SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Physics Area - PhD course in
Astrophysics and Cosmology

Searching for our origins in the Cosmic Microwave Background polarisation

Candidate:
Claudio Ranucci

Advisors:
Prof. Nicoletta Krachmalnicoff
Prof. Carlo Baccigalupi

Co-advisors:
Dr. Léo Vacher
Dr. Alessandro Carones

Academic Year 2025-2026





Searching for our origins in the Cosmic Microwave Background polarisation

Scuola Internazionale Superiore di Studi Avanzati (SISSA)
PhD in Astrophysics and Cosmology

Claudio Ranucci

Advisors
Prof. Nicoletta Krachmalnicoff
Prof. Carlo Baccigalupi

Co-Advisors
Dr. Léo Vacher
Dr. Alessandro Carones

Coordinator of the PhD course
Prof. Carlo Baccigalupi

Academic Year 2025/2026

Dissertation submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

Thesis defended on 22 September 2026

Searching for our origins in the Cosmic Microwave Background polarisation
PhD thesis. Scuola Internazionale di Studi Superiori Avanzati (SISSA).

© 2026 Claudio Ranucci. This thesis is released under the CC BY-NC-SA 4.0 (Creative Commons Attribution-Non Commercial-Share Alike 4.0) International license. To read a copy of the license visit https://creativecommons.org/licenses/by-nc-sa/4.0/.

This thesis has been typeset by LaTeX.

Version: September 16, 2026

Author's email: claudio.ranucci.cr@gmail.com

Abstract

This Thesis is the final product of the work I have carried out in the last three years of my PhD journey. The central topic of this manuscript is the astrophysical observable known as the Cosmic Microwave Background (CMB) radiation, one of the pillars on which our knowledge of the Universe is built. I will describe some of the huge collective effort being carried out every day by the scientific community in order to improve our understanding of the CMB and shed light on the mysteries of Nature. Nowadays, the frontier of CMB science is represented by its polarised signal, which contains crucial information about the very first phases of the history of the Universe. In particular, current and future CMB experiments are focusing on the hunt for primordial B-modes, a peculiar pattern thought to be induced by gravitational waves in the primordial plasma, possibly generated by an inflationary phase happened immediately after the birth of the Universe. Usually, this signal is parametrised through the tensor-to-scalar ratio r.

The extraction of this very faint signal requires the crafting of extremely sensitive instrumentation, with a complete control over experimental systematic, as well as a full characterisation of the contamination due to the polarised emission of our Galaxy. In this field, the instrumental state-of-art is represented by the *Simons Observatory* (*SO*), a newly built CMB experiment located in Chile which is currently taking data of the sky. I have been a member of the *SO* collaboration since 2023; in these years, I participated to the development of the data analysis pipeline, whose main task is to measure the primordial B-mode signal from the observed multi-frequency sky maps. I contributed to the implementation of cosmological datasets from other CMB experiments (*Planck*, *WMAP*, *S-PASS*) in the pipeline, in order to combine *SO* data with external probes for an improved characterisation of the Galactic foregrounds. Within the *SO* project, I have also contributed to the production of several sets of simulations, which have been used both for the investigation of instrumental systematics, and for the quantification of the uncertainties in the measurement of the statistical properties of the CMB and of cosmological parameters.

The continuous improvement in experimental sensitivity will bring the community closer to a potential detection of primordial B-modes. The measurement is extremely challenging and the contaminating contributions could bias the inference of the cosmological signal, possibly giving rise to a false detection. To avoid this scenario, we should be prepared for it with a set of procedures aimed at rejecting or confirming an eventual detection of primordial B-modes. Therefore, this kind of null tests is becoming more important every year as our experiments get better and better. In the last two years of my PhD, I worked on the development of robustness tests for the detection of the tensor-to-scalar ratio from CMB observations, to be applied

in current and future CMB experiments. I studied the possibility of using higher-order statistics, like Minkowski functionals and scattering transform statistics, to determine if a potential B-modes detection is contaminated by Galactic foregrounds or systematics. I used scattering transforms to build and provide to the community a working robustness test methodology applicable to current and future experiments, which represents a step forward in the measurement of primordial inflation.

Ringraziamenti

I miei primi ringraziamenti vanno senza dubbio ai miei supervisori, Nicoletta e Carlo, senza i quali non ci sarebbe stato nessun dottorato. Gli sarò sempre immensamente riconoscente per avermi accolto in SISSA e trattato alla pari fin dal primo giorno, permettendomi di fare questo lavoro nel miglior modo possibile. Grazie a Nico per avermi insegnato cosa vuol dire sporcarsi le mani per fare veramente scienza, e per essere stata anche la mia psicologa, aiutandomi a capire cosa fare da grande. Grazie a Carlo per avermi insegnato i trucchi del mestiere e come affrontare le sfide della vita con entusiasmo. E grazie ad entrambi per avermi mostrato che fare scienza è molto di più che stare davanti a un computer. Un grazie particolare a Nico, che mi ha permesso di stare vicino alla mia famiglia quando ce n'era bisogno.

From the bottom of my heart, I want to thank my older brothers (and co-supervisors) Léo and Alessandro. I really don't know where would I be without them. They started as simple collaborators and later became part of my family, helping, sustaining, and also protecting me in the most difficult moments of my PhD. They were able to transmit me the passion for this job and the joy of doing science together, achieving results that I could not even approach on my own. I will forever cherish the memories we shared in Trieste and around the world.

I also want to thank all the people I crossed paths with in the SO *collaboration, and in particular the BB Analysis Working Group. I had the opportunity to work with many great scientists, which taught me what it is like to work all together as human beings to achieve historical discoveries for mankind, feeling part of something bigger.*

Special thanks are reserved for the CMB group at IJCLab. Matthieu, Sophie, Thibaut, Merry, Clément, and everyone else. They welcomed me in their group in Orsay, helping and supporting me for the last months of my PhD, while I was trying to finish projects and write the Thesis, in an extremely stimulating environment.

Heartfelt thanks go to my collaborators in Paris, Erwan and Sébastien. They allowed me to work with them at ENS as a colleague, explaining me the secrets of scattering transforms over and over again. Thanks also to Alexandros, who was always available for answering my multiple naïve questions and doubts on the topic. I really want to thank Erwan, who believed in me at the point of offering the opportunity of a lifetime. It was incredibly important for me, and I will never forget it.

I also want to thank Erwan and Giuseppe for accepting the role of reviewers of the present Thesis. I really appreciate the time they spent in reading it, and their comments and suggestions were able to improve the final version of the manuscript.

Un grazie speciale ai miei compagni di avventura di questi quattro anni in SISSA: Carlos, Emma, Florie, Francesco, Francescopaolo, Lorenzo, Martina. Per tutti i

momenti in cui ci siamo sostenuti a vicenda, nei successi e nelle depressioni. Per tutte le serate insieme e per le sagre friulane.

Il mio ringraziamento più grande va alla mia famiglia, che fin da piccolo mi ha sempre sostenuto in ogni cosa, e che mi ha sempre spronato a raggiungere i miei obiettivi, per quanto assurdi e lontani. Dal più profondo del mio cuore, grazie ai miei Genitori, per avermi dato tutto ciò che potevano, e anche di più. Per tutti i sacrifici fatti per non farmi mai mancare nulla, e per aprirmi tutte le strade possibili. Per avermi cresciuto fino a diventare la persona che sono felice di essere oggi. Questo traguardo è anche e soprattutto merito loro, e questa Tesi è a loro dedicata.

Ed infine, alla persona più importante, a mia Moglie. Che dopo più di dieci anni, è ancora al mio fianco come il primo giorno, disposta ad abbandonare tutto e tutti per cambiare vita e sostenere i miei progetti. Grazie per aver creduto in me dal primo istante, e per avermi sostenuto nei momenti più duri di questi anni, anche durante i miei lunghi e difficili periodi di assenza. E per rendermi ogni giorno una persona migliore. Questa Tesi è dedicata anche a lei.

Contents

List of publications associated to the Thesis xi

Introduction 1

I Cosmology: Theory and Observations 5

1 The concordance model of Cosmology 11
1.1 General Relativity 11
1.2 The ΛCDM model 15
1.2.1 Thermal history of the Universe 24
1.3 The inhomogeneous Universe 29
1.3.1 Scalar-vector-tensor decomposition 29

2 Cosmic inflation 37
2.1 Inflation as a scalar field 39
2.2 Inflationary perturbations 43
2.2.1 Scalar perturbations 44
2.2.2 Tensor perturbations 46
2.3 Models of inflation 47

3 The Cosmic Microwave Background 51
3.1 From the primordial plasma to the CMB 52
3.2 Main features of the CMB 53
3.3 CMB anisotropies 55
3.3.1 Temperature anisotropies 55
3.3.2 Polarisation anisotropies 57
3.3.3 CMB secondary anisotropies 64
3.4 The CMB angular power spectrum 65
3.5 Observing the CMB 73
3.5.1 Past experiments 74

3.5.2 Current / active experiments 75
3.5.3 Future experiments 76

4 Foregrounds as a contaminant to CMB observations 77
4.1 Inventory of the contaminants 77
4.2 Polarised foregrounds 79
4.2.1 Synchrotron 80
4.2.2 Thermal dust 82
4.2.3 Point sources 83
4.3 Component separation 84
4.3.1 ILC 86
4.3.2 NILC 87
4.3.3 MC-NILC 89
4.4 Simulating the contaminated sky 90
4.4.1 Dust 91
4.4.2 Synchrotron 91

II Looking at the sky with the *Simons Observatory* 93

5 The *Observatory* 97

6 External data for the *SO* 103
6.1 External data preprocessing 105
6.1.1 Downgrading 108
6.1.2 Coordinates transformation 109
6.1.3 Masking 110
6.1.4 Additional splits building for *WMAP* 111
6.2 External data characterisation 112

7 Simulations 119
7.1 Covariance estimation 119
7.2 Transfer function estimation 123

III The *B*-sure project: robustness tests for primordial *B*-modes 127

8 The need for robustness tests in CMB science 129
8.1 The *BICEP* accident 130

9 How to build a robustness test **133**
9.1 Input sky simulations 135
9.2 Component separation 141
9.3 Likelihood 144
9.4 Reference Gaussian simulations 148
9.4.1 `CMBtens` 149
9.4.2 `FGgauss` 150
9.5 Minkowski functionals 151
9.6 Scattering transform statistics 156
9.7 Compatibility analysis 161
9.7.1 For Minkowski functionals 162
9.7.2 For ST coefficients 164

10 Results of the tests characterisation **167**
10.1 Minkowski functionals test 167
10.2 Scattering transform test 173

IV Conclusions **183**

11 Some final words **185**

List of acronyms **189**

List of Figures **190**

List of Tables **199**

Bibliography **227**

Ai miei Genitori,
e a mia Moglie

List of publications associated to the Thesis

1. **C. Ranucci**, S. Pierre, L. Vacher, A. Carones, N. Krachmalnicoff, E. Allys, J.-M. Delouis, P. Campeti, G. Puglisi, and C. Baccigalupi, *"B-sure. Part II. Scattering transforms as robustness test for tensor-to-scalar ratio detection from CMB observations"*. Submitted to: *Journal of Cosmology and Astroparticle Physics*, September 2026. arXiv: 2609.17531 [astro-ph.CO].

2. The *SO* Collaboration, *"The Simons Observatory: forecasted constraints on primordial gravitational waves with the expanded array of Small Aperture Telescopes"*. Published in: *Journal of Cosmology and Astroparticle Physics*, April 2026. DOI: 10.1088/1475-7516/2026/04/051. arXiv: 2512.15833 [astro-ph.CO].

3. **C. Ranucci**, A. Carones, L. Vacher, N. Krachmalnicoff, and C. Baccigalupi, *"B-sure. Part I. Minkowski functionals as robustness test for tensor-to-scalar ratio detection from CMB observations"*. Published in: *Journal of Cosmology and Astroparticle Physics*, January 2026. DOI: 10.1088/1475-7516/2026/01/045. arXiv: 2510.18868 [astro-ph.CO].

4. The *SO* Collaboration, *"The Simons Observatory: science goals and forecasts for the enhanced Large Aperture Telescope"*. Published in: *Journal of Cosmology and Astroparticle Physics*, August 2025. DOI: 10.1088/1475-7516/2025/08/034. arXiv: 2503.00636 [astro-ph.IM].

5. L. Vacher, A. Carones, J. Aumont, J. Chluba, N. Krachmalnicoff, **C. Ranucci**, M. Remazeilles, A. Rizzieri, *"How bad could it be? Modelling the 3D complexity of the polarised dust signal using moment expansion"*. Published in: *Astronomy and Astrophysics*, May 2025. DOI: 10.1051/0004-6361/202453066. arXiv: 2411.11649 [astro-ph.CO].

Introduction

The field of observational Cosmology focus on the understanding of the structure, content, and history of our Universe, by harvesting this information from observational data. During the last century, humanity has been able to accumulate an enormous amount of data, which ultimately led to the shape of the concordance model of Cosmology, dubbed the Λ Cold Dark Matter (ΛCDM). This model is a beautiful and minimal phenomenological description of the ~ 14 billion of years of evolution of the Universe, from a few seconds after its beginning until our days, and even a little into the future. The quantity of data at our disposal allowed us to enter the "precision Cosmology" era, enabling the measurement of the ΛCDM parameters to percent accuracy.

