$ with the effect that $\mathcal{N}_{\mathrm{a}}^{-1} \to 0$ leaving

$$\mathbf{N}^{-1} = \mathbf{P}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{P} - \mathbf{P}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{F} \left(\mathbf{F}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{F} \right)^{-1} \mathbf{F}^{\mathrm{T}} \mathcal{N}_{\mathrm{w}}^{-1} \mathbf{P}. \tag{3.6}$$

This line of thinking implies that both optimal and generalized destriping noise models are assumed accurate and no unmodelled noise is present. In reality destriping always ignores the part of noise that cannot be modelled by the constant baseline offsets. Paper **II** quantifies in detail the effects of unmodelled noise and shows the approximation to be good for short baselines and low knee frequencies.

3.2 To noise covariance from its inverse

Applying equation (3.3) and its successors lead to a situation where one needs to invert a symmetric $3\mathrm{N}_{\mathrm{pix}} \times 3\mathrm{N}_{\mathrm{pix}}$ matrix. Potentially the problem is huge: the matrix size scales as $\mathrm{N}_{\mathrm{pix}}^2 \sim \mathrm{N}_{\mathrm{side}}^4$ and the actual inversion by most methods as $\mathrm{N}_{\mathrm{pix}}^3 \sim \mathrm{N}_{\mathrm{side}}^6$.

Map-making algorithms attempt to resolve the correlated noise component by comparing samples measured from the same pixels at different times and by different detectors. The methods are unable to resolve the overall average amplitude, the monopole, of the temperature map. The same limitation applies to the analytical noise covariance matrices: when no measures are taken to prevent it, the ill-determined temperature monopole renders the covariance matrix singular. We have chosen to regularize the matrices by projecting out the ill-conditioned eigenmode:

$$\mathbf{N}^{-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\mathrm{T}} = \sum_{i=0} \lambda_i \widehat{\mathbf{e}}_i \widehat{\mathbf{e}}_i^{\mathrm{T}} \quad \to \quad \mathbf{N}_{\mathrm{reg}} = \sum_{i=1} \frac{1}{\lambda_i} \widehat{\mathbf{e}}_i \widehat{\mathbf{e}}_i^{\mathrm{T}}, \tag{3.7}$$

[1] $(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U}\right)^{-1}\mathbf{VA}^{-1}$

where we have written the covariance matrix as a product of orthogonal and diagonal matrices, $\mathbf{U}$ and $\mathbf{\Lambda}$, and ultimately in terms of its eigenvalues and eigenvectors λ_i and $\widehat{\mathbf{e}}_i$. Because $\mathbf{N}$ is real and symmetric, it always has $3\mathrm{N}_{\mathrm{pix}}$ linearly independent eigenvectors.

Both the eigenvalue analysis and the consecutive matrix inversion are demanding operations when the matrices are computed even for moderate resolutions such as $\mathrm{N}_{\mathrm{side}}$=32. To invert such a 11 GB matrix requires 30 CPU hours on a modern desktop. Advancing to a step better resolution of $\mathrm{N}_{\mathrm{side}}$=64 multiplies these requirements by 16 and 64 respectively. Fortunately the ScaLAPACK[2] linear algebra package provides efficient MPI parallelized routines for these operations on a parallel platform. Currently the inversion and regularization of an $\mathrm{N}_{\mathrm{side}}$=64 matrix is computationally feasible and taken to suffice.

3.3 Structure of the noise covariance

We wrote a code implementing equation (3.5), the generalized destriper case, and used it to compute the residual noise covariance matrix for a realistic scanning of the sky using the Planck 70 GHz channel comprising 12 detectors. The noise parameters for the simulation were chosen as $f_{\mathrm{s}} = 76.8\,\mathrm{Hz}$, $f_{\mathrm{min}} = 1.15 \times 10^{-5}\,\mathrm{Hz}$, $f_{\mathrm{knee}} = 50\,\mathrm{mHz}$, and $\alpha = 1.7$. Since each row and column is really a covariance map of the corresponding pixel, we can plot them using standard HEALPix visualization tools. For clarity we normalize the covariance map into correlation factors:

$$\mathrm{Corr}(m_p, m_q) \equiv \frac{\langle m_p m_q \rangle}{\sqrt{\langle m_p^2 \rangle}\sqrt{\langle m_q^2 \rangle}}, \tag{3.8}$$

where $\mathbf{m}$ is to be taken a noise-only map. Figures 3.1 and 3.2 exhibit the estimated correlation factors. We have chosen two reference pixels fixing $p = 0$ or $p = 2047$. Given HEALPix resolution $\mathrm{N}_{\mathrm{side}}$=32 and NESTED ordering scheme, these two pixels are located at the equator and north pole respectively. In a single map, q assumes all values between 0 and $\mathrm{N}_{\mathrm{pix}} - 1 = 12287$.

3.4 Downgrading maps and matrices

It is not trivial to produce useful low resolution maps from a high resolution observation. One has the choice of a number of methods all leading to different signal and noise structures. If we examine the problem from power spectrum estimation viewpoint, the most useful methods are those that have analytical descriptions (covariances) of the induced signal and noise distortions.

Another desired virtue of a low resolution map is being bandwidth limited. Pixelized spherical harmonic functions, $Y_{\ell m}(\theta, \phi)$, form an approximately orthogonal set of functions on a pixelized sphere up to $\ell_{\mathrm{max}} \approx 3\mathrm{N}_{\mathrm{side}} - 1$. When a high resolution map containing small angular scale power is re-pixelized to lower resolution, power beyond

[2] `http://www.netlib.org/scalapack/scalapack_home.html`

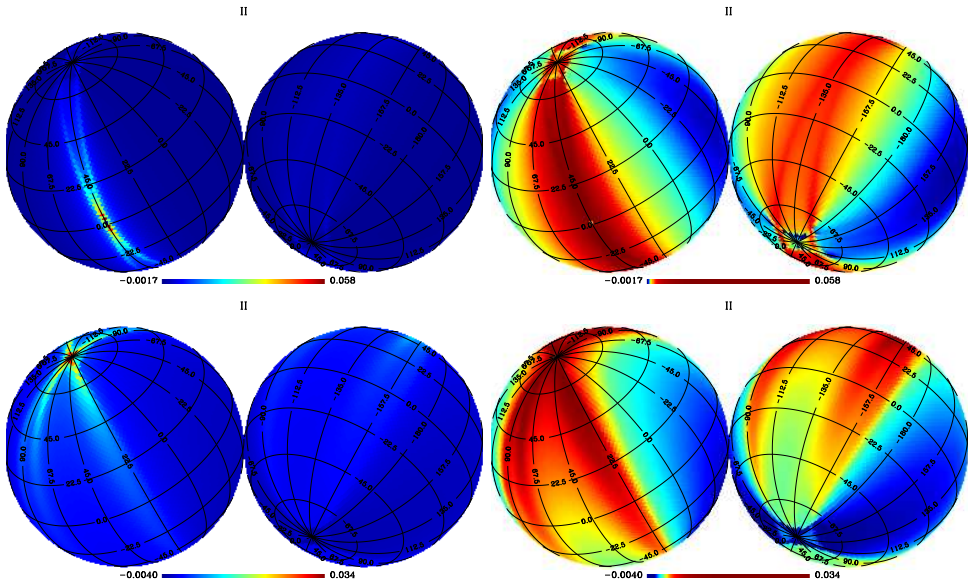


Figure 3.1: Visualization of the noise covariance matrix. *Top:* Noise correlation of an equatorial pixel. *Bottom:* Noise correlation of a polar pixel. *Left:* Linear color scale highlights the regions of strongest correlation. *Right:* Histogram equalized color scale reveals a rich pattern of correlation over the entire map. Cf. Figure 3.2 for correlation with polarization maps.

this $\ell_{\max}$ gets aliased to lower angular scales distorting the power spectrum estimates. For iterative power spectrum estimators such as the Gibbs sampler [67], it is imperative to make the signal bandwidth limited.