Nonetheless, human beings never stop to ask themselves questions, and many secrets still awaits to be uncovered. Perhaps one of the most fascinating mystery is the origin of our Universe: *how did everything start?* The search for this answer is one of the frontiers tackled by observational Cosmology today, and it also is the main topic that will be faced in this Thesis. In particular, I will focus on the observable known as the Cosmic Microwave Background (CMB), one of the richest trove of information the sky can offer. It is the oldest light that we can observe from Earth, being the detectable remnant of the Universe when it was in its primeval form, a very hot and dense plasma made of photon and baryons. Many physical theories suggest the presence of a very faint signal hidden in the polarised signal of the CMB, called the primordial B-mode signal. It is thought to carry crucial clues on the unknown inflationary phase (or cosmic inflation), a theorised brief epoch during which the Universe underwent through an exponential expansion, stretching quantum fluctuations to macroscopic scale.

The measurement of the faint primordial B-modes is one of the major targets of current and future microwave instrumentation. The state-of-the-art is represented by the *Simons Observatory* (*SO*), a newly built experiment situated in the Atacama desert of Chile. The *SO* is currently observing the sky, gathering data with the most refined telescopes ever built for CMB science. I have been a member of the *SO* Collaboration since the start of my PhD, where I mostly contributed to the

characterisation of Galactic and extra-Galactic foregrounds, which represent some of the main contaminations of CMB observations. Our own Galaxy prevents us from having a clear look at the microwave sky, with dust grains and relativistic electrons emitting in the same range of frequencies as the CMB photons. The combination of *SO* observations with external data from other microwave experiment is very useful for characterising Galactic foregrounds and for their subtraction or mitigation, necessary in order to have a "cleaned" view of the CMB.

The physics of Galactic foregrounds is so complex that only a certain amount of mitigation can be achieved. It is not possible to completely remove them from microwave sky maps, and residuals contamination will always be present. Even if very small, the foregrounds leftover has a level comparable to the primordial B-mode signal, and their presence is able to hinder the extraction of the cosmological signal and, even worse, can lead to a false detection, which has already happened in recent years. To tackle this problem, in the last two years of my PhD I worked on the development of robustness tests apt at the rejection of an eventual detection of primordial B-modes, when the measurement is contaminated by the presence of Galactic foreground residuals left by the cleaning procedure. To this purpose I explored mathematical tools sensitive to the specific features characterising the foregrounds emission, which are different from the pure CMB. I employed summary statistics like Minkowski functionals (MFs) and scattering transforms (ST) to build robustness tests able to warn us about possible bias in the estimation of cosmological parameters.

The Thesis is organised in three major Parts. In Part I I provide a general introduction to the modern picture of Cosmology, with a particular focus on my research interests. The starting point is the theory of General Relativity, which is combined with decades of celestial observations to build the concordance model of Cosmology that we have today. I present the major milestones accomplished by human beings in the quest to understand the history and content of our Universe, together with the remaining mysteries and the challenges we have to overcome to solve them. Chapter 3 is dedicated to the CMB radiation, where I describe its properties and the amount of crucial information we can extract from it. Chapter 4 is dedicated to Galactic foregrounds and their role in CMB observations.

Part II focuses on the *Simons Observatory*, spanning from the experimental capability of the project, to my role in the Collaboration. In particular, I report my contributions in chapters 6 and 7: the former presents the characterisation of datasets coming from other CMB experiments, in order for them to be combined with *SO* observations; the latter describes the several types of sky simulations that I have produced for the instrument characterisation, and for the validation and testing

of the B-modes data analysis pipeline.

Part III presents the other main project I carried out during my PhD, focused on the development of robustness tests for the detection of primordial B-modes. At first I describe why such an effort is needed in this historical moment, and then I present the (simulated) experimental setup I adopted for the analysis, and the mathematical tools explored to build such robustness tests.

Finally, I give a summary of the Thesis and draw some conclusions in Part IV, also providing insights on the road ahead and the challenges that awaits to be faced.

Throughout the manuscript I will adopt natural units $c = k_{\rm B} = \hbar = 1$ unless otherwise stated, where c is the speed of light in vacuum, $k_{\rm B}$ is the Boltzmann constant, and $\hbar$ is the reduced Planck constant.

Part I

Cosmology: Theory and Observations

Cosmology is one of the most ancient branches of science, naturally arising from some of the primitive questions that human beings started to pose themselves: *Where do we come from? Where are we going? Why are we here?* Ancient civilisation would look at the sky, trying to provide answers to existential doubts through the fabrication of myths and tales. But they would also interrogate themselves on the mechanics of celestial objects, with the example of the Greco-Roman philosopher Claudius Ptolemy proposing the earliest form of a cosmological model, the geocentric system, in the 2nd century AD. This was the very primitive form of Cosmology, intended as the study of the "cosmos" and Nature as a whole.

Modern Cosmology is the scientific evolution of the ancient Cosmology, and its goal is to answer the same doubts we had thousands of years ago, but with the power of mathematics and physical laws. It concerns the study of the history of the Universe, when and how it began, its content and structure, and its ultimate fate. The birth of modern Cosmology is intrinsically linked to the development of General Relativity (GR), the theoretical framework describing the gravitational field as the interconnection between the matter content and the geometry of the Universe. The theory sets the foundation to interpret the experimental discoveries of the last century, leading to the consolidation of the nowadays concordance model of Cosmology, called the Λ Cold Dark Matter (ΛCDM).

The ΛCDM framework is just an evolved form of the earlier hot Big Bang model, built upon a series of critical experimental evidences known as "observational pillars". The story starts in 1917, when Albert Einstein used the field equations of GR to derive a cosmological model [68] describing a static (not contracting and not expanding) Universe, containing uniformly distributed matter and a balance term Λ, named the cosmological constant. This extra quantity was manually added by Einstein in order to preserve the idea of a static Universe, which was the generally accepted idea at the time. A Universe filled only with matter would have rapidly collapsed under its own gravitational attraction, unless counteracted by a repulsive force provide by the Λ term. At the time, Einstein did not recognise the underlying instability of this model if small perturbations are introduced.

The first groundbreaking cosmological discovery happened in 1929, when Edwin Hubble and Milton Humason measured the distance and radial velocity of nearby galaxies (at the time, just "nebulae") through spectroscopic observations with the telescope of Mount Palomar. They found a positive linear correlation between distances and velocities of these galaxies [110]: larger the distance, larger the velocity. Moreover, the positive sign of the measured velocities means that nearly all the observed objects are receding with respect to the Earth. This discovery was totally clashing with the at the time predominant idea of a static Universe,

and it ignited years of discussions aimed at its explanation. The key of the puzzle was to consider the expansion of the Universe, already proposed by Alexander Friedman [79] and Georges Lemaître [119] years before Hubble and Humason's measurement[1]: each object is receding from the others because of the continuos expansion of spacetime.

This prompted decades of debates about the origin and dynamics of the Universe, mainly between two contrasting pictures: the steady state model and the Hot Big Bang model. In the former, continuos matter creation was invoked to maintain a constant density of the Universe, exactly balancing the expansion rate and subject to instability in a similar way to the original Einstein's proposal. In the latter, the Universe has been expanding since its beginning, starting from an initial high-density state [120].

The accidental discovery of the Cosmic Microwave Background (CMB) radiation in 1964 by Arno Penzias and Robert Wilson [155] (and physically interpreted by Robert Dicke's group [60]) provided the experimental evidence favouring the Hot Big Bang model[2], with these detected photons being the remnants of a primordial Universe in thermal equilibrium (also called "primordial fireball").

The CMB measurement was also helpful to corroborate the Big Bang nucleosynthesis (BBN) theory, outlined some years before by Ralph Alpher, Hans Bethe, and George Gamow in the famous $\alpha\beta\gamma$ paper of 1948 [13]. During the last decades, the theory has been able to successfully explain the primordial abundances of light elements (hydrogen, helium, deuterium, with some caveats on lithium) and how they could have been generated during the first $\sim$ 20 minutes of the Universe, based only on the Big Bang hypothesis and on the knowledge of the reaction rates of these nuclear processes.

The most recent observational breakthrough is represented by the discovery of the accelerated expansion of the Universe reported in 1998 and 1999 by two independent research teams led by Adam Riess [183] and Saul Perlmutter [156]. In experiments similar to the one of Hubble and Humason, they built a cosmological distance ladder by measuring the redshift and luminosity of different so-called "standard candles", as Cepheids and type Ia supernovae, in order derive information on cosmological parameters from measurements of the local Universe. They found direct evidences for our Universe to be in a phase of accelerated expansion, with supernovae being fainter than expected for a standard flat Universe. This contrasted with the idea of gravitational interaction slowing down the expansion of the Universe, and led to the acceptance of the mysterious dark energy (in the form of

[1] Actually, Lemaître experimentally discovered the expansion of the Universe years before Hubble.

[2] Even if Penzias and Wilson measured the radiation at just one wavelength. The full confirmation arrived years later with the mapping of the blackbody spectrum.

a cosmological constant Λ) as the dominant component of spacetime, contributing to $\sim 68\%$ of the total energy density budget.

These observational pillars now constitute the foundation of modern Cosmology and of the concordance cosmological model, where different physical components drive the dynamics of the Universe, interacting between each other in the theoretical framework of General Relativity. Nonetheless, many mysteries still awaits to be uncovered, and for this, many challenges need to be faced. In the following of this Part, I will describe some of these frontiers, with a focus on the ones I worked on during my PhD.

Chapter 1

The concordance model of Cosmology

In this chapter, I will describe the ΛCDM model starting from the mathematical framework behind it, represented by General Relativity. I will then move to the applications of this formalism to our Universe and to the different observed phenomena driving the shape of the concordance model. The inflationary paradigm will also be described as an extension of the ΛCDM scenario. Hereafter, I will employ the Einstein notation, meaning that repeated indices are summed, and the $(-,+,+,+)$ metric signature is adopted.

1.1 General Relativity

General Relativity (GR) is the geometric theory describing gravitational interactions in our Universe. It thus constitutes the first building block of the theoretical framework upon which the majority of the cosmological models are constructed, by explaining the dynamics of the Universe depending on its content, at given times. All the other physical interactions are usually neglected when we talk about the Universe as a whole. Strong and weak interactions are only important on very small scales ($\lesssim 10^{-15}\,\mathrm{m}$). Moreover, the Universe appears to be neutral on average, meaning that electromagnetic forces can not play a role in the large-scale dynamics of the Universe. Thus, GR can be considered as the only relevant long-range forces, neglecting individual particles and focusing only on the collective average behaviour of matter, which can be described by statistical mechanics. In other words, Cosmology can be thought as the application and combination of two equations, the Einstein equations for gravity and the Boltzmann equations for matter and radiation.

In the following I will discuss the most important GR notions needed for the

development of the Thesis, omitting many interesting mathematical and technical points. I refer the reader to other sources [145, 214, 62, 75] for complete and pedagogical descriptions of GR.

The starting point is the definition of the metric. The metric takes two infinitesimally close points in space (or spacetime) defined in some arbitrary coordinate system, and returns the actual physical distance between them. This physical distance is an invariant, in the sense that two observers in different coordinate systems will measure the same distance. Thus, the metric turns observer-dependent coordinates into invariants. For the usual Cartesian coordinate system in a 2D plane, the distance between two points (x^1, x^2) separated by $\mathrm{d}x^1$ and $\mathrm{d}x^2$ is given by the familiar Pythagorean theorem:

$$\mathrm{d}s^2 = \sum_{i,j=1,2} g_{ij}\mathrm{d}x^i\mathrm{d}x^j = \left(\mathrm{d}x^1\right)^2 + \left(\mathrm{d}x^2\right)^2 , \tag{1.1}$$

where the metric g_{ij} in this case is simply the identity matrix $\mathbb{1}$:

$$g_{ij} \overset{\text{Cartesian}}{=} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1} . \tag{1.2}$$

If we instead consider polar coordinates (r, θ) in the same 2D plane, the distance between two points separated by $\mathrm{d}r$ and $\mathrm{d}\theta$ will be

$$\mathrm{d}s^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j = (\mathrm{d}r)^2 + r^2\,(\mathrm{d}\theta)^2 , \tag{1.3}$$

where the metric g_{ij} is now

$$g_{ij} \overset{\text{polar}}{=} \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} . \tag{1.4}$$

Both forms of the metric describe the same space, a 2D plane, and the distance they provide is the same.