In Paper **II** we describe and study three low resolution strategies: direct map-making, noise-weighted downgrading, and smoothing downgrading. The first method is the simplest case and the appropriate noise covariance was developed in equations (3.3) and (3.5). It has the significant downside of inducing signal distortions due to subpixel structure. A more refined method is to recover the original noise-weighted map, $\mathbf{P}^{\mathrm{T}}\mathcal{N}^{-1}\mathbf{d}'$ (see Eq. (2.8)), from the high resolution map and rebin it according to estimated noise levels. In this approach the signal distortion is limited to pixelization smoothing and the noise covariance is very close to direct map-making.

If it is necessary to limit the signal bandwidth to angular scales supported by the low resolution, it is possible to downgrade the map through harmonic analysis. In this approach the harmonic expansion of the high resolution map is convolved with a symmetric Gaussian beam designed to suppress high-ℓ power and the smoothed expansion is resynthesized to lower resolution. In this last considered case both noise and signal covariances need to be modified accordingly with the smoothing beam window function.

The presented matrices describe the pixel-pixel correlations directly at the map-making resolution. However it is obvious that this will not be the case for low resolution maps downgraded from high resolution, a procedure called for by the signal

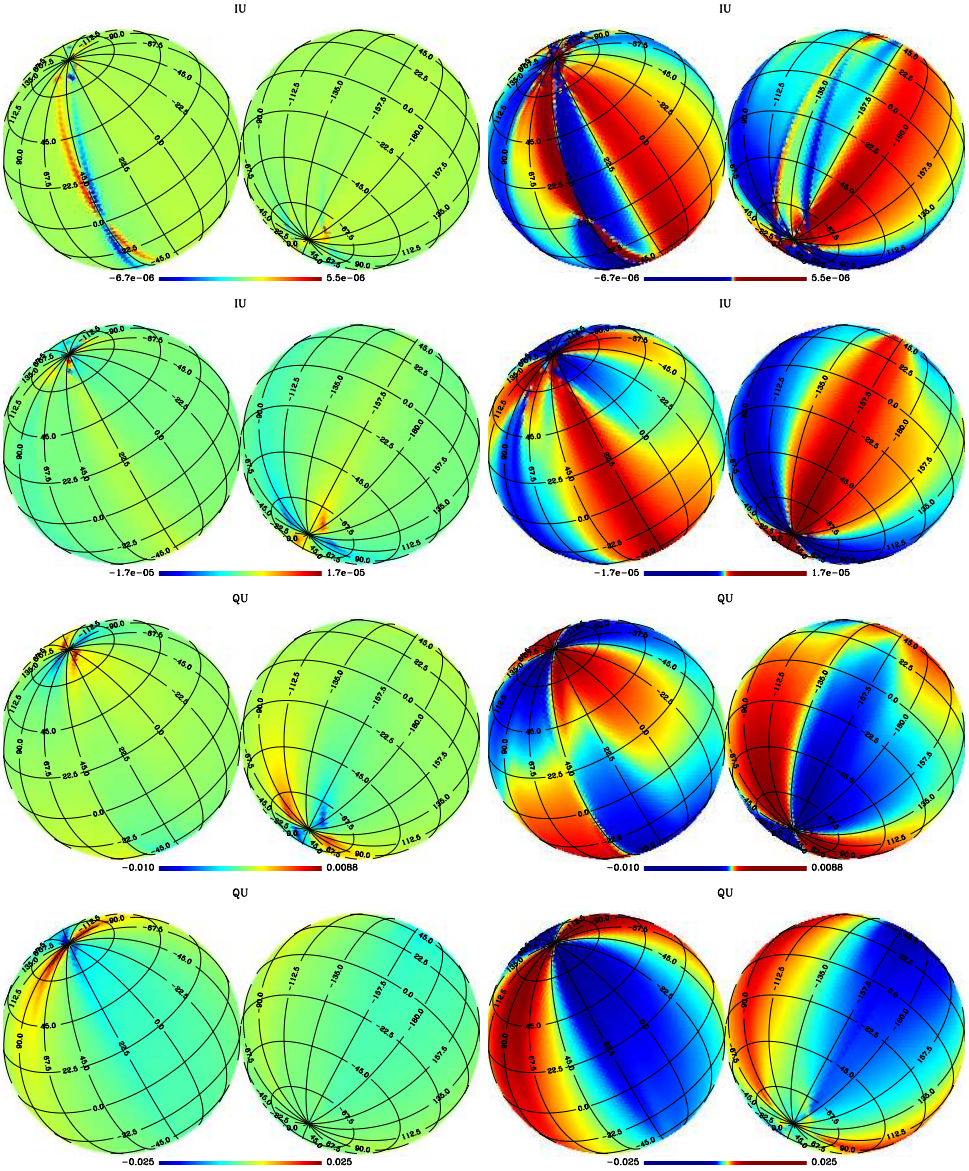


Figure 3.2: Visualization of the noise covariance matrix. *1st and 3rd row:* IU and QU noise correlation of an equatorial pixel. *2nd and 4th row:* IU and QU correlation of a polar pixel. *Left:* Linear color scale highlights the regions of strongest correlation. *Right:* Histogram equalized color scale reveals a rich pattern of correlation over the entire map. Cf. Figure 3.1 for correlation of temperature pixels.

distortion effects. We can device a downgrading operator that re-pixelizes a map into lower resolution and, when applied to a noise covariance matrix, yields the statistical noise description in the downgraded case:

$$\mathbf{D} = \left(\mathbf{P}_{\rm low}^{\rm T}\mathcal{N}_{\rm w}^{-1}\mathbf{P}_{\rm low}\right)^{-1}\mathbf{X}\left(\mathbf{P}^{\rm T}\mathcal{N}_{\rm w}^{-1}\mathbf{P}\right). \tag{3.9}$$

Matrix $\mathbf{X}$ simply sums all subpixel[3] temperatures into a superpixel. It can be formulated as

$$\mathbf{X}_{qp} = \begin{cases} 1, & p \text{ subpixel of } q \\ 0, & \text{otherwise.} \end{cases} \tag{3.10}$$

Now the low resolution maps and covariance matrices become

$$\widetilde{\mathbf{m}}_{\rm low} = \mathbf{D}\widetilde{\mathbf{m}} \quad \text{and} \quad \mathbf{N}_{\rm low} = \mathbf{D}\mathbf{N}\mathbf{D}^{\rm T}. \tag{3.11}$$

Because of the sheer size of the full resolution covariance matrix, it is not feasible to iterate to the low target resolution through huge intermediate products. For both optimal and generalized destriping map-makers it means that the downgraded covariance can only be estimated by the use of some intermediate resolution matrix (e.g. going through $\mathrm{N_{side}}$=64 to $\mathrm{N_{side}}$=32), or skipping the downgrading procedure altogether by producing a low resolution matrix directly. We found that even the direct approach produced satisfactory results.

3.5 Making use of residual noise covariance

Due to resource limitations the full statistical noise description can only be applied at large angular scales. Fortunately this is the region where noise residuals have the most prominent effect on the angular power spectrum of the map estimate.

For a cosmologist the low-ℓ region of the angular power spectra contains a host of interesting features such as the effects of integrated Sachs-Wolfe effect (ISW), reionization and the primordial gravity waves. Misestimated or neglected noise correlations in the maps increase errors and have the potential to bias the power spectrum estimates in the low-ℓ regime.

3.5.1 Noise bias

A useful tool in assessing the significance of the residual noise correlations is the noise bias. It is defined as the average of the angular power spectra of noise-only maps. Since noise and the sky are uncorrelated, the noise bias is additive to the angular power spectrum of the true sky.

One can produce Monte Carlo estimates of the noise bias by producing noise time streams, processing them into maps according to some scanning strategy, and computing the average angular power spectra of the noise maps.

[3] For HEALPix NESTED pixel ordering scheme the task of determining all subpixel pixel numbers is straightforward. One large $\mathrm{N_{side}}$=L superpixel contains $n = (H/L)^2$ $\mathrm{N_{side}}$=H subpixels. The subpixels, p, have consecutive pixel numbers that habit the range $nq \le p \le (n+1)q - 1$, where q is the superpixel number.

The noise covariance matrix, containing the full statistical description of the residual noise, provides an analytical estimate of the noise bias. Let $\mathbf{m}$ be a noise-only map, computed by applying any of the available linear map-making operations to a noise-only TOD and $\mathbf{N}$ is its covariance matrix:

$$\mathbf{N} = \langle \mathbf{m}\mathbf{m}^{\mathrm{T}} \rangle. \tag{3.12}$$

Our noise bias, $\mathbf{N}_\ell$ is the average power spectrum of these noise maps. By theory of spherical harmonic transforms we can write the spherical harmonic expansion (the $a_{\ell m}$ coefficients) using yet another linear operator[4],

$$\mathbf{a} = \mathbf{Y}\mathbf{m}, \tag{3.13}$$

and recalling the definition of the angular power spectrum then leads to

$$\mathbf{N}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle \mathbf{a}_{\ell m} \mathbf{a}^*_{\ell m} \rangle = \frac{1}{2\ell+1} \sum_i \left(\mathbf{Y}\mathbf{N}\mathbf{Y}^\dagger \right)_{ii}, \tag{3.14}$$

where the rather symbolic sum over i implies a sum over proper diagonal elements.

3.5.2 Large scale power spectrum estimation

Current paradigms of the CMB power spectrum estimation split the problem into low and high-ℓ regimes that overlap around $\ell \sim 100$. Low-ℓ analysis is driven to produce maximum likelihood estimates making use of all the available information, whereas the high-ℓ estimates employ a Gaussian or similar approximation of the C_ℓ likelihoods. Also, the noise is modelled as pure white noise at these small angular scales. The high-ℓ approximations are known to be accurate at their region of application, but break down as the angular scale increases.

This thesis deals with the distinctly non-white residual noise correlations of the low-ℓ regime. Therefore it is necessary to understand how the statistical noise description is put into use in the optimal power spectrum estimation.

Isotropy of the CMB signal means that all statistical Gaussian information about the CMB anisotropies can be conveyed as an angular power spectrum C_ℓ (c.f. Eq. (1.19)). It follows that a mapping can be found between the power spectrum and the signal pixel-pixel covariance matrix ($\mathbf{m}$ is the sky map):

$$\mathbf{S} = \langle \mathbf{m}\mathbf{m}^{\mathrm{T}} \rangle = \sum_\ell \mathbf{P}^\ell C_\ell. \tag{3.15}$$

The form of $\mathbf{P}^\ell$ depends on whether we are dealing with polarization maps or not. For the temperature-only case the matrix is written using Legendre polynomials P_ℓ:

$$\mathbf{P}^\ell_{qp} \equiv \frac{2\ell+1}{4\pi} P_\ell(\hat{\mathbf{r}}_p \cdot \hat{\mathbf{r}}_q), \tag{3.16}$$

[4] Matrix $\mathbf{Y}$ has a row for every $a_{\ell m}$, $a^{\mathrm{E}}_{\ell m}$ and $a^{\mathrm{B}}_{\ell m}$ coefficient and each row is the map of the corresponding (spin-weighted) spherical harmonic function.

where $\widehat{\mathbf{r}}_x$ is a unit vector pointing to the center of pixel x.

With both noise and sky covariance matrices we may then write a likelihood function based on the multivariate Gaussian approximation:

$$\log \mathcal{L}(C_\ell) = -\frac{1}{2}\widetilde{\mathbf{m}}^{\mathrm{T}} \left(\mathbf{S}+\mathbf{N}\right)^{-1} \widetilde{\mathbf{m}} - \frac{1}{2}\log\left[(2\pi)^{N_{\mathrm{pix}}} \det\left(\mathbf{S}+\mathbf{N}\right)\right]. \tag{3.17}$$

For power spectrum estimation we take $\mathbf{S} = \mathbf{S}(C_\ell)$ where the free parameters are all C_ℓ up to some moderate ℓ_{max} no larger than $3\mathrm{N}_{\mathrm{side}}$ of our covariance matrices. The same expression can be used in low resolution parameter estimation, but now the set of free parameters is defined by cosmology and the angular power spectrum is computed from them using numerical codes such as CAMB. In both cases the procedure is simply to maximize the likelihood by any means available and possibly characterize the isocurves of the likelihood in the surrounding parameter space. Paper **III** deals with the parameter estimation problem, although mostly in the high-ℓ regime.

Chapter 4

Isocurvature models

In a CMB experiment, parameter estimation is one of the very last steps the data flows through and likely the first step where particular cosmological models are applied to the data. Inputs to this phase are the estimated, component separated CMB map, power spectrum estimates and the residual noise covariance matrix, now hosting residual information from both map-making and component separation. The task is to choose a cosmological model defined by a handful of parameters and optimize those parameters to model the inputs (maps and C_ℓ) to the best possible agreement. It is possible to use additional cosmological datasets to break parameter degeneracies and wield additional constraining power (c.f. Section 1.3 for a review).

In this chapter we study the parameters of a particular cosmological model that slightly extends the standard cosmological model by relaxing the initial conditions.

4.1 Cosmological parameter estimation

A CMB experiment can be considered successful after raw data is processed into the frequency maps and the CMB component of these maps is resolved with reliable understanding of all residuals present in the CMB sky estimate. Even the most accurate CMB sky map does little to promote cosmological understanding without careful comparison of theoretical predictions and the observed sky.

An extremely useful tool in comparing theory and observation in cosmology has been the angular power spectrum. We possess the tools to develop initial conditions set at the inflationary epoch, only small fraction of a second since the Big Bang, into a pattern of angular correlation in the CMB sky of the present, 14 billion years later.

Parametrizing the initial conditions with a handful of parameters allows us to determine the values of these parameters with variable precision following from their impact on the correlation pattern of the sky. The parameter estimation itself provides little information about whether or not a chosen parametrization reflects the physical setting of the Universe as long as the parametrization has *sufficient* degrees of freedom to model the initial conditions.