If we now move to a four-dimensional spacetime, comprising three spatial coordinates and one time coordinate, the invariant includes time intervals:

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu , \tag{1.5}$$

where the first coordinate is time-like ($x^0 = t$) and the other three are spatial coordinates. $g_{\mu\nu}$ is symmetric, so it has four diagonal and six independent off-diagonal components. The interval $\mathrm{d}s^2$ is also called proper time, and again, different observers will measure this very same quantity, no matter the coordinate system. One of the good reasons to use the metric is that it naturally describes

curved spacetimes. In this framework, instead of considering gravity as an external force acting on particles, it can be included in the metric with particles that now move freely in curved spacetimes. This is just a simple way to think of the geometrisation of gravity. The underlying principle is the general covariance realised by Einstein more than a century ago: an observer in a uniform gravitational field makes exactly the same measurements as one in an accelerated reference frame. Special relativity is described by the $\eta_{\mu\nu}$ metric, a simpler form of the general metric:

$$g_{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.6}$$

This is the metric implicitly used to derive the equations of electromagnetism. It describes a spacetime that is not curved. In Minkowski space, particles travel in straight lines unless they are acted on by a force. In a curved space, particles travel along a geodesic, the shortest path (or extremal path) between two points. In GR, this is the path followed by a particle in the absence of any forces apart from gravity. The geodesic equation can be derived as a generalisation of Newton's law with no forces, accounting for general coordinate systems and spacetimes, obtaining

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = 0 \,, \tag{1.7}$$

where λ is an affine parameter which monotonically increases along the particle's path, since time is now one of the coordinates and can not be used as the evolution parameter. In the above equation I also introduced the Christoffel symbols $\Gamma^\mu_{\alpha\beta}$ (or Levi-Civita connection coefficients, or affine connection coefficients), which play here the same role as the Newtonian gravitational field. The Christoffel symbols are directly linked to the geometry of spacetime, as they can be derived using only first derivatives of the metric:

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left(g_{\alpha\nu,\beta} + g_{\beta\nu,\alpha} - g_{\alpha\beta,\nu} \right) , \tag{1.8}$$

where the comma indicates the derivative with respect to the coordinate $g_{\alpha\nu,\beta} \equiv \partial g_{\alpha\nu} / \partial x^\beta$. They are symmetric in the lower indices α and β. Thus, the geodesic equation is the generalised version of Newton's law of motion, where the left-hand side represent the acceleration of a particle. The motion of a particle is then completely specified by the metric describing the curvature of the spacetime, which in turn depends on the spacetime content. This relation is condensed into the

Einstein equations:

$$\mathcal{G}_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \,, \tag{1.9}$$

where $\mathcal{G}_{\mu\nu}$ is the Einstein tensor, defined as

$$\mathcal{G}_{\mu\nu} \equiv \mathcal{R}_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \,, \tag{1.10}$$

$\mathcal{R}_{\mu\nu}$ is the Ricci tensor

$$\mathcal{R}_{\mu\nu} = \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} - \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\alpha\nu} \,, \tag{1.11}$$

and its contraction $R \equiv g^{\mu\nu}\mathcal{R}_{\mu\nu}$ is the curvature scalar (or Ricci scalar). On the right-hand side of eq. (1.9), $T_{\mu\nu}$ is the stress-energy (or energy-momentum) tensor, describing the energy contributions from the constituents of the spacetime. The remaining Λ term is the cosmological constant, an additive term allowed by Einstein equations. It can also be moved to the other side of the field equations by incorporating it into the stress-energy tensor as $T^{\Lambda}_{\mu\nu} = -\left(\Lambda/8\pi G\right) g_{\mu\nu}$. The $8\pi G$ factor allows for the recovery of the usual Newtonian theory of gravity in the weak field or low-velocity limits. The Einstein equations are tensor equations relating a set of symmetric 4 $\times$ 4 tensors, where each tensor has 10 independent components. The four Bianchi identities (not reported here) can be used to reduce the number of independent equations from 10 to 6, leaving the metric with four gauge-fixing degrees of freedom corresponding to the freedom to choose a coordinate system. One of this constraints is the local conservation of energy and momentum (or continuity equation), which reads

$$\nabla_{\nu} T^{\mu\nu} = \frac{\partial T^{\mu\nu}}{\partial x^{\nu}} + \Gamma^{\mu}_{\nu\beta} T^{\beta\nu} + \Gamma^{\nu}_{\nu\beta} T^{\mu\beta} = 0 \,. \tag{1.12}$$

The Einstein tensor $\mathcal{G}_{\mu\nu}$ is the simplest symmetric tensor able to relate the geometry of a spacetime ($g_{\mu\nu}$) with its energetic content ($T_{\mu\nu}$), while also satisfying the energy-momentum conservation law. Thus, the left-hand side of the Einstein equations is a function of the metric, while the right-hand side depends on the content of the spacetime; the Einstein equations are the connection between the two. The solution of the Einstein equations is a metric, which then is plugged into the geodesic equations in order to know the evolution of a particle in a spacetime curved by its content through the stress-energy tensor.

This very brief introduction to the GR formalism I laid down in this section is enough to explain, model, and predict a huge amount of physical phenomena, with the majority of them being observed after the birth of GR and thus categorised as "confirmations" of the theory. Some of them include the passage of time, the geom-

etry of space, the motion of bodies in free fall, the propagation of light, gravitational time dilation, gravitational lensing, gravitational redshift of light, the Shapiro time delay, singularities and black holes, gravitational waves. So far, remarkably, all tests of GR have been in agreement with the theory. Last but not least, GR provides the foundation of the concordance model of Cosmology, which we discuss in the following section.

1.2 The ΛCDM model

As already introduced before, the flat Λ Cold Dark Matter (ΛCDM) is nowadays considered as the concordance model of Cosmology, being the simplest, most minimal, and easily interpretable description of our Universe. It is a phenomenological model, meaning that the majority of its structure has been invoked "a posteriori" in order to explain puzzling observed phenomena, rather than predicting them with an "a priori" approach. Examples are the dark matter and dark energy components, which now play a major role in the cosmological landscape but have been included only to mitigate the lack of explanation for several observations. Nonetheless, the ΛCDM model possesses solid theoretical foundations, provided by physical theories, which have been able to explain and also predict some of the most crucial phenomena observed in the last century. It relies on GR as the underlying theory of gravity explaining the dynamics of the Universe.

The basic assumption for the construction of the flat ΛCDM model is the cosmological principle, which states that *at sufficiently large scales, the properties of the Universe are the same for all observers, no matter where we are and where we look* [123]. More technically, it implies that the Universe is homogeneous and isotropic on very large scales. Homogeneity means that the density of the Universe is approximately constant at all positions, while isotropy means that there is no preferred direction. In other words, a homogeneous Universe is invariant under translations, and an isotropic Universe is invariant under rotations. Isotropy implies homogeneity, but a homogeneous Universe could be anisotropic. Philosophically, the cosmological principle is a formal request by cosmologists for the observable Universe to be a fair representation of the unseen, unknown Universe, and that the same physical laws apply throughout it[1]. The cosmological principle seems to be verified by observations of the large-scale structure of the Universe, even if it still questioned and open for debate.

The other building block of the ΛCDM model is the consideration of an expanding

[1] Karl Popper is one of the philosophers that have criticised the cosmological principle, stating "it makes our lack of knowledge a principle of knowing something" [116].

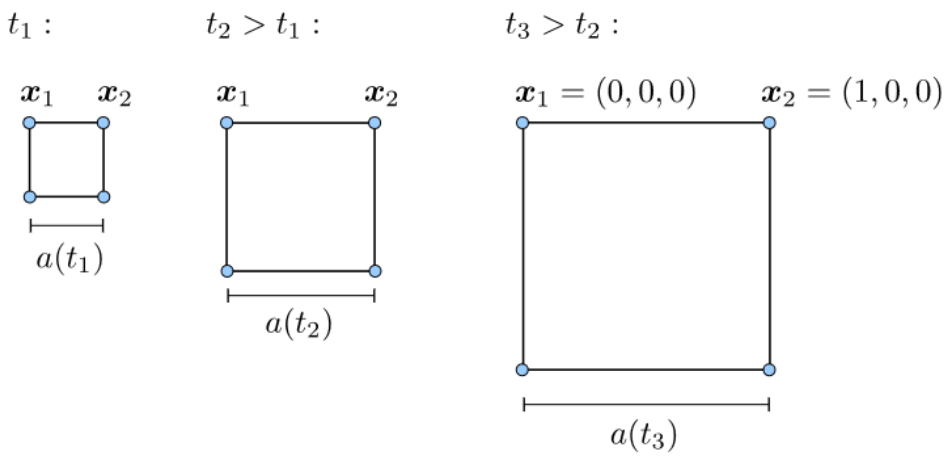


Figure 1.1. Expansion of the Universe. The comoving distance between points x_1, x_2 on a grid remains constant as the Universe expands. The physical distance is proportional to the comoving distance times the scale factor $a(t)$. Figure taken from [62].

Universe: if the Universe is homogeneous and isotropic as per the cosmological principle, then distances between points are not fixed, but they become larger with time, as the spacetime stretches. As I briefly sketched at the beginning of this Part, the idea of a static Universe was already criticised by theoretical cosmologists for instability reasons, favouring the idea of a spacetime in continuos expansion [79]. The measurements of receding galaxies made by Lemaître [119], Hubble, and Humason [110] was the experimental evidence of the expansion of the Universe. They also registered the positive correlation between distances d and velocities v of the galaxies, known as the Hubble-Lemaître law:

$$v = H_0 d\,, \tag{1.13}$$

where H_0 is the proportionality factor of the relation, dubbed the Hubble constant. The effect of the expansion can be encapsulated in a dimensionless, time-dependent scale factor a, which is conventionally set to 1 at present time, and it was smaller in the past. If we consider two points on a grid in space that evolves with time as in figure 1.1, we see that their comoving distance (measuring differences between coordinates) remains the same, while the physical distance changes with time, through the scale factor. A comoving frame is a coordinate system that follows the expansion of the Universe, with the time t called the comoving or cosmic time. The age of the Universe at present time is indicated by t_0. Another useful time variable usually adopted in Cosmology is the conformal time (or conformal horizon), related

to the cosmic time t through the scale factor a:

$$\mathrm{d}\tau = \frac{\mathrm{d}t}{a(t)} \quad \rightarrow \quad \tau = \int_0^{t_*} \frac{\mathrm{d}t}{a(t)} \,. \tag{1.14}$$

This is the maximum distance a photon can travel from the beginning of time until t_*. We will indicate derivative with respect to the cosmic time with overdots ($\partial x/\partial t = \dot{x}$), while derivatives with respect to conformal time are expressed with primes ($\partial x/\partial \tau = x'$).

A direct consequence of the expansion is that the physical wavelength of light emitted from a distant object is stretched out proportionally to the scale factor; the observed wavelength is larger than the one at which the light was emitted. This stretching factor is usually defined as the redshift z:

$$\begin{aligned} z \equiv \frac{\lambda_{\mathsf{obs}} - \lambda_{\mathsf{emit}}}{\lambda_{\mathsf{emit}}} = \frac{\lambda_{\mathsf{obs}}}{\lambda_{\mathsf{emit}}} - 1 \Rightarrow \\ \Rightarrow 1 + z = \frac{\lambda_{\mathsf{obs}}}{\lambda_{\mathsf{emit}}} = \frac{a_{\mathsf{obs}}}{a_{\mathsf{emit}}} = \frac{1}{a_{\mathsf{emit}}} \,. \end{aligned} \tag{1.15}$$

Following the convention on the scale factor, redshift is $z = 0$ at present time and it is larger in the past.

In a flat (Euclidean) space, the comoving distance between two points is $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$, where the metric is the same we discussed in eq. (1.1) but with a third spatial dimension $g_{ij} = \mathbb{1} = \mathsf{diag}(1, 1, 1)$. If each coordinate is expanding proportionally to the scale factor, the physical distance becomes:

$$\begin{gathered} d(t) = \sqrt{[a(t)x_1]^2 + [a(t)x_2]^2 + [a(t)x_3]^2} = a(t)\sqrt{x_1^2 + x_2^2 + x_3^2} = a(t)r \,, \\ \text{with} \quad a(t = t_0) = a_0 \equiv 1 \,, \end{gathered} \tag{1.16}$$

where the metric now includes the background expansion through the scale factor

$$g_{ij} = \begin{pmatrix} a^2(t) & 0 & 0 \\ 0 & a^2(t) & 0 \\ 0 & 0 & a^2(t) \end{pmatrix} = a^2(t) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = a^2(t)\,\mathbb{1} \,. \tag{1.17}$$

The velocity is just the derivative of the distance, being

$$v(t) = \dot{d}(t) = \dot{a}(t)r + a(t)\dot{r} = \dot{a}(t)r \,, \tag{1.18}$$

where $\dot{r} = 0$ as the comoving distance does not depend on time. If we introduce the

Hubble parameter (or Hubble rate)

$$H(t) \equiv \frac{1}{a(t)} \frac{\mathrm{d}a(t)}{\mathrm{d}t} = \frac{\dot{a}(t)}{a(t)}\,, \tag{1.19}$$

we can rewrite the velocity as

$$v(t) = \dot{a}(t) r = H(t) a(t) r = H(t) d(t)\,. \tag{1.20}$$

Looking at nearby galaxies ($t \sim t_0$), we are able to retrieve the empirical Hubble-Lemaître law:

$$\begin{aligned} H(t \simeq t_0) &\simeq H_0 \\ v(t \simeq t_0) &= H_0 d\,. \end{aligned} \tag{1.21}$$

The Hubble parameter quantifies how rapidly the scale factor changes, and its value at present time is, again, the Hubble constant. Thus, the combination of cosmological principle and expansion of the Universe alone already gives a fundamental prediction which can be experimentally tested.