Characterizing the primordial curvature perturbations by a featureless power spectrum like

$$\mathcal{P}_{\mathcal{R}}(k) = A\left(\frac{k}{k_{\mathrm{p}}}\right)^{n_{\mathrm{s}}-1} \tag{4.1}$$

is an example of a parametrization that is not directly physical. For most inflationary scenarios both A and n_{s} can be calculated from actual physical quantities such as the form of inflaton potential. In a sense, the parametrization (4.1) covers a wide class of inflationary scenarios that produces relatively featureless perturbation spectrum, i.e.:

$$\left|\frac{\mathrm{d}n_{\mathrm{s}}}{\mathrm{d}\ln k}\right| \ll 1. \tag{4.2}$$

4.1.1 Bayesian inference

Cosmological parameter estimation is based on the Bayes' Theorem,

$$P(T|D) = \frac{P(D|T)P(T)}{P(D)}, \tag{4.3}$$

that gives the *conditional probability* of a random event T, *given that* D occurs. It features an important deviation from the more traditional frequentist statistics in the form of *prior probabilities* $P(T)$ and $P(D)$. For our purposes we take T to be some cosmological theory with specific parameter values and D as the observed cosmological data.

In cosmological parameter estimation the model is fixed and one maximizes the probability (4.3) to determine the best fitting parameter values and their confidence limits.

In cosmological model selection the models themselves vary and the degree at which a certain model fits the data is compared. Bayesian statistics defines the *evidence* of a model as

$$E = \frac{1}{V}\int \mathrm{d}\boldsymbol{\theta}\, P(T(\boldsymbol{\theta})|D), \tag{4.4}$$

where the integral is over the allowed parameter space and V is the multidimensional volume of that space:

$$V = \int \mathrm{d}\boldsymbol{\theta} P(T(\boldsymbol{\theta})). \tag{4.5}$$

The evidence is very dependent on the chosen priors. Many cosmologists refrain from using anything but the most conservative priors, giving its interpretation more than a sense of controversy. Be as it may, model selection by comparing evidence facilitates with ease the penalty of unnecessary parameters. Another widely used means of model selection is comparison of the best fit probabilities. Unlike Bayesian comparison these methods need to be augmented by some penalty function usually in the form of higher expectation value for the probability.

4.1.2 Markov chain Monte Carlo

A key ingredient in evaluating the probability in Eq. (4.3) is calculating the probability of the observed data, given some model and its parameter values. If we take $\mathbf{d}$ to be a vector of observed data points, e.g. C_ℓ amplitudes, we assume that they contain some degree of uncertainty. This uncertainty (noise) is quantified by a covariance matrix $\mathbf{N}$. Now we may compute the χ^2 statistic of the observed data as:

$$\chi^2 = (\mathbf{d}-\mathbf{t})^{\mathrm{T}}\,\mathbf{N}^{-1}\,(\mathbf{d}-\mathbf{t})\,, \tag{4.6}$$

where $\mathbf{t}$ contains the theoretical predictions of the corresponding amplitudes. Under the null hypothesis our theory is correct and the difference $\mathbf{d}-\mathbf{t}$ is a realization of noise accurately described by the noise covariance matrix. If noise is sufficiently Gaussian, then the expectation value of the χ^2 statistic is the number of data points[1].

We use the multivariate Gaussian probability to write

$$P(D|T) = \left[(2\pi)^{n_{\mathrm{data}}}\det\mathbf{N}\right]^{-1/2}\exp\left(-\frac{1}{2}\,(\mathbf{d}-\mathbf{t})^{\mathrm{T}}\,\mathbf{N}^{-1}\,(\mathbf{d}-\mathbf{t})\right). \tag{4.7}$$

For most purposes the normalization of the probability is irrelevant and it suffices to write

$$P(D|T) \propto \mathcal{L}_D(T) = e^{-\frac{1}{2}\chi^2}, \tag{4.8}$$

where we have defined the *likelihood function* $\mathcal{L}_D$. The maximum probability and likelihood occur for the same values of model parameters and the same applies also for the log-likelihood function:

$$\log\mathcal{L}_D(T) = -\frac{1}{2}\chi^2. \tag{4.9}$$

We note that the maximal likelihood coincides with the minimum χ^2. Had we chosen to keep the normalization constant in the definition of the likelihood function, it would have appeared here as an unnecessary additive constant.

The probability distribution in (4.7) contains much more information than just the best fit values of the model parameters. In fact, it provides us with a joint probability distribution of all the model parameters assuming the null hypothesis. Extracting that information can be costly in terms of computing resources. That on the account of both number of parameters and cost of producing theoretical estimates. Instead of simple gridding of the distribution a cosmologist resorts to a Markov chain Monte Carlo (MCMC) that scales linearly with the number of parameters.

A particular instance of the MCMC method is the Metropolis-Hastings algorithm:

1 Pick a random starting point $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ in the allowed parameter space.

2 Draw a random (Gaussian) step from current location into a new location $\boldsymbol{\theta}'$.

3 Evaluate the likelihood ratio $r = \mathcal{L}(\boldsymbol{\theta}')/\mathcal{L}(\boldsymbol{\theta})$ and assume the new position with probability r. Return to 2.

[1] The statement is easy to understand by writing the noise realization as $\mathbf{N}^{1/2}\mathbf{u}$ where $\mathbf{u}$ is a vector of uniform uncorrelated random numbers. Then $\langle\chi^2\rangle = \langle\mathbf{u}^{\mathrm{T}}\mathbf{u}\rangle = \sum_n 1$.

It has been shown that after an initial burn-in period the samples $\boldsymbol{\theta}$ in the chain follow the probability distribution defined by (4.8). A great deal of care is required in assessing whether or not a chain or a set of chains have actually converged to an equilibrium distribution. One of the often encountered measures of convergence is the Gelman-Rubin R statistic [68] that for the multivariate case is [69]:

$$\widehat{R}^{\mathrm{p}} = \max_{\mathbf{a}} \frac{\mathbf{a}^{\mathrm{T}}\widehat{\mathbf{V}}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\mathbf{W}\mathbf{a}}, \tag{4.10}$$

where $\widehat{\mathbf{V}}$ is a pooled estimate of the parameter covariance matrix and $\mathbf{W}$ is an averaged estimate of the same quantity. Essentially $\widehat{R}^{\mathrm{p}}$ is a distance measure of the two matrices and approaches unity as the chains mix and begin to provide converged estimates. It should be noted that since convergence is evaluated by measuring chain mixing, the statistic is blind to completely unsampled modes of the parameter space. A popular implementation of the MCMC sampler is COSMOMC[2] [70].

At equilibrium, the regions of equal probability have approximately the same amount of samples drawn from them. If the run parameters are chosen poorly or the chain starting point distribution is not dispersed enough, the chains may spend considerable time sampling from a single mode of the distribution neglecting all other regions of acceptable parameter values. An important factor in speeding up the convergence of an MCMC run is using an estimated parameter-parameter covariance matrix to find the real non-degenerate sampling directions and to adjust the step size so that the parameter space is traversed at a feasible pace.

4.2 Isocurvature degrees of freedom

When the primordial curvature perturbations are characterized by a power spectrum, a common approximation is to assume zero entropy perturbation. These adiabatic initial conditions occur naturally when a single inflaton field is solely responsible for the Universe's energy density and decays into all species of energy: photons, neutrinos, dark and baryonic matter, relevant for early evolution of the Universe. Thus the species all inherit the same energy density perturbations. If one does not wish to limit to single field inflationary models then the possibility of entropy perturbations arises.

The term *isocurvature* refers to the fact that instead of zero entropy perturbation one may keep the curvature perturbation constant and make use of the additional entropy degrees of freedom to alter individual species' energy density perturbation.

The comoving curvature perturbation (1.43) in terms of the conformal Newtonian gauge metric perturbation is

$$\mathcal{R} = -\Psi - \frac{2}{3(1+w)}\left(\frac{1}{\mathcal{H}}\cdot\frac{\partial\Psi}{\partial t} + \Psi\right), \tag{4.11}$$

and we define the entropy perturbation between photons and cold dark matter as

$$\mathcal{S} = \frac{\delta_{\mathrm{c}}}{1+w_{\mathrm{c}}} - \frac{\delta_\gamma}{1+w_\gamma} = \delta_{\mathrm{c}} - \frac{3}{4}\delta_\gamma. \tag{4.12}$$

[2] `http://cosmologist.info/cosmomc`

Although a number of other pairs of energy species can be considered, we limit the study to this particular case.

We define the adiabatic mode as $\mathcal{S} = 0$ and the isocurvature mode as $\mathcal{R} = 0$. The separation is defined some time after inflation but well before the cosmological scales enter the horizon (cf. Fig. 1.1 for evolution of scales).

During inflation, the entropy perturbation evolves and may seed another adiabatic mode. In total, we end up having four correlation spectra instead of the usual adiabatic one: primordial adiabatic, primordial isocurvature, and secondary adiabatic autocorrelations, and an isocurvature-adiabatic cross correlation. We take the primordial modes to be uncorrelated. Thus we define our initial conditions by

$$\mathcal{P}_\mathcal{R}(t,k) = A_\mathrm{r}^2 \left(\frac{k}{k_\mathrm{p}}\right)^{n_\mathrm{adi1}} + A_\mathrm{s}^2 \left(\frac{k}{k_\mathrm{p}}\right)^{n_\mathrm{adi2}}, \tag{4.13}$$

$$\mathcal{P}_\mathcal{S}(t,k) = B^2 \left(\frac{k}{k_\mathrm{p}}\right)^{n_\mathrm{iso1}} \quad \text{and} \tag{4.14}$$

$$\mathcal{C}_{\mathcal{R}\mathcal{S}}(t,k) = A_\mathrm{s} B \left(\frac{k}{k_\mathrm{p}}\right)^{n_\mathrm{cor}}, \tag{4.15}$$

and evolve the perturbation equations from beyond the horizon into CMB photon density correlation at present. Here, the correlated spectral index n_cor is the simple average $(n_\mathrm{adi2} + n_\mathrm{iso})/2$.

Instead of absolute amplitudes A_r, A_s and B it is useful to define relative amplitudes that explicitly contain the isocurvature fraction of total power:

$$\alpha = \frac{B^2}{A_\mathrm{r}^2 + A_\mathrm{s}^2 + B^2}, \quad \text{and} \quad \gamma = \mathrm{sign}(A_\mathrm{s} B) \frac{A_\mathrm{s}^2}{A_\mathrm{r}^2 + A_\mathrm{s}^2}, \tag{4.16}$$

where α is the fraction of pure isocurvature power at the reference scale k_p and γ is the correlated fraction of adiabatic power at the same scale.

4.2.1 Parametrization

Great care is required when physical degrees of freedom are parametrized for an MCMC study. In order to be as conservative as possible a cosmologist often assigns only wide uniform probabilities to his chosen primary set of parameters. Top hat prior to one parameter is usually something completely different for a derived parameter. A derived parameter is a quantity fixed by the values of the primary parameters in a model. Relation of likelihoods between two alternate parametrizations, $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, is given by

$$g(\boldsymbol{\phi}) = f(\boldsymbol{\theta})|J|, \quad \text{where} \quad |J| = \begin{vmatrix} \frac{\partial \boldsymbol{\theta}_1}{\partial \boldsymbol{\phi}_1} & \frac{\partial \boldsymbol{\theta}_1}{\partial \boldsymbol{\phi}_2} & \cdots \\ \frac{\partial \boldsymbol{\theta}_2}{\partial \boldsymbol{\phi}_1} & \frac{\partial \boldsymbol{\theta}_2}{\partial \boldsymbol{\phi}_2} & \cdots \\ \vdots & \vdots & \ddots \end{vmatrix}, \tag{4.17}$$

and g and f are probability density functions (pdf). There must exist a one-to-one mapping between the two sets of parameters.

Often it is hard to determine which is physically the primary and which is derived. Parametrization effects are strongest when the data itself does not constrain the parameters sufficiently. One must pay extra attention to parametrization priors when the prior likelihood is distinctly non-flat within the allowed region of the derived parameter.

Figure 4.1 contains a case study of data drawn from simple $y = x$+noise model. Two alternate parametrizations are explored and the parametrization induced priors are shown.

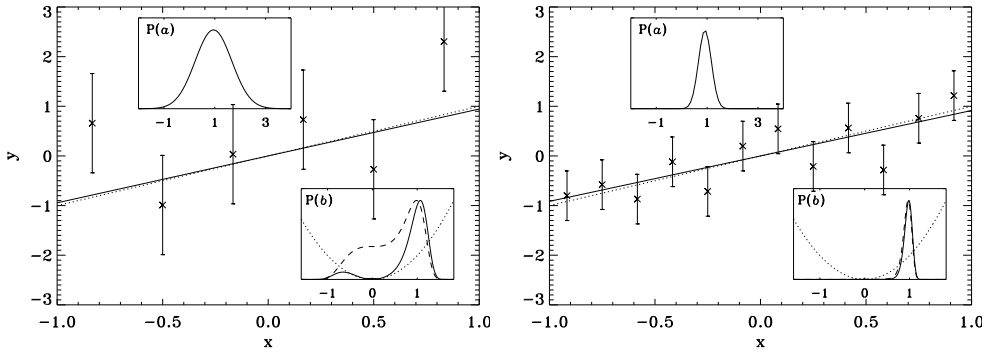


Figure 4.1: Parametrization effects in parameter fitting. *Left:* sparse data. *Right:* High precision data. The insets depict likelihoods of two one parameter models, $y = ax = \sqrt[3]{b}x$, computed from the shown data points. The dotted line for $P(b)$ is the parametrization induced prior and the dashed line is the likelihood of b using a flat prior on b. Conclusive data negates the parametrization effects.