Moving to a four-dimensional flat spacetime, the metric need to also consider the time coordinate:

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & a^2(t) & 0 & 0 \\ 0 & 0 & a^2(t) & 0 \\ 0 & 0 & 0 & a^2(t) \end{pmatrix}, \tag{1.22}$$

which leads to the following invariant interval

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t) \left(\mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2 \right)\,. \tag{1.23}$$

If then we allow for the spacetime to be curved and not simply flat any more, and we consider polar coordinates (i.e., a spherical comoving reference frame in spacetime), the metric assumes the shape

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t) \left(\frac{1}{1 - \kappa r^2}\, \mathrm{d}r^2 + r^2\, \mathrm{d}\theta^2 + r^2 \sin^2\theta\, \mathrm{d}\phi^2 \right) \tag{1.24}$$

where (t, r, θ, ϕ) is the set of coordinates adopted. This is known as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, which is the most general metric describing a homogeneous and isotropic expanding spacetime, thus satisfying the cosmological principle. The $\kappa = 1/R_0^2$ parameter specifies the curvature of the spacetime:

- $\kappa = 0$ is the flat (or Euclidean) Universe we already examinated. Two free

particles travelling parallel to each other will continue to travel on parallel trajectories. In the ΛCDM model, spacetime is flat;

- $\kappa > 0$ describes a closed (or spherical) Universe, where the initially parallel trajectories will gradually converge;

- $\kappa < 0$ describes an open (or hyperbolic) Universe, where the parallel paths will diverge.

R_0 is the curvature radius of the Universe at t_0. The value of κ, and thus the geometry, is determined by the total energy density in the Universe. If the density is higher than a critical value (approximately $10^{-29}\,\mathrm{g\,cm^{-3}}$), the universe is closed; if the density is lower, it is open. For the Universe to be flat, the energy density has to be precisely equal to the critical one, which is absolutely unlikely to happen. Yet, this is actually what current observations are suggesting, leading to the adoption of a flat spacetime for the ΛCDM scenario. The inflationary mechanism (discussed in chapter 2) provides a possible explanation of this fact.

We now move to discussing the implications of the cosmological principle on the content of an expanding Universe. The cosmological principle actually places strong constraints on the behaviour of matter, radiation, and other components. These fluids filling the Universe must be perfect fluids, meaning they can not have any flux on large scale, as it would break isotropy by selecting a preferred direction. They have isotropic pressure, no shear, no viscosity, and no heat conduction. Considering a perfect fluid at rest in the comoving frame, with four-velocity $u^\mu = (1, 0, 0, 0)$, its stress-energy tensor is restricted to take the form:

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix} , \tag{1.25}$$

where ρ and P are the energy density and pressure of the fluid, respectively. The relation between these two quantities is called equation of state, expressed as

$$P = w(t)\rho\,, \tag{1.26}$$

where $w(t)$ (or $w(a)$) is the equation of state parameter, different for each component contained in the Universe, and which in general can be time-dependent. The time evolution of the components of the stress-energy tensor can be derived by working out the conservation of energy and momentum of eq. (1.12), which is a set of four equations. Applying the first equation $\nabla_\mu T^{\mu 0}$ to an expanding Universe described

by the FLRW metric, we find

$$\begin{aligned} \frac{\partial \rho}{\partial t} + 3\frac{\dot{a}}{a}\left(\rho + P\right) = 0\,, \\ \frac{\partial \rho}{\partial t} + 3H\left[1 + w(t)\right]\rho = 0\,. \end{aligned} \tag{1.27}$$

Integrating this equation, we get the evolution of any constituent, which reads:

$$\rho(a) = \rho_0 \, \exp\left\{-3\int^a \left[1 + w\left(a'\right)\right]\frac{\mathrm{d}a'}{a'}\right\}. \tag{1.28}$$

If the equation of state parameter does not depend on time $w(a) = w$, the above equation simplifies to

$$\rho(a) = \rho_0 a^{-3(1+w)}\,. \tag{1.29}$$

The equation of state parameter then completely specifies the time evolution of a fluid present in the Universe for the specific cosmological model, and for each of the several components it can be inferred through thermodynamics and statistical physics arguments.

We are now able to work out the explicit form of the Einstein equations for a perfect fluid in an expanding Universe, assuming the general (curved) FLRW metric as an ansatz. If we insert the stress-energy tensor of eq. (1.25) into eq. (1.9), we obtain what are known as the first and second Friedmann equations (or Friedmann-Lemaître equations):

$$\begin{aligned} H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} \\ \frac{\ddot{a}}{a} = H^2 + \dot{H} = -\frac{4\pi G}{3}\left(\rho + 3P\right) \end{aligned} \tag{1.30}$$

where H is the Hubble parameter already defined in eq. (1.19), $\rho = \sum_s \rho_s$ and $P = \sum_s P_s$ are the total energy density and pressure from all the components s, and κ is the curvature parameter introduced in eq. (1.24). These equations can be solved to obtain the evolution of the scale factor as a function of time $a(t)$, and thus the evolution of the Universe. Incidentally, the combination of the first and second Friedmann equations gives back the energy and momentum conservation of eq. (1.27).

If we introduce the critical density ρ_{cr} and the density parameter Ω_s for each component s as

$$\rho_{\mathrm{cr}} \equiv \frac{3H_0^2}{8\pi G} \quad \Rightarrow \quad \Omega_s = \frac{\rho_{0_s}}{\rho_{\mathrm{cr}}} \tag{1.31}$$

the first Friedmann equation can be recast in a more compact and intuitive form:

$$\begin{aligned}\left(\frac{H}{H_0}\right)^2 &= \frac{8\pi G}{3H_0^2}\sum_s \rho_{0_s} a^{-3(1+w_s)} - \frac{\kappa}{H_0^2}a^{-2} \\ &= \sum_s \Omega_s a^{-3(1+w_s)}\,,\end{aligned} \tag{1.32}$$

where we explicitly wrote the densities time evolution, and we defined a density parameter also for the curvature as $\Omega_\kappa = -\kappa/H_0^2$, which can be seen as an additional fluid with an equation of state parameter $w_\kappa = -1/3$. Evaluating the above equation at present time ($a = a_0 = 1$ and $H = H_0$), we get a constraint for the different parameters

$$\Omega = \sum_s \Omega_s = 1\,, \tag{1.33}$$

which provides an interpretation of the density parameter as the contribution percentage of each component to the cosmic energy budget.

Friedmann equations describe the dynamics of the homogeneous Universe. If a single component is present (or dominant over the other components), the equations can be integrated to obtain the evolution of the scale factor as a function of time and equation of state parameter $w \neq -1$:

$$a(t) \propto t^{\frac{2}{3(1+w)}}\,. \tag{1.34}$$

In the flat ΛCDM scenario, the Universe is populated with the following landscape of constituents:

- **Baryons and dark matter:** "baryons" is the very general word used by cosmologists to address massive particles present in the Standard Model of particle physics. It usually refers to all the ordinary, visible matter in the Universe which emits measurable electromagnetic signals. This includes everything made of protons, neutrons, and electrons, such as stars, planets, interstellar gas, etc. The baryons abundance is constrained by two independent probes: the Big Bang Nucleosynthesis (BBN) [53] and the CMB [166]. These two very precise measurements agree on the baryons density being just $\Omega_{\rm b} \sim 0.05$, 5% of the total density budget of the Universe. With this low value, ordinary matter is not able to account for a large number of observations gathered in the last century. Already in 1933 the pioneer Fritz Zwicky [225] proposed, among other things[2], the existence of a dark matter (DM) component, non-electromagnetically interacting with other particles. The reason

[2]Zwicky was also the first to propose supernovae as the final, exploding stage of a star's life, as well as the origin of cosmic rays. He then proceeded to find supernovae by himself comparing photographic plates by eye (120 detections over 50 years, a list is on this website). He also proposed

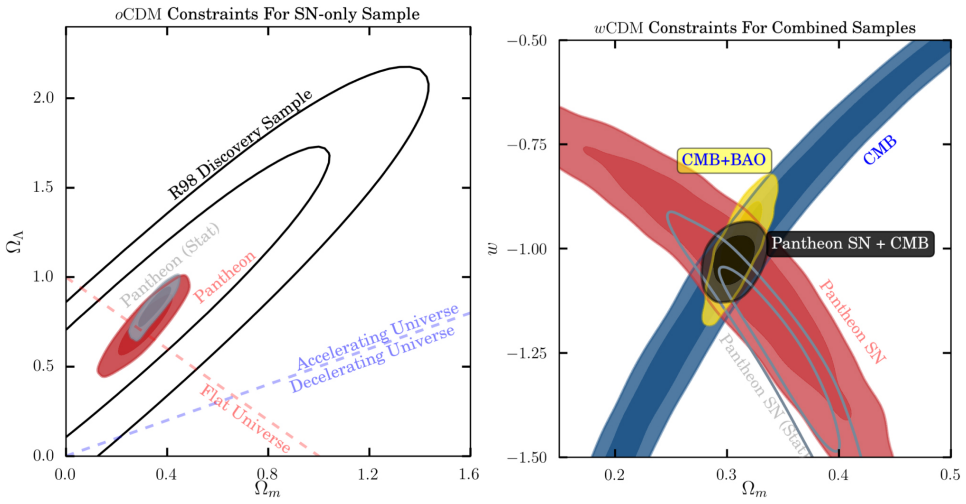


Figure 1.2. Constraints on the energy densities of matter (Ω_m) and dark energy (Ω_Λ), and the dark energy equation of state parameter (w). *Left:* constraint contours for Ω_m and Ω_Λ, from a recent analysis of type Ia supernovae distances, allowing for an open Universe. Filled contours are the more recent results, while the black lines show constraints from the original 1998 sample of [183]. Dashed lines indicate the parameters combination to have a flat Universe and the separation between an accelerating or decelerating one. Data clearly prefer a flat, accelerating scenario. *Right:* contours for w and Ω_m, using type Ia supernovae, CMB, and BAO measurements, and assuming a flat Universe. The combination of different probes points to the cosmological constant ($w = -1$). Both figures are from [189].

was to explain the velocity dispersion measured for the Coma galaxy cluster, way too high for visible matter to account for it. Since then, the existence of dark matter has been largely confirmed by data covering gravitational lensing, galaxy velocities, and CMB, accounting for 30% ($\Omega_\mathrm{DM} \sim 0.3$) [166] of the total energy density budget of the Universe. In ΛCDM, non-baryonic matter is modelled as cold dark matter (CDM), that is a massive, cold, and weakly interacting fluid. It needs to be "cold" (low velocities) in order to efficiently clump and form structures in the early Universe; "hot" (large velocities) dark matter instead would not be able to cluster and explain the large-scale structures we observe in the Universe. The energy density of matter (baryonic and non-baryonic) is dominated by its mass, as[3] $\rho_\mathrm{m} c^2 \gg P_\mathrm{m} \propto v_\mathrm{m}^2$. Thus, matter in the Universe is usually taken to be pressure-less, with an equation of state parameter $w_\mathrm{m} = P_\mathrm{m}/\rho_\mathrm{m} \simeq 0$. Plugging this into eq. (1.28), we get the density evolution for the matter component $\rho_\mathrm{m} \propto a^{-3}$, which is sometimes split into baryonic and CDM contributions $\rho_\mathrm{m} = \rho_\mathrm{b} + \rho_\mathrm{DM}$. According to eq. (1.34), a matter-dominated Universe evolves as $a \propto t^{2/3}$.

neutron stars as possible supernovae remnants, and suggested the idea that galaxies could act as gravitational lenses as a consequence of GR.

[3] I reintroduced here the speed of light c for the sake of clarity.

- **Radiation:** we dub as "radiation" all the components in the Universe which are relativistic, in the sense that their mass can be considered as negligible when compared with their velocity. Photons and neutrinos can be incorporated in the same radiation component. The majority of the radiation contribution is from photons of the Cosmic Microwave Background (CMB), with their numerical density today being $n_\gamma \sim 400\,\mathrm{cm}^{-3}$ [77]. Their energy density has been measured to be $\Omega_\mathrm{r} \sim 10^{-5}$ [166]; for this reason, in many calculations the radiation contribution to the budget is safely ignored, being negligible with respect to other components. Regarding cosmic neutrinos, they have not been measured directly, but there are solid theoretical arguments about their contribution to the energy density. The CMB is able to constrain the total energy density of relativistic species (photons and neutrinos), and the measurements seems to agree with theoretical predictions. For relativistic particles, the radiation pressure is dominant with respect to the negligible mass of photons and neutrinos. Using statistical arguments, the radiation pressure assumes the form $P_\mathrm{r} = \rho_\mathrm{r}/3$, leading to $w_\mathrm{r} = 1/3$. Plugging this into eq. (1.28), we get the density evolution for the radiation component $\rho_\mathrm{r} \propto a^{-4}$. We have an extra $1/a$ term compared with the matter case, coming from the stretching of the photon wavelength described in eq. (1.15). A radiation-dominated Universe evolves as $a \propto t^{1/2}$ following eq. (1.34).