An obvious choice for parameters of the additional isocurvature degrees of freedom would be to use the new amplitudes and spectral indices, $A_{\rm s}$, B, $n_{\rm adi2}$ and $n_{\rm iso}$. That approach, however, bears significant shortcomings. One is forced to choose a pivot scale $k_{\rm p}$ that fixes the point of reference for the amplitudes. Different choice of the pivot scale induces non-flat priors to amplitudes defined at another scale. Take for example single spectrum defined by $\boldsymbol{\theta} = (A^2, n)$ and another parametrization $\boldsymbol{\phi} = (B^2, n)$, related by a mapping $A^2 = B^2(k_{\rm A}/k_{\rm B})^{n-1}$. The probability densities become

$$g(B^2, n) = f(A^2, n)\left(\frac{k_{\rm A}}{k_{\rm B}}\right)^{n-1} \quad \Rightarrow \quad \mathcal{L}(B^2) \propto \left(\frac{k_{\rm A}}{k_{\rm B}}\right)^{n}. \tag{4.18}$$

If the spectral index is well constrained by the data, this poses little problem for amplitude determination. With current data, such is not the case especially for the isocurvature degrees of freedom. In Paper **III** we explored a novel parametrization where the primordial spectra are defined not in just one but two separate reference scales. The spectral index of the spectrum follows from the two reference amplitudes and is not a primary parameter. Reference scales still need to fixed but their actual values induce less extreme parametrization priors. Let $\boldsymbol{\theta} = (A^2, n, k_{\rm A})$ define the original spectral index parametrization, then our new amplitude parametrization will be $\boldsymbol{\phi} = (B^2, C^2, k_{\rm B}, k_{\rm C})$,

where

$$B^2 = \mathcal{P}(k_{\rm B}) = A^2 \left(\frac{k_{\rm B}}{k_{\rm A}}\right)^{n-1} \quad \text{and} \quad C^2 = \mathcal{P}(k_{\rm C}) = A^2 \left(\frac{k_{\rm C}}{k_{\rm A}}\right)^{n-1}, \tag{4.19}$$

and the mapping can be inverted to give

$$n - 1 = \frac{\log\left(B^2/C^2\right)}{\log\left(k_{\rm B}/k_{\rm C}\right)} \quad \text{and} \quad A^2 = B^2 \left(\frac{k_{\rm A}}{k_{\rm B}}\right)^{\log\left(B^2/C^2\right)/\log(k_{\rm B}/k_{\rm C})}. \tag{4.20}$$

To study the parametrization priors we can take two sets of amplitude parameters and the corresponding pivot scales, (A, B, k_A, k_B) and (C, D, k_C, k_D). After computing the Jacobian (4.17), we find

$$g(C^2, D^2) = f(A^2, B^2) \cdot 2\frac{\log\left(k_{\rm A}/k_{\rm B}\right)}{\log\left(k_{\rm C}/k_{\rm D}\right)} \left(\frac{k_{\rm A}k_{\rm B}}{k_{\rm C}k_{\rm D}}\right)^{\log\left(A^2/B^2\right)/\log(k_{\rm A}/k_{\rm B})}, \tag{4.21}$$

a manifestly invariant probability density at many relevant limits such as $k_{\rm A}k_{\rm B} \approx k_{\rm C}k_{\rm D}$, $A^2 \approx B^2$ and $|\log\left(k_{\rm A}/k_{\rm B}\right)| \gg 1$.

Amplitude and spectral index parametrizations imply distinctly non-flat priors to each other. We took great care to analyze this effect and found perceptible changes in the primary and derived parameter likelihoods when employing the different parametrizations. However, the changes were always well within the inferred confidence limits.

4.3 Results and implications for cosmology

Pure isocurvature perturbations were shown to be inconsistent with the CMB early on [71, 72, 73]. After that a considerable amount of work [e.g. 74, 75, 76, 77, 78, 79, to name a few] has been exerted to find stringent upper limits to the amount of primordial isocurvature perturbations. A reliable detection of even a trace amount of isocurvature would have an impact on the realm of inflationary models since such initial conditions do not arise naturally in single field (inflaton) inflationary models. Fortunately, multi-field inflationary models are subject to active research [80, 81, 82].

In Paper **III** we combined CMB and LSS data to infer posterior likelihoods of our six perturbation amplitude parameters $(A_1, A_2, \alpha_1, \alpha_2, \gamma_1, \gamma_2)$ and four background parameters, the baryon and cold dark matter densities, ratio of the sound horizon and angular diameter distance and the optical depth to reionization ($\omega_{\rm b}$, $\omega_{\rm c}$, θ and τ respectively). Following [83] we evaluated the simulated Markovian chains to estimate the marginalized posterior likelihood of the total non-adiabatic fraction of present CMB temperature fluctuations,

$$\alpha_{\rm T} = \frac{\sum_{\ell=2}^{\ell_{\rm max}} (2\ell+1)\left(C_\ell^{\rm TTiso} + C_\ell^{\rm TTcor}\right)}{\sum_{\ell=2}^{\ell_{\rm max}} (2\ell+1)\, C_\ell^{\rm TT}}, \tag{4.22}$$

and found indications of isocurvature modes. The resulting posterior likelihood for $\alpha_{\rm T}$ is reproduced in Figure 4.2. In our studies we found the data to prefer a non-zero isocurvature contribution by more than two sigma confidence.

There are, however, reasons why we do not take this as a conclusive detection of primordial isocurvature mode. Firstly, the slight deviations from adiabatic model could be accommodated by more complex structure of the primordial adiabatic spectrum. The argument was explored by the WMAP team who found some indications of a more complex primordial spectrum [84]. Secondly, the detection is based to considerable extent on the TT power spectrum around the third acoustic peak. Observations of that region of the spectrum do not allow percent level differences to be used as basis of conclusive model selection. Therefore better data is called for.

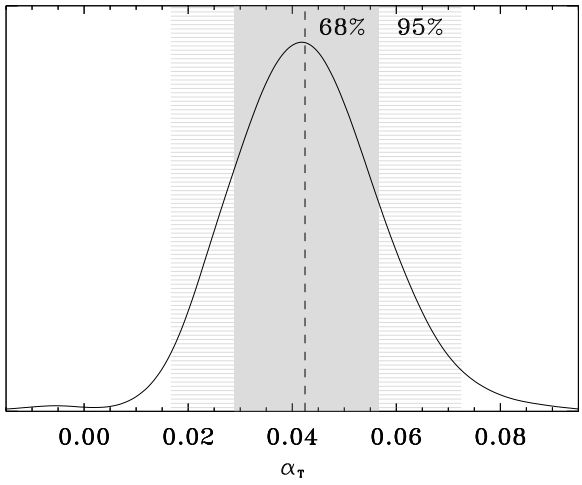


Figure 4.2: Our marginalized posterior likelihood for total non-adiabatic contribution to CMB temperature fluctuations, α_{T}. It is non-zero by two sigma.

Chapter 5

Conclusions

Along with increasingly precise measurements, the need for well validated, approximate analysis tools becomes imminent. The destriping approach with its derivants offer a very promising, scalable solution to the map-making problem in the context of multi-detector, high resolution, CMB surveys. The linearity of the destriping method makes it transparent how different signal components propagate and how they are affected by the underlying approximations. An especially attractive feature of the variable baseline length, generalized destriper is its ability to produce results, almost arbitrarily close to the optimal method, when sufficient resources are available and the noise power spectrum is sufficiently white at high frequencies.

A quick and efficient map-making tool as itself is not enough. In order to derive reliable estimates of the angular power spectrum or cosmological parameters from the maps, there must exist a model of residual noise. We studied and compared three different map-making methods and the associated residual noise covariance matrices, and found three very different levels of both the computational resource consumption and accuracy of the noise model. The methods considered were an optimal map-maker and a classical and a generalized destriper. Not unexpectedly, we showed that the optimal method will produce the most accurate correspondence between the produced maps and the residual noise covariance matrix designed to describe them. However, that precision comes at a considerable resource cost, making it unfeasible to apply the tool on every-day basis to large datasets.

Our noise covariance matrix for the generalized destriper shows great promise because of its generality: it can be used both at a high resource, high accuracy mode and the opposite, effectively reproducing both the optimal map-maker and standard destriper results. Nevertheless, it must be remembered that the estimated noise covariance is only as good as the destriping approximation that states that all correlations between noise samples can be modelled by the constant baseline offsets of fixed length. Such a noise model can not cover high frequency correlated noise to arbitrary precision as it ultimately just falls back to the optimal noise model (offsets of single samples). It follows that, provided that the correlated noise knee frequency is low enough and there is no significant high frequency correlation, the generalized destriper is a sound choice

for the map-making tool and the corresponding noise covariance evaluation.

When conducting an experiment as large as PLANCK, it pays to consider that it will not be remembered by the maps, catalogs and power spectra that it produces but by the science done, based on those products. Nevertheless, all the exciting science will need to be based on robust analysis and refinement of the observations. PLANCK has all the potential to discover many long sought features in the CMB such as isocurvature perturbations, primordial gravitational waves and non-gaussianity. A reliable detection of any of those features would certainly compare to all the past cosmological discoveries since these signals originate all the way from the inflationary era.

Bibliography

[1] Edwin Hubble. A relation between distance and radial velocity among extra–galactic nebulae. *Proc. Nat. Acad. Sci.*, 15:168–173, 1929.

[2] Georges Lemaître. Un Univers homogéne de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales Soc.Sci.Brux.*, 47:49, 1927.

[3] Albert Einstein. Die Feldgleichungen der Gravitation. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915:844–847, 1915.

[4] Albert Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1917:142–152, 1917.

[5] Alexander A. Friedmann. Über die Krümmung des Raumes. *Z. Phys. A*, 10:377–386, 1922.

[6] Georges Lemaître. The expanding universe. *Mon. Not. Roy. Astron. Soc.*, 91:490–501, 1931.

[7] Howard P. Robertson. Kinematics and world-structure. *Astrophys. J.*, 82:284, nov 1935.

[8] Howard P. Robertson. Kinematics and world-structure ii. *Astrophys. J.*, 83:187, apr 1936.

[9] Howard P. Robertson. Kinematics and world-structure iii. *Astrophys. J.*, 83:257, may 1936.

[10] Fred Hoyle. A new model for the expanding universe. *Mon. Not. Roy. Astron. Soc.*, 108:372–382, 1948.

[11] Ralph A. Alpher, Hans Bethe, and George Gamow. The origin of chemical elements. *Phys. Rev.*, 73(7):803–804, 1948.

[12] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic black-body radiation. *Astrophys. J.*, 142:414–419, 1965.

[13] Arno A. Penzias and Robert W. Wilson. A Measurement of excess antenna temperature at 4080 Mc/s. *Astrophys. J.*, 142:419–421, 1965.

[14] E. Komatsu et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys. J. Suppl.*, 180:330–376, 2009.

[15] Alan H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. *Phys. Rev.*, D23:347–356, 1981.

[16] Alexei A. Starobinsky. A new type of isotropic cosmological model without singularity. *Phys. Lett.*, B91:99–102, 1980.

[17] Sean Carroll. *Spacetime and Geometry: An Introduction to General Relativity.* Addison Wesley, 2004.

[18] Viatcheslav F. Mukhanov, Hume A. Feldman, and Robert H. Brandenberger. Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions. *Phys. Rept.*, 215:203–333, 1992.

[19] Wayne Hu and Martin J. White. A CMB Polarization Primer. *New Astron.*, 2:323, 1997.

[20] Wayne Hu. CMB temperature and polarization anisotropy fundamentals. *Ann. Phys.*, 303:203–225, 2003.

[21] K. M. Gorski et al. The healpix primer. *ArXiv e-prints*, 1999. astro-ph/9905275.

[22] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient Computation of CMB anisotropies in closed FRW models. *Astrophys. J.*, 538:473–476, 2000.

[23] James G. Bartlett, Alain Blanchard, Joseph Silk, and Michael S. Turner. The Case for a Hubble constant of 30-kms**-1/Mpc**-1. *Science*, 267:980–983, 1995.

[24] Allan Sandage. H0 = 43 +/- 11 km/s/Mpc based on angular diameters of high-luminosity field spiral galaxies. *Astrophys. J.*, 402:3–14, January 1993.

[25] W. L. Freedman et al. Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant. *Astrophys. J.*, 553:47–72, 2001.

[26] Gustav A. Tammann, Allan Sandage, and Bernd Reindl. The expansion field: The value of H_0. *Astron. Astrophys. Rev.*, 15:289–331, 2008.

[27] Gennaro Miele and Ofelia Pisanti. Primordial Nucleosynthesis: an updated comparison of observational light nuclei abundances with theoretical predictions. *ArXiv e-prints*, 2008. 0811.4479.

[28] E. Hawkins et al. The 2dF Galaxy Redshift Survey: correlation functions, peculiar velocities and the matter density of the Universe. *Mon. Not. Roy. Astron. Soc.*, 346:78, 2003.

[29] Max Tegmark et al. Cosmological Constraints from the SDSS Luminous Red Galaxies. *Phys. Rev.*, D74:123507, 2006.

[30] P. James E. Peebles and J. T. Yu. Primeval adiabatic perturbation in an expanding universe. *Astrophys. J.*, 162:815–836, 1970.

[31] Nick Kaiser. Weak Lensing and Cosmology. *Astrophys. J.*, 498:26, 1998.

[32] Wayne Hu and Max Tegmark. Weak Lensing: Prospects for Measuring Cosmological Parameters. *Astrophys. J.*, 514:L65–L68, 1999.

[33] Matthias Bartelmann and Peter Schneider. Weak Gravitational Lensing. *Phys. Rept.*, 340:291–472, 2001.

[34] Brian P. Schmidt et al. The High-Z Supernova Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type Ia Supernovae. *Astrophys. J.*, 507:46–63, 1998.

[35] Adam G. Riess et al. Type Ia Supernova Discoveries at z>1 From the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution. *Astrophys. J.*, 607:665–687, 2004.

[36] P. McDonald et al. The Linear Theory Power Spectrum from the Lyman-alpha Forest in the Sloan Digital Sky Survey. *Astrophys. J.*, 635:761–783, 2005.

[37] Uros Seljak, Anze Slosar, and Patrick McDonald. Cosmological parameters from combining the Lyman-alpha forest with CMB, galaxy clustering and SN constraints. *JCAP*, 0610:014, 2006.

[38] Julien Lesgourgues, Alexei A. Starobinsky, and Wessel Valkenburg. What do WMAP and SDSS really tell about inflation? *JCAP*, 0801:010, 2008.

[39] Scott Dodelson. *Modern Cosmology: Anisotropies and Inhomogeneities in the Universe*. Academic Press, second edition, 2003.

[40] Hannu Kurki-Suonio. Cosmological perturbation theory. Electonically available at `http://theory.physics.helsinki.fi/~cmb/CosPer.ps`, (Referenced on February 13th 2009), 2005.

[41] Andrew R. Liddle and David H. Lyth. *Cosmological inflation and large-scale structure*. Cambridge University Press, 2000.

[42] George Efstathiou, Steven Gratton, and Francesco Paci. Impact of Galactic polarized emission on B-mode detection at low multipoles. *ArXiv e-prints*, 2009. 0902.4803.