- **Dark energy:** this is the predominant component of the ΛCDM model. It is thought to be responsible for the accelerated expansion of the Universe, observed for the first time through measurements of the distances of type Ia supernovae [183, 156], and later also confirmed by CMB, Baryonic Acoustic Oscillations (BAO), and large-scale structure growth measurements. Moreover, we have strong evidences for the Universe to be flat, meaning that the total cosmic energy density Ω is very close to 1. Since the matter energy density is roughly 30%, it is necessary for dark energy to contribute for the missing $\sim 70\%$. It has to be a non-clustering fluid with a negative equation of state parameter, and thus with negative pressure (or positive tension). Observations (see figure 1.2) have constrained it to be close to $w_\mathrm{DE} = w_\Lambda = -1$, when dark energy is modelled as a cosmological constant Λ, which is the same extra term appearing in the Einstein equations (eq. (1.9)). In general, dark energy can also be modelled in other ways, allowing for a time-dependant equation of state, or considering additional parameters. The cosmological constant is the simplest and most natural scenario in accordance with observations. As the name suggests, if we plug $w_\mathrm{DE} = w_\Lambda = -1$ into eq. (1.28), we get a constant evolution of the energy density $\rho_\mathrm{DE} = \rho_\Lambda$. In the case of dark

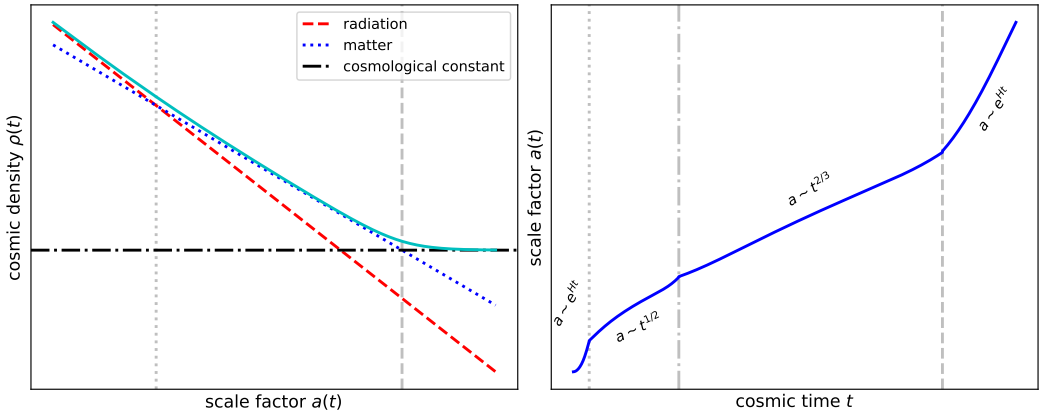


Figure 1.3. Dynamics of the Universe provided by the Friedmann equations. *Left:* energy density ρ as a function of the scale factor a. Different constituents of the universe dominates at different time of the history: radiation (red dashed), non-relativistic matter (blue dotted) and a cosmological constant (black dash-dotted). *Right:* behaviour of the scale factor a as a function of the cosmic time t, in the frame of the ΛCDM model.

energy being the dominant component with $w_\Lambda = -1$, the integration of the Friedmann equations provides a Universe exponentially expanding as $a \propto e^{Ht}$. The cosmological constant scenario has been frequently associated with the notion of vacuum energy, based on the fact the it is something not diluted by the continuous expansion of the spacetime. Sadly (or interestingly), the quantum-mechanical theoretical expectation value for the cosmological constant turns out to be hundreds of order of magnitude (10^{120}) larger than the value required by cosmological observations [213], leaving the story still open as one of the most puzzling problems of modern physics.

Thus, in the ΛCDM model there are different fluids evolving in different ways, and the evolution of the Universe is driven by the dominant component at a specific epoch. Figure 1.3 reports the behaviours of the energy density and the scale factor as a function of time, solutions of the Friedmann equations for different dominating components. The Universe dynamics is first driven by radiation ($\rho \sim a^{-4}$, $a \sim t^{1/2}$); after the matter-radiation equality at $a_{\rm eq}$, non-relativistic matter become the dominant contribution ($\rho \sim a^{-3}$, $a \sim t^{2/3}$); at present-time, dark energy (as a cosmological constant) is dominating the expansion ($\rho \sim$ const, $a \sim e^{Ht}$).

1.2.1 Thermal history of the Universe

We now have all the ingredients to draw an approximate chronology of the Universe according to the ΛCDM model, starting from a few moments after the Big Bang ($t \sim 0$) until the present day ($t \sim 14\,\mathrm{Gyr}$). The observed expansion of the Universe

implies that at earlier times the Universe was smaller, denser, and necessarily hotter. Consequently, the temperature and energy of the Universe scales as the inverse of the scale factor $T \propto 1/a(t) = 1 + z$. Depending on the considered time (or scale factor), we can define different phases which have occurred in the Universe:

- **Planck epoch ($t \lesssim 10^{-34}\,\mathrm{s}$):** at this energies, all the fundamental forces are merged into quantum gravity, and the physics of the Standard Model is not valid. We do not have access to this era, just speculations.

- **Inflationary epoch ($10^{-33}\,\mathrm{s} \lesssim t \lesssim 10^{-14}\,\mathrm{s}$):** the earliest time we can access through observational data. Immediately after its birth, the Universe undergoes through a hypothetical phase of exponential expansion ($a \sim e^{Ht}$) that stretches the infant spacetime from a radius of $10^{-29}\,\mathrm{m}$ to 0.9 m, on a timescale of $10^{-36}\,\mathrm{s}$. At this energies, electroweak and strong forces are unified together (Grand Unified Theory, GUT), and GR is still considered to be valid. At the end of inflation, a reheating mechanism converts the energy of the field into a very hot and dense mixture of particles present in the Standard Model. Formally, the inflationary theory is not part of the ΛCDM, but rather a non-standard extension invoked to solve some of its problems. It will be discussed in more details in chapter 2.

- **Electroweak phase transition ($10^{-14}\,\mathrm{s} < t < 10^{-10}\,\mathrm{s}$):** this is the earliest stage explainable within the ΛCDM scenario. At energies $> 160\,\mathrm{GeV}$, all elementary particles are massless. At energies $\sim 160\,\mathrm{GeV}$ ($t \sim 10^{-12}\,\mathrm{s}$) the electroweak symmetry breaking happens, and particles start to acquire mass through the Higgs mechanism [87, 212, 185]. The W and Z bosons (mediators of the weak force) become abruptly massive, while the photon (mediator of the electromagnetic force) remains massless. The electroweak interaction is disentangled into electromagnetic and weak interactions. All the fundamental forces have now taken their present forms. Temperature is still too high to form neutron and protons, and the content of the Universe is a hot quark-gluon plasma.

- **Quantum chromodynamics phase transition ($10^{-12}\,\mathrm{s} < t < 10^{-5}\,\mathrm{s}$):** this range of energies can be accessed by particle physics experiment like particle accelerator. As the Universe expand, the quark-gluon plasma cools down and quarks are able to form hadrons (baryons and mesons).

- **Neutrino decoupling ($t \sim 1\,\mathrm{s}$):** as the Universe becomes less dense and colder, neutrinos decouple from other particles and begin travelling freely through space. The Cosmic Neutrino Background (CNB or CνB) is released.

As neutrinos rarely interact with matter, this primordial background still exists today. Due to cosmological redshift, the neutrinos from this event have now a very low energy, around 10^{-10} times the amount of those observable with present-day direct detection. Since even high-energy neutrinos are notoriously difficult to detect, it is very likely that the CNB may not be possible to observe for many years, if at all.

- **Big Bang Nucleosynthesis (BBN) ($3\,\mathrm{min} \lesssim t \lesssim 20\,\mathrm{min}$):** the temperature of the plasma is now low enough to allow nuclear reactions. Neutrons and protons are fused into deuterium (^{2}H), helium-3 (^{3}He), helium-4 (^{4}He), and traces of lithium (^{7}Li) and beryllium (^{7}Be). The reaction rates can be measured in nuclear physics laboratories. The temperature drops too quickly for heavier elements to form, which will be then forged by nuclear reactions in the interiors of stars. The abundance of primordial light elements can be predicted and then measured through observations of old galaxies, less polluted by stellar activity. A very good agreement is registered between predictions and measurements, making the BBN one of the most important confirmations of the Big Bang model. Observations of lithium amount show some discrepancy with the theoretical forecast, even if it is a notoriously difficult measurement due to the strong pollution by stellar-generated lithium.

- **Recombination ($50,000\,\mathrm{yr} \lesssim t \lesssim 380,000\,\mathrm{yr}$):** as we saw in the cosmic inventory, the radiation component evolves with an extra $1/a$ factor compared to matter. This means that going back in time, radiation is the first component to guide the expansion of the Universe, after the hypothetical inflationary period. Thus, until now the Universe evolved as a radiation-dominated spacetime, with $\rho \sim a^{-4}$ and $a \sim t^{1/2}$. As the Universe cools, matter-radiation equality occurs at around 50,000 years after the Big Bang, and from this point onwards the dynamics of the Universe is driven by non-relativistic matter, with $\rho \sim a^{-3}$ and $a \sim t^{2/3}$. Structures can start to form and grow in amplitude, as they are not wiped out by radiation pressure any more. The Universe cools down enough for the ionised particles to start forming the first atoms; photons gradually stop interacting with matter (decoupling) and can stream freely in the "transparent" Universe. Similarly to what happened with neutrinos, the CMB radiation is emitted: a thermal bath of primordial photons which can be observed today as the oldest light in the Universe accessible to us. It is a blackbody thermal radiation with an average temperature of $T_{\mathrm{CMB}} = 2.725\,\mathrm{K}$, with temperature variations (anisotropies) at the 10^{-5} level. Being the main focus of this Thesis, the CMB will be described in more details in the coming section 3.

Table 1.1. Best-fit values of the concordance ΛCDM model parameters combining CMB data from the *Planck* experiment with BAO measurements. We report the flat ($\Omega_\kappa = 0$) `base_plikHM_TTTEEE_lowl_lowE_lensing_post_BAO` set from [166]. The six parameters in the upper part are the primary parameters of the concordance cosmology, while those below are derived parameters. The last column lists the 95% confidence-level parameter limits. More details are on the *Planck* wiki.

Parameter	Symbol	Best fit	95% C.L. limits
Baryon density	$\Omega_\mathrm{b}h^2$	0.02245	$\pm$ 0.00027
Dark matter density	$\Omega_\mathrm{c}h^2$	0.11928	$\pm$ 0.0018
Reionisation optical depth	τ_rei	0.057	$\pm$ 0.014
Hubble constant	h	0.6770	$\pm$ 0.0081
Scalar spectral index	n_s	0.9682	+ 0.0076 / $-$ 0.0073
Scalar spectrum amplitude	$\ln\left(10^{10}A_\mathrm{s}\right)$	3.048	$\pm$ 0.028
Dark energy density	Ω_Λ	0.689	$\pm$ 0.011
Matter density	Ω_m	0.311	$\pm$ 0.011
Matter spectrum normalisation	σ_8	0.811	$\pm$ 0.012
Age of the Universe [Gyr]	t_0	13.784	+ 0.040 / $-$ 0.037

- **Dark ages and reionisation ($380,000\,\mathrm{kyr} \lesssim t \lesssim 1\,\mathrm{Gyr}$):** after recombination, all the matter in the Universe is in the form of a neutral gas of atoms, interacting through gravity only. Apart from 21 cm spin-down photons emitted by neutral hydrogen, there are no sources of light during this epoch. CMB photons shift from visible to infrared light after 3 million years. The Universe is hence devoid of visible light. As the neutral gas gradually clumps, the first stars (Population III) begin to form some million years after the Big Bang. These stars mainly emit UV photons, with energies able to ionise the neutral gas present in the interstellar medium. The hydrogen atoms are split back into a plasma of free electrons and protons for the first time since recombination and decoupling, which is the current status of interstellar and intergalactic gas. This process is known as reionisation.

- **Large-scale structures formation ($1\,\mathrm{Gyr} \lesssim t \lesssim 14\,\mathrm{Gyr}$):** matter continues to clump and collapse following the gravitational influence of dark matter halos. Galaxies, clusters of galaxies, and super-clusters of galaxies slowly acquire the shape and distribution we observe today, forming the filamentary structure known as the Cosmic Web. After dark energy-matter equality ($t \sim 10\,\mathrm{Gyr}$), the cosmological constant starts dominating the dynamics of the Universe, leading to the accelerated expansion.

Hence, the ΛCDM model is an extremely powerful phenomenological description of our Universe. It can provide a consistent explanation of the thermal history of the Universe, in agreement with most observations and relying on just 6 minimal

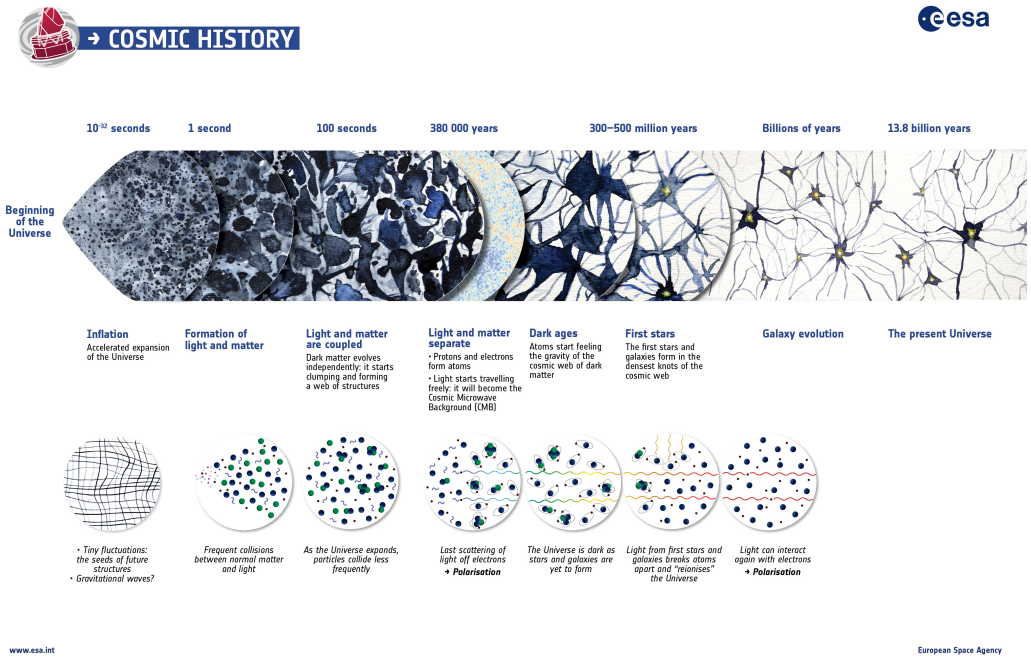


Figure 1.4. Simplified chronology of the Universe, as per the ΛCDM concordance model. The timeline in the upper part shows an artistic view of the evolution of the cosmos on large scales, passing through inflation, CMB emission, "Dark Ages", birth of the first stars and galaxies, and until present time. The inserts below show some of the microscopic processes taking place during cosmic history. Credit: ESA.

parameters. Their values are constrained by the wealth of cosmological data we have at our disposal today, combining CMB, type Ia supernovae, and large-scale structures measurements. The best-fit values for the concordance model are reported in table 1.1, where we introduced a few parameters: the reionisation optical depth τ, describing some features of the reionisation process; the reduced Hubble constant $h = H_0/100$; the scalar spectra index n_s, describing how the density fluctuations of the early Universe vary by size or scale; the scalar power spectrum amplitude A_s, which quantifies the magnitude of the primordial quantum fluctuations; the matter power spectrum normalisation σ_8, which parametrise the amplitude of the density fluctuations, normalised to the present Universe. Moreover, observational data suggest a flat Universe, with the measured curvature being compatible with zero ($\Omega_\kappa = 0.0007 \pm 0.0037$). The radiation density parameter is constrained to be very small, $\Omega_\mathrm{r} \simeq 1.67 \times 10^{-5}$, and thus negligible with respect to the other components (even if dominant at very high redshift).

Even if amazingly successful, the ΛCDM is not the definitive answer to all of the features of the Universe. There are still a lot of open questions for which we do not have an explanation; some of them will be explored in chapter 2.

1.3 The inhomogeneous Universe

Until now, I have dealt with the evolution of the Universe under the assumption of homogeneity, with the density of the fluid filling the spacetime being the same everywhere. But as we can clearly see by looking at the sky, the Universe is not homogeneous on the small scale, with structures, galaxies, filaments, and voids being present. Thus, spatial variations must occur in the early Universe in order for structures to form. I will now discuss the most important concepts about the evolution of these inhomogeneities, omitting many technical details and almost all the mathematical derivations, which can easily be found in many standard textbooks (e.g, [214, 62]).

In order for us to observe the present-day structures, the inhomogeneities must have grown with time through gravitational interaction; going back in the past, they should have been much smaller. This implies that at some point in the early Universe, these spatial variations are small enough to be considered as simple, linear perturbations around a mean value provided by the approximately homogeneous background. Thus, perturbed quantities like the metric or the stress-energy tensor can be separated as

$$\begin{aligned} g_{\mu\nu}(t,x) &= \bar{g}_{\mu\nu}(t) + \delta g_{\mu\nu}(t,x) = \bar{g}_{\mu\nu}(t) + h_{\mu\nu}(t,x)\,, \\ T_{\mu\nu}(t,x) &= \bar{T}_{\mu\nu}(t) + \delta T_{\mu\nu}(t,x)\,, \end{aligned} \tag{1.35}$$

where the mean backgrounds $\bar{g}_{\mu\nu}$ and $\bar{T}_{\mu\nu}$ depend only on time, being approximated as homogeneous at the zeroth order, and the perturbations are small in magnitude ($\delta g_{\mu\nu}\,,\ \delta T_{\mu\nu} \ll 1$). We also define an overdensity (or underdensity) in a given point of the spacetime as

$$\delta(t,x) = \frac{\rho(t,x) - \bar{\rho}}{\bar{\rho}} \tag{1.36}$$

with respect to the average background density $\bar{\rho}$.

1.3.1 Scalar-vector-tensor decomposition

The evolution of the above perturbations can be calculated using the Einstein equations, similarly to what is done for the homogenous Universe, but now expanding the perturbations at linear order around a mean value. In general, Einstein equations are tensor equalities comprising of 16 coupled equations, which are reduced to 10 thanks to the symmetry of the Einstein tensor. These equations can be decoupled by means of the scalar-vector-tensor (SVT) decomposition[4]. It allows to decom-

[4]which follows from the Helmholtz's theorem, or Helmholtz decomposition.

pose the evolution equations for the general linearised perturbations of the FLRW metric into four scalars, two divergence-free spatial vector fields, and a traceless, symmetric spatial tensor field with vanishing longitudinal components. The vector and tensor fields each have two independent components, so this decomposition encodes all 10 degrees of freedom of the metric perturbation. Moreover, the principle of general covariance at the basis of GR (and in general, the gauge invariance of gauge theories) requires that the dynamics of the field does not change under local transformations and choice of coordinates (gauge). This means that one can change the coordinate frame and switch between different gauges in order to simplify calculations and find gauge invariant quantities which can than be translated between different gauges. Choice of convenient gauges allow to reduce the degrees of freedom of the metric perturbation from 10 to 6, where two scalar and one vector fields of the decomposition may be set to zero.

The components of the perturbed metric can be written as

$$\begin{aligned} g_{00}(t,x) &= -1 + h_{00}(t,x)\,, \\ g_{0i}(t,x) &= a(t)h_{0i}(t,x) = a(t)h_{i0}(t,x)\,, \\ g_{ij}(t,x) &= a^2(t)\left[\delta_{ij} + h_{ij}(t,x)\right]\,, \end{aligned} \tag{1.37}$$

where h_{00}, h_{0i}, h_{ij} are metric perturbations functions of space and time, all of whose components are assumed to be small in magnitude. Adopting the so-called Newtonian gauge, the components of $h_{\mu\nu}$ assume the form

$$\begin{aligned} h_{00} &= -2\Psi\,, \\ h_{0i} &= 0\,, \\ h_{ij} &= 2\left(\Phi\delta_{ij} + h_{ij}^{\mathrm{TT}}\right)\,, \end{aligned} \tag{1.38}$$

where Ψ corresponds to the Newtonian potential and governs the motion of slow-moving (non-relativistic) bodies, and Φ is the perturbation to the spatial curvature (it can also be interpreted as a local perturbation to the scale factor). Vector fields are usually ignored in cosmological calculations as they are associated to vorticity, and the relative perturbations are damped by the expansion of the Universe. They do not contribute to structure formation and have no important effect compared to the other fields, in the majority of cosmological scenarios. Tensor perturbations are described by h_{ij}^{TT}, which is a traceless ($g^{ij}h_{ij}^{\mathrm{TT}} = 0$), transverse tensor field.

An important consequence of the decomposition framework is the so-called decomposition theorem, which states that scalar, vector, and tensor perturbations evolve independently at linear order. That is, if some physical process in the early Universe generates tensor perturbations, these do not induce scalar perturbations

as they evolve. Conversely, the evolution of scalar perturbations is not touched by possible vector or tensor perturbations. The decomposition theorem holds because the background FLRW metric is spatially isotropic. Thanks to this theorem, scalar and tensor perturbations can be studied independently.

In general, a cosmological perturbation theory does not attempt at predicting the presence or amount of overdensities in any given spot on the sky, which would be nearly impossible. Rather, theories focus on achieving a statistical description of the perturbed field, using quantities as correlation functions and the power spectrum, which we can hopefully measure through observations. To do so, it is very convenient to work in Fourier space: we define the spatial Fourier transform and inverse transform of the real-valued field $f(t, \mathbf{x}),\ f \in \mathbb{R}^3$ as

$$\begin{aligned} f(\mathbf{x}) &= \int \frac{\mathrm{d}^3\mathrm{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{f}(\mathbf{k})\,, \\ \tilde{f}(\mathbf{k}) &= \int \mathrm{d}^3\mathbf{x}\, e^{-i\mathbf{k}\cdot\mathbf{x}} f(\mathbf{x})\,, \end{aligned} \tag{1.39}$$

where $\mathbf{k}$ is the associated wavevector, or mode, usually characterised by its magnitude $k = |\mathbf{k}|$. Each mode identifies a particular wavelength, or spatial scale; being the inverse of x, small k correspond to large scales, while large k correspond to small scales. In Fourier space, derivatives of the field $f(t, \mathrm{x})$ with respect to $\mathbf{x}$ become simpler algebraic relations:

$$\frac{\partial f(t, \mathrm{x})}{\partial x^i} = i k_i \tilde{f}(t, \mathrm{k}) \tag{1.40}$$

The power of the Fourier transform is to decouple the mode of the field, if isotropy is assumed: each Fourier mode k evolves independently from the other modes k', and their evolution can be calculated without knowing the solution for other values.

The two-point correlation function describes the excess probability of having two overdensities at positions $\mathbf{x}_1$ and $\mathbf{x}_2$ separated by a certain distance $\mathbf{d} = |\mathbf{x}_1 - \mathbf{x}_2|$, with respect to the uniform probability given by a random distribution. It is defined as

$$\xi(\mathbf{d}) = \langle \delta(\mathbf{x}_1)\, \delta(\mathbf{x}_2) \rangle = \frac{1}{V} \int \mathrm{d}^3\mathbf{x}\, \delta(\mathbf{x}_1)\, \delta(\mathbf{x}_2) \tag{1.41}$$

with V being the normalising cosmic volume. The power spectrum $P(k)$ is the Fourier transform of the correlation function:

$$\begin{aligned} \xi(\mathbf{d}) &= \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\, P(k) e^{i\mathbf{k}\cdot\mathbf{x}}\,, \\ P(k) &= \int \mathrm{d}^3\mathbf{x}\, \xi(\mathbf{d}) e^{-i\mathbf{k}\cdot\mathbf{x}}\,. \end{aligned} \tag{1.42}$$

Equivalently, if $\tilde{\delta}(\mathbf{k})$ is the Fourier transform of the overdensity, the power spectrum can also be expressed as an average in Fourier space, as

$$\langle \tilde{\delta}(\mathbf{k}_1)\, \tilde{\delta}^*(\mathbf{k}_2) \rangle = (2\pi)^3 \delta_{\rm D}^3(\mathbf{k}_1 - \mathbf{k}_2) P(k)\,, \tag{1.43}$$

where δ^* indicates the complex conjugate of δ, and $\delta_{\rm D}^3$ is not an overdensity but the Dirac delta function which is zero when $\mathbf{k}_1 \neq \mathbf{k}_2$, meaning that different modes are independent. If the considered field is approximately Gaussian-distributed, all its statistical information is encoded in the 2-point correlation function, and thus in the power spectrum, which then represents a sufficient statistics. Any higher-order correlation functions, for example the 3-point correlation function involving distances between three different positions, are highly suppressed. The power spectrum $P(k)$ has dimensions of a volume (length^3); a widely used dimensionless expressions is

$$\Delta^2(k) = \frac{k^3}{2\pi^2} P(k)\,, \tag{1.44}$$

which gives the variance of the modes in a logarithmic wavenumber interval. It is also useful to quantify deviations from a scale invariant (or scale-free) spectra which goes like $P(k) \propto k^{-3}$. Throughout this Thesis, we follow the convention to drop the tilde on Fourier-transformed quantities:

$$\tilde{f}(\mathbf{k}) \to f(\mathbf{k})\,. \tag{1.45}$$

Scalar perturbations

The perturbed metric in Newtonian gauge and without tensor perturbations is

$$\begin{aligned} g_{00} &= -1 - 2\Psi\,, \\ g_{0i} &= 0\,, \\ g_{ij} &= a^2 \delta_{ij}\,(1 + 2\Phi)\ . \end{aligned} \tag{1.46}$$

Inserting this metric into the Einstein equations, together with a stress-energy tensor containing perturbations of density and pressure, we can obtain two relationships between the metric perturbations Ψ and Φ and the density perturbations for the different components populating the early Universe:

$$\begin{aligned} k^2\Phi + 3\frac{a'}{a}\left(\Phi' - \Psi\frac{a'}{a}\right) &= 4\pi G a^2 \left[\rho_{\rm m}\delta_{\rm m} + 4\rho_{\rm r}\Theta_{\rm r_0}\right]\,, \\ \Phi + \Psi &= 0 \quad \text{(no anisotropic stress)}\,, \end{aligned} \tag{1.47}$$

where the subscript m stands for all matter such as baryons and dark matter, and r includes photons and relativistic (massless) neutrinos. Θ_{r_0} are the monopole terms of the radiation (photons and neutrinos) fractional temperature changes $\Theta = \delta T/T$, linked to the density fluctuations through

$$\rho_{\mathrm{r}} \propto T^4 \quad \rightarrow \quad \frac{\delta\rho_{\mathrm{r}}}{\rho_{\mathrm{r}}} = 4\frac{\delta T}{T} = 4\Theta\,. \tag{1.48}$$

In the second equation we have neglected the anisotropic stress induced by photons and neutrinos quadrupole terms Θ_{r_2} in the right-hand side. In this simplified case the two gravitational potentials are equal and opposite. In the limit of no expansion ($a =$ constant), the first relation of eq. (1.47) reduces to the ordinary Poisson equation for gravity (in Fourier space): the left-hand side is the Laplacian $-\nabla^2$ while the right-hand side is $4\pi G a^2 \delta\rho$. We only have two equations because all the other Einstein equations are zero or redundant, as they do not add any information about the evolution of Φ and Ψ, similarly to the Friedmann equations we discussed in section 1.2.