[43] Chung-Pei Ma and Edmund Bertschinger. Cosmological perturbation theory in the synchronous and conformal Newtonian gauges. *Astrophys. J.*, 455:7–25, 1995.

[44] Uros Seljak and Matias Zaldarriaga. A Line of Sight Approach to Cosmic Microwave Background Anisotropies. *Astrophys. J.*, 469:437–444, 1996.

[45] Sara Seager, Dimitar D. Sasselov, and Douglas Scott. How exactly did the Universe become neutral? *Astrophys. J. Suppl.*, 128:407–430, 2000.

[46] Planck Collaboration. Planck: The scientific programme. *ESA-SCI*, 1, 2005. e-print astro-ph/0604069.

[47] S. M. Leach et al. Component separation methods for the Planck mission. *Astron. Astrophys.*, 491:597–615, 2008.

[48] George F. Smoot et al. Structure in the COBE differential microwave radiometer first year maps. *Astrophys. J.*, 396:L1–L5, 1992.

[49] C. L. Bennett et al. First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. *Astrophys. J. Suppl.*, 148:1, 2003.

[50] D. Maino et al. The Planck-LFI instrument: analysis of the 1/f noise and implications for the scanning strategy. *Astron. Astrophys. Suppl. Ser.*, 140:383, 1999.

[51] Michael Seiffert et al. 1/f noise and other systematic effects in the Planck-LFI radiometers. *Astron. Astrophys.*, 391:1185–1197, 2002.

[52] Xavier Dupac and Jan Tauber. Scanning strategy for mapping the Cosmic Microwave Background anisotropies with Planck. *ArXiv e-prints*, 2004. astro-ph/0409405.

[53] K. M. Gorski et al. HEALPix – a Framework for High Resolution Discretization, and Fast Analysis of Data Distributed on the Sphere. *Astrophys. J.*, 622:759–771, 2005.

[54] Charmaine Armitage and Benjamin D. Wandelt. Deconvolution Map-Making for Cosmic Microwave Background Observations. *Phys. Rev.*, D70:123007, 2004.

[55] A. Kogut et al. Calibration and Systematic Error Analysis For the COBE-DMR Four-Year Sky Maps. *Astrophys. J.*, 470:653, 1996.

[56] Max Tegmark. How to make maps from CMB data without losing information. *Astrophys. J.*, 480:L87–L90, 1997.

[57] Theodore S. Kisner. *Boomerang 2003: Measuring the polarization of the Cosmic Microwave Background Radiation.* PhD thesis, University of California, Santa Barbara, 2008.

[58] M. A. J. Ashdown et al. Making sky maps from Planck data. *Astron. Astrophys.*, 467:761–775, 2007.

[59] M. A. J. Ashdown et al. Making Maps from Planck LFI 30GHz Data. *Astron. Astrophys.*, 471:361–380, 2007.

[60] M. A. J. Ashdown et al. Making Maps from Planck LFI 30GHz Data with Asymmetric Beams and Cooler Noise. *Astron. Astrophys.*, 493:753–783, 2009.

[61] T. Poutanen et al. Comparison of map-making algorithms for CMB experiments. *Astron. Astrophys.*, 449:1311–1322, 2006.

[62] Elina Keihänen, Hannu Kurki-Suonio, Torsti Poutanen, Davide Maino, and Carlo Burigana. A maximum likelihood approach to the destriping technique. *Astron. Astrophys.*, 428:287–298, 2004.

[63] C. Burigana et al. A preliminary study on destriping techniques of PLANCK/LFI measurements versus observational strategy. *ArXiv e-prints*, 1997. astro-ph/9906360.

[64] Jacques Delabrouille. Analysis of the accuracy of a destriping method for future cosmic microwave background mapping with the PLANCK SURVEYOR satellite. *Astron. Astrophys. Suppl. Ser.*, 127:555–567, February 1998.

[65] B. Revenu et al. Destriping of polarized data in a cmb mission with a circular scanning strategy. *Astron. Astrophys. Suppl. Ser.*, 142:499–509, March 2000.

[66] W. H. Press et al. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*. Cambridge University Press, second edition, 1992.

[67] H. K. Eriksen et al. Power spectrum estimation from high-resolution maps by Gibbs sampling. *Astrophys. J. Suppl.*, 155:227–241, 2004.

[68] Andrew Gelman and Donald B. Rubin. Inference from Iterative Simulation using Multiple Sequences. *Stat. Sci.*, 7:457–511, 1992.

[69] Stephen Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.*, 7:434–455, 1998.

[70] Antony Lewis and Sarah Bridle. Cosmological parameters from CMB and other data: a Monte Carlo approach. *Phys. Rev.*, D66, 2002.

[71] George Efstathiou and J. Richard Bond. Isocurvature cold dark matter fluctuations. *Mon. Not. Roy. Astron. Soc.*, 218:103–121, 1986.

[72] Kari Enqvist, Hannu Kurki-Suonio, and Jussi Väliviita. Limits on isocurvature fluctuations from Boomerang and MAXIMA. *Phys. Rev.*, D62:103003, 2000.

[73] Kari Enqvist, Hannu Kurki-Suonio, and Jussi Väliviita. Open and closed CDM isocurvature models contrasted with the CMB data. *Phys. Rev.*, D65:043002, 2002.

[74] Patrick Crotty, Juan Garcia-Bellido, Julien Lesgourgues, and Alain Riazuelo. Bounds on isocurvature perturbations from CMB and LSS data. *Phys. Rev. Lett.*, 91:171301, 2003.

[75] M. Bucher et al. The initial conditions of the universe: how much isocurvature is allowed? *Phys. Rev. Lett.*, 93:081301, 2004.

[76] Maria Beltran, Juan Garcia-Bellido, Julien Lesgourgues, and Matteo Viel. Squeezing the window on isocurvature modes with the Lyman- alpha forest. *Phys. Rev.*, D72:103515, 2005.

[77] Rachel Bean, Joanna Dunkley, and Elena Pierpaoli. Constraining Isocurvature Initial Conditions with WMAP 3- year data. *Phys. Rev.*, D74:063503, 2006.

[78] Roberto Trotta. The isocurvature fraction after WMAP 3-year data. *Mon. Not. Roy. Astron. Soc. Lett.*, 375:L26–L30, 2007.

[79] Masahiro Kawasaki and Toyokazu Sekiguchi. Cosmological Constraints on Isocurvature and Tensor Perturbations. *Prog. Theor. Phys.*, 120:995–1016, 2008.

[80] Andrei D. Linde and Viatcheslav F. Mukhanov. Nongaussian isocurvature perturbations from inflation. *Phys. Rev.*, D56:535–539, 1997.

[81] David H. Lyth and David Wands. Generating the curvature perturbation without an inflaton. *Phys. Lett.*, B524:5–14, 2002.

[82] Takeo Moroi and Tomo Takahashi. Effects of cosmological moduli fields on cosmic microwave background. *Phys. Lett.*, B522:215–221, 2001.

[83] Hannu Kurki-Suonio, Vesa Muhonen, and Jussi Väliviita. Correlated Primordial Perturbations in Light of CMB and LSS Data. *Phys. Rev.*, D71:063005, 2005.

[84] D. N. Spergel et al. Wilkinson Microwave Anisotropy Probe (WMAP) three year results: Implications for cosmology. *Astrophys. J. Suppl.*, 170:377, 2007.