Differently from tensor perturbations, both equations do not contain any second time derivatives acting on Φ or Ψ, being constraint equations. Φ and Ψ do not represent propagating degrees of freedom (neither does the usual Newtonian gravitational potential).

Tensor perturbations

Tensor perturbations can be characterised by a metric perturbation as in eq. (1.37) [62], with $h_{00} = -1$, $h_{0i} = 0$, and

$$\delta g_{ij}(t, \mathbf{x}) = a^2(t) h_{ij}^{\mathrm{TT}}\,, \quad h_{ij}^{\mathrm{TT}} = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.49}$$

where we have chosen the perturbations to be in the x-y plane, with the z-axis being the direction of the wavevector $\mathbf{k}$. The perturbations to the metric are described by two functions, h_+ and $h_\times$, assumed small. They are two components of a divergence-less, traceless, and symmetric tensor. Divergence-less means that $k^i h_{ij}^{\mathrm{TT}} = k^j h_{ij}^{\mathrm{TT}} = 0$.

Inserting this perturbed metric into the Einstein equations, we obtain the expression for the evolution of the tensor modes:

$$h_{\mathrm{t}}'' + 2\frac{a'}{a} h_{\mathrm{t}}' + k^2 h_{\mathrm{t}} = 0\,, \tag{1.50}$$

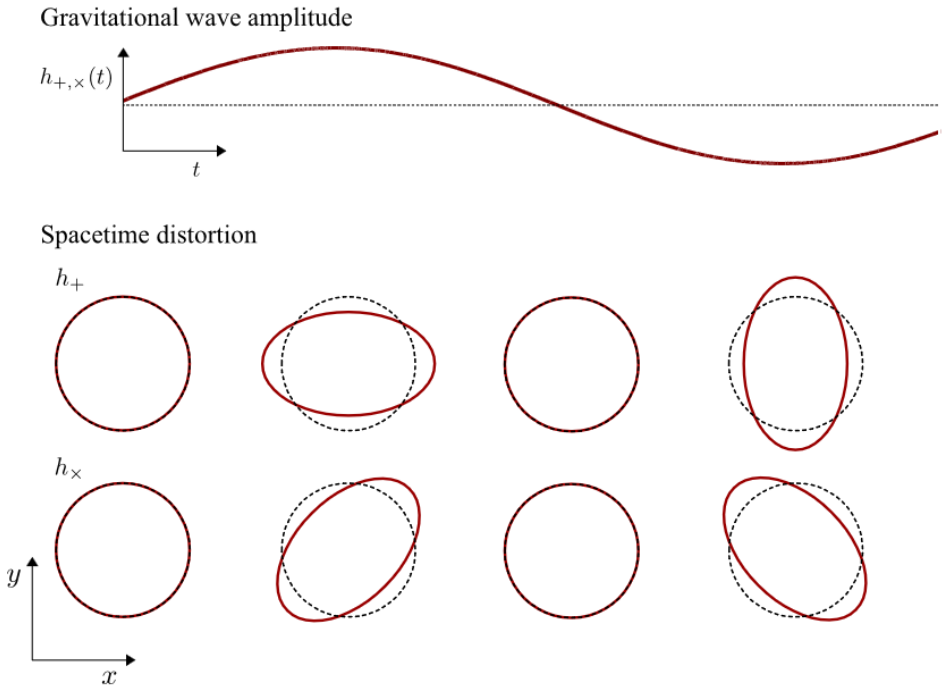


Figure 1.5. Illustration of a propagating gravitational wave (tensor mode), perturbating the spacetime. The k-vector is along the z axis, coming out of the page. The upper panel shows the time evolution of the wave amplitude (neglecting the damping due to the expansion), while the lower panel shows the stretching and compression of spacetime perpendicular to the wave's direction of propagation at the different points in time during a single wave cycle. Figure from [62].

with $\mathrm{t} = +, \times$. The stress-energy tensor contribution in the right-hand side of the Einstein equations is zero, as we assume to not have anisotropic stress, as we did for the scalar perturbations. This means that tensor modes are not produced by matter perturbations[5].

Eq. (1.50) is a wave equation, and the corresponding solutions are called gravitational waves. The two independent components represent the two polarisations of gravitational waves, $+$ and $\times$. Since they are transported by a massless carrier (the graviton), the two polarisations are orthogonal to the direction of propagation, just like the photon. Neglecting the expansion of the Universe, we immediately see that the two solutions are $h_\mathrm{t} \propto e^{\pm ik\tau}$. The two modes correspond to waves travelling in the $\pm z$ direction at the speed of light. The spacetime distortions induced by these modes are illustrated in figure 1.5, where the elliptical pattern in the plane perpendicular to the wavevector is a fundamental distinction between tensor

[5]Anisotropies in the radiation components (photons and neutrinos) do have an anisotropic stress, given by their quadrupole. For photons, the quadrupole is suppressed during the radiation-dominated era, so their source term can be ignored. The neutrino anisotropic stress is more relevant, and indeed does provide a source term for gravitational waves, leading to a damping of tensor modes on small scales. We neglect it here since we will focus on large-scale tensor modes.

modes and scalar perturbations, as the latter induce distortion which are always azimuthally symmetric around the wavevector. From eq. (1.50), we see that when the wavelength of a perturbation is larger than the causal horizon (simplifying, in the limit of $k \to 0$), $h_{\rm t} = \text{constant}$ is a solution; when the wavelength of the mode starts instead to be comparable with the horizon, its amplitude oscillates and begins to decay as $1/a$. This implies that the energy density in gravitational waves redshifts as a^{-4}, exactly as any form of radiation. This means that very large scales are "frozen" and preserve their amplitude during the time they are outside the horizon, while smaller scales enter the horizon way before and are thus subject to more damping. We will see in the next chapter how these tensor perturbation can be produced in the primordial Universe as a consequence of the inflationary period.

Chapter 2

Cosmic inflation

As we have seen in the previous chapter, the ΛCDM scenario is widely considered as the concordance model of Cosmology. It is able to explain a huge amount of observational data, and it provides a phenomenological description of the content of the Universe. Even if extremely successful, there is a certain number of questions the ΛCDM framework is not able to answer. Dark matter and dark energy are evidently the most prominent ones, since they represent $\sim 90\%$ of the constituent of our Universe, and we have no clue about their nature. But they are not the only problems in the Hot Big Bang model. Here we report some of them:

- **Initial conditions:** the Big Bang theory postulates an initial very hot uniform plasma, which expands according to the equations described in section 1.2 and ultimately forms stars and galaxies. For structures to be assembled, density perturbations in the primordial plasma are required, otherwise the expansion of the Universe would pull the mass apart before reasonable clumps are formed. A completely uniform mass density is not able to produce the structure we see today. What is the origin of the perturbations seeds?
- **Magnetic monopoles:** at $t < 10^{-14}$, the electromagnetic, weak, and strong forces are unified in a single interaction, arising due to spontaneous symmetry breaking from a single gauge theory. Grand Unified Theories (GUT) predict a number of heavy, stable particles, with the most notorious being the magnetic monopole, a "charge" of magnetic field. A huge amount of these magnetic monopoles is predicted to be produced in the early Universe, and they should have persisted to the present day as one of the main constituent of the Universe. This is not the case, as they have never been observed in nature, and all searches for them have failed, placing stringent limits on the density of relic magnetic monopoles in the Universe.
- **Horizon problem:** this is one of the indirect consequences of the cosmological

principle. Microwave experiments from space have mapped the CMB with exquisite precision, finding that $t_{\rm CMB} \sim 380,000$ years after the Big Bang the Universe was very smooth and at the same temperature of $T_{\rm CMB} = 2.27255\,{\rm K}$ [77] everywhere we look, with small differences (anisotropies) at the 10^{-5} level. This agrees well with the notion of homogeneous and isotropic Universe at large scales. But while we can observe the full CMB surface since we are looking with a $\sim 13\,{\rm Gyr}$ delay, at the time of its emission a particle could have only travelled a certain distance since the beginning of the Universe, given by the conformal distance we introduced in eq. (1.14):

$$r_H = c\tau = \int_0^{t_{\rm CMB}} \frac{c\,{\rm d}t}{a(t)} \sim 100\,{\rm Mpc}\,, \tag{2.1}$$

where c is the speed of light. This means that today, patches separated by more than ~ 1 degree were outside of their reciprocal causal horizon. At earlier times, two widely apart CMB regions had never been in causal connection, with no possibility of interacting. How comes that they are almost at the exact same temperature? In other words, what mechanism can thermalise the whole Universe even before the reciprocal interaction of its regions?

- **Flatness problem:** CMB observations have constrained our Universe to be very close to a flat geometry within errors, with $\Omega_\kappa = 0.0007 \pm 0.0037$ [166]; in other words, the energy density is very close to the critical density. From the Friedmann equations:

$$\Omega(t) - 1 = \frac{\kappa}{H^2(t)a^2(t)}\,. \tag{2.2}$$

  Given the expansion of the Universe, any past small deviations of the total density from the critical one would rapidly increase over cosmic time. This implies that in the early times, the Universe was much closer to flatness than today, with the total density extremely close to the critical one, with a difference of 10^{60} or less at Planck scales. Why and how has the total density been fine-tuned to this exact value at the beginning of time?

These and other minor problems led to the development of inflation as the widely accepted theoretical framework for the solution and explanation of the Hot Big Bang model issues. The general mechanism is nowadays a consensus in the community, even if the physical theory driving it is still unknown. As such, inflation is not part of the concordance model of Cosmology, but rather a ΛCDM extension indirectly supported by observations, but currently looking for experimental confirmation.

2.1 Inflation as a scalar field

The inflationary paradigm was mainly developed in the late 1970s and 1980s, with contributions by several physicists, including Brout [32], Starobinsky [194, 195], Guth [91, 92], Hawking [99], and Linde [125, 126]. The theory explains why our Universe is so homogeneous, flat, and isotropic, and why its different parts began their expansion simultaneously. It provides a mechanism explaining the origins of structure formation and solves the different problems we gathered above. The foundational idea of all the inflationary theories is the occurrence of an exponentially accelerated expansion of the Universe happening immediately after the Big Bang, at $t < 10^{-15}\,\mathrm{s}$, and lasting at least $10^{-32}\,\mathrm{s}$. Before inflation, the whole Universe was contained in a sphere of radius $\sim 10^{-29}\,\mathrm{m}$; by the end of inflation, it grew up to a radius of $\sim 1\,\mathrm{m}$. This idea automatically solves the horizon problem: at very early times, the entire Universe was in causal contact, and all the different regions were connected and thus able to thermalise; the faster-than-light space expansion brings these regions outside of the causal horizon, they are not connected any more, but they still share the same temperature which we observe in the CMB. The abrupt expansion of the Universe also alleviates the magnetic monopole problem, as inflation would greatly dilute the density of these peculiar particles, making them extremely rare in the current Universe and virtually undetectable. The inflationary mechanism is able to solve the flatness problem because of the violent exponential expansion, where the scale factor grows as $a(t) \sim e^{\lambda t}$; whatever is the original curvature of spacetime, inflation will force it to be extremely close to zero, as it scales like $\sim a^{-2}$. We notice here that the expansion due to inflation has the same shape as the accelerated expansion induced by dark energy at present time, as they describe the same phenomena. The difference is in the magnitude of the acceleration, which is hundred of times larger for the inflationary scenario. Solving the horizon problem require the Universe to grow exponentially for roughly 60 e-folds, where e-folds are defined as the logarithm of the ratio between the scale factors at the end ($a(t_\mathrm{e})$) and beginning ($a(t_\mathrm{i})$) of inflation:

$$N_* = \ln\left[a(t_\mathrm{e})/a(t_\mathrm{i})\right]. \tag{2.3}$$

Finally, inflation also provides an explanation for the origins of density perturbations in the early Universe. Primordial quantum fluctuations (guaranteed by the Heisenberg's uncertainty principle) present around the homogenous field before the beginning of the inflationary phase are stretched to macroscopic scales by the exponential expansion of spacetime, becoming the seeds for the density contrast necessary to ignite the formation of structures in the later Universe.

We now summarise the basic formalism needed for discussing the main topics of the Thesis. We refer the interested reader to aforementioned seminal works in the field for complete descriptions [194, 91, 126]. We start by noticing that in order to have a period of accelerated expansion, the second time derivative of the scale factor must be positive, $\ddot{a} > 0$. From the second Friedmann equation (1.30), we see that the fluid responsible for the acceleration must satisfy

$$\rho + 3P < 0 \quad \Rightarrow \quad P/\rho = w < -1/3\,, \tag{2.4}$$

thus requiring a negative effective pressure $P < 0$. This implies that whatever is driving inflation can not be matter or radiation, as they have a non-negative pressure (we recall that $P_{\rm m} \simeq 0$ and $P_{\rm r} = \rho_{\rm r}/3$ for matter and radiation, respectively). Moreover, it cannot be a cosmological constant either: a cosmological constant would lead to perpetual inflation, while inflation needs to end at some point and transition to the radiation- and then matter-dominated phases which we observe.

The simplest possibility to generate such a transitory epoch of accelerated expansion is via the potential energy of a scalar field. This is also applicable in a similar way to build models of dark energy, which are generally dubbed "quintessence". In Nature, there is no known scalar field able to drive inflation. The Higgs boson is a known example of a scalar field, but it is very complicated to use it to drive inflation, as many of its properties have already been constrained. We note that a lot of proposed inflationary theories contain additional scalar fields, as well as other different frameworks; the interested reader can look at [135] for a review of 74 different models of inflation. Here we focus on the single scalar field as the majority of current work on inflation is based on this idea. It is worth stressing that even if inflation is driven by something other than a scalar field, the general underlying inflation scenario is very difficult to avoid, being also corroborated by observational data.

The Lagrangian for a canonical scalar field $\phi = \phi(x,t)$ self-interacting through a potential $V(\phi)$ is

$$\mathcal{L}_\phi = -\frac{1}{2} g^{\mu\nu} \partial_\mu\phi\, \partial_\nu\phi - V\left(\phi\right)\,, \tag{2.5}$$

from which the stress-energy tensor for the scalar field can be derived following Noether's theorem[1] as

$$\begin{aligned} T^\phi_{\mu\nu} &= \frac{\partial \mathcal{L}_\phi}{\partial\left(\partial_\mu\phi\right)} \partial_\nu\phi - g_{\mu\nu}\mathcal{L}_\phi \\ &= \partial_\mu\phi\, \partial_\nu\phi - \frac{1}{2} g_{\mu\nu}\, \partial_\alpha\phi\, \partial^\alpha\phi - g_{\mu\nu} V(\phi) \end{aligned} \tag{2.6}$$

[1]The theorem states that every differentiable symmetry of the action (in this case $S = \int \sqrt{-g}\, \mathcal{L}\, \mathrm{d}^4\mathrm{r}$) of a physical system has a corresponding conservation law.

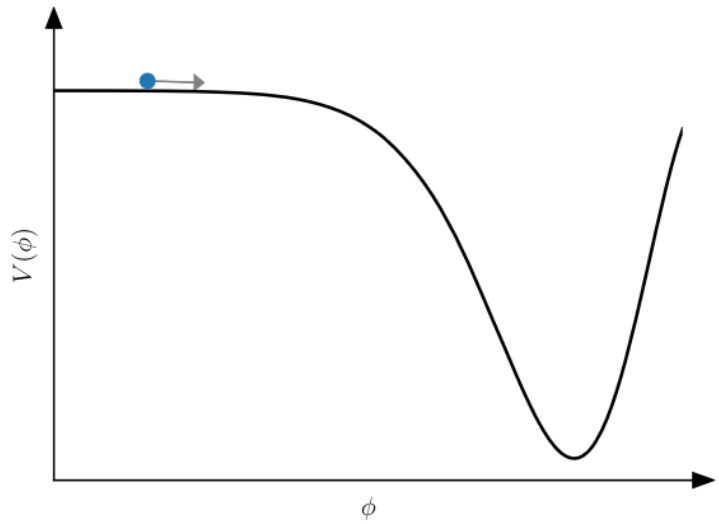


Figure 2.1. A scalar field slowly rolling down a potential $V(\phi)$. Since it rolls slowly, it has little kinetic energy. The potential energy is non-zero, however, so the pressure is negative. The inflationary epoch ends once the field has reached the minimum of the potential. Figure taken from [62].

In general, a field can be approximated as homogeneous at the zeroth order and perturbed at the first order, as $\phi(x,t) = \phi(t) + \delta\phi(x,t)$. If we neglect spatial variations, we get a homogeneous scalar field $\delta\phi = 0$, and ϕ acts as a perfect fluid filling the Universe. In this case the components of the stress-energy tensor reduce to

$$\begin{aligned} \rho_\phi &= \frac{1}{2}\dot{\phi}^2 + V(\phi)\,, \\ P_\phi &= \frac{1}{2}\dot{\phi}^2 - V(\phi)\,, \end{aligned} \tag{2.7}$$

where we inserted the time and spatial components given by eq. (1.25) for the stress-energy tensor of a perfect fluid in a FLRW metric. The equation of state parameter satisfying the condition for having an accelerated expansion is

$$w_\phi = \frac{P_\phi}{\rho_\phi} = \frac{\dot{\phi}^2 - V(\phi)}{\dot{\phi}^2 + V(\phi)} < -\frac{1}{3} \tag{2.8}$$

The first term in the expressions of density and pressure is the kinetic energy density of the field, while the second is its potential energy density. A homogeneous scalar field has then a similar dynamics as a single particle moving in a potential, with $\phi(t)$ being the position of the particle and $\dot{\phi}$ the analogous of the velocity. A field configuration with negative pressure is therefore one with more potential energy than kinetic.

The most popular scenarios of inflation assume a scalar field slowly rolling toward its true ground state [125]. The slow rolling condition $\dot{\phi} \ll V(\phi)$ is required

in order to have little kinetic energy, or in any case much smaller than the potential energy, so to have a negative pressure (which will be nearly -1). The potential energy of such field is very close to constant (if the potential is not too steep) so it dominates over the kinetic energy (and the energy of all other particles). An example of such potential is illustrated in figure 2.1; inflation happens when the field is slowly rolling on the plateau, and it ends once the field has reached the minimum of the potential, where it will oscillate and lose its energy by creating pairs of elementary particles. These particles interact with each other to reach a state of thermal equilibrium with some temperature T. This process is called "reheating". The empty, diluted Universe aftermath of the expansion will be populated with the particles described by the Standard Model.

The evolution of the field ϕ for any potential is derived from the Einstein equations and the conservation of the stress-energy tensor. For the continuity, since ϕ is a homogenous field in an expanding Universe, we get the same result we obtained in the previous section in eq. (1.27), that we report again here:

$$\frac{\partial \rho_\phi}{\partial t} + 3H\left(\rho_\phi + P_\phi\right) = 0\,. \tag{2.9}$$

Evaluating this equation with the density and pressure of the scalar field obtained above yields

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\mathrm{d}V(\phi)}{\mathrm{d}\phi} = 0\,. \tag{2.10}$$

This is the non-linear Klein-Gordon equation of motion for a scalar homogeneous field in an expanding spacetime, self-interacting through the potential $V(\phi)$. It is similar to the equation of motion for a harmonic oscillator, with an additional friction term $3H\dot{\phi}$. The Einstein equations provide

$$H^2 = \frac{8\pi}{3M_\mathrm{P}^2}\left[\frac{1}{2}\dot{\phi}^2 + V(\phi)\right] - \frac{k}{a^2}\,, \tag{2.11}$$

where $M_\mathrm{P}^2 = 1/G$ is the Planck mass in natural units. According to this expression, if the scalar field ϕ is initially large, the Hubble parameter H must be large too. This means that the friction term in the field equation is very large, and thus the scalar field moves very slowly, against the viscosity. Due to the extremely rapid growth of the scale of the Universe and the slow motion of the field ϕ, soon after the beginning of inflation the acceleration and curvature terms become negligible, while the potential energy dominates over the kinetic energy:

$$\ddot{\phi} \ll 3H\dot{\phi}\,, \quad H^2 \gg \frac{k}{a^2}\,, \quad \dot{\phi}^2 \ll V(\phi)\,. \tag{2.12}$$

Thus, the system of equations can be simplified as

$$\begin{aligned} 3\frac{\dot{a}}{a}\dot{\phi} &\simeq -\frac{\mathrm{d}V(\phi)}{\mathrm{d}\phi}\,, \\ \left(\frac{\dot{a}}{a}\right)^2 &\simeq \frac{8\pi}{3M_{\mathrm{P}}^2}V(\phi)\,. \end{aligned} \tag{2.13}$$

The last equation shows that the size of the Universe in this regime grows approximately as $e^{\lambda t}$, with $\lambda = (2/M_{\mathrm{P}})\sqrt{2\pi V(\phi)/3}$.

It is possible to quantify the slow roll assumption by defining two parameters [124] which can be directly linked to observations:

$$\begin{aligned} \epsilon_{\mathrm{sr}}(\phi) &= \frac{M_{\mathrm{P}}^2}{16\pi}\left(\frac{1}{V}\frac{\mathrm{d}V}{\mathrm{d}\phi}\right)^2\,, \\ \delta_{\mathrm{sr}}(\phi) &= \frac{M_{\mathrm{P}}^2}{8\pi}\left(\frac{1}{V}\frac{\mathrm{d}^2V}{\mathrm{d}\phi^2}\right)\,, \end{aligned} \tag{2.14}$$

where the first parameter measures the slope of the potential, and the second its curvature. For the slow roll approximation to hold, the required conditions are

$$\epsilon_{\mathrm{sr}} \ll 1\,, \quad |\delta_{\mathrm{sr}}| \ll 1\,. \tag{2.15}$$

Usually, the end of the inflationary period is identified with the two parameters reaching $\epsilon_{\mathrm{sr}} \sim 1$ and $\delta_{\mathrm{sr}} \sim 1$ (or equivalently, the kinetic energy starts to play a role $\dot{\phi}^2 \sim V$), when the reheating phase happens and the Universe becomes radiation-dominated. From this moment, the history of the Universe can be described through the standard Hot Big Bang theory. Figure 2.2 illustrates the evolution of scales before, during, and after inflation, where they are expanded from sub-horizon to super-horizon, transitioning to the standard Hot Big Bang theory.

2.2 Inflationary perturbations

One of the most important features of the inflationary mechanism is its ability to generate primordial perturbations around the smooth Universe, planting the seeds for the formation of structures. Inflation predicts that quantum-mechanical perturbations in the very early Universe are first produced when the relevant scales are causally connected (within the causal horizon). These scales are successively stretched outside the horizon by inflation, and will then re-enter much later to act as initial conditions for the growth of structures. In Fourier space, the quantum perturbation field will oscillate around zero, with a null average. However, any given mode has non-zero variance, encapsulated in the power spectrum $P(k)$ we

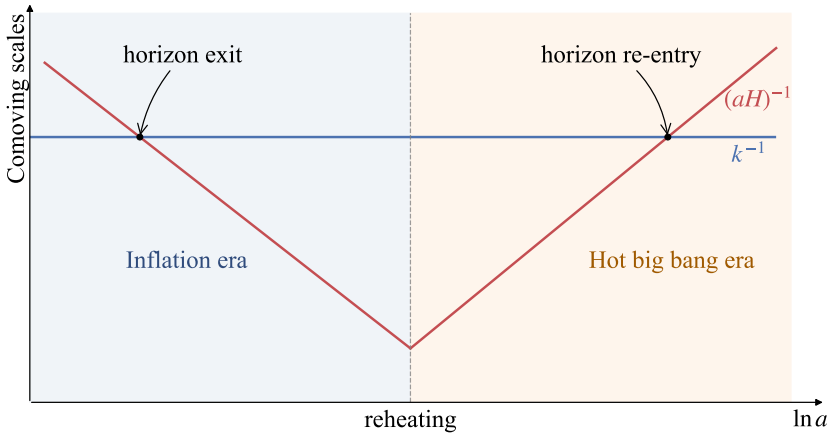


Figure 2.2. Comoving scales evolution during and after the inflationary period. Scales are initially into the horizon and thus in casual connection. Inflation stretches the scales and make them exit the horizon, where they are not casually connected any more, and remain frozen. Finally, scales re-enter the horizon during the hot Big Bang era. Figure from [177].

introduced in eq. (1.43). The perturbations power spectrum is the main prediction of inflationary theories. As we saw in section 1.3, perturbations are decomposed in scalar and tensor modes, which can be studied independently thanks to the decomposition theorem.

2.2.1 Scalar perturbations

One of the primary predictions of single-field inflation is that it generates adiabatic perturbations: different patches of the Universe have different overdensities, but the fractional density perturbations are the same for all species s:

$$\frac{\delta\rho_s}{\rho_s} = \frac{\delta\rho}{\rho}\,, \tag{2.16}$$

and the same for velocities. If they are instead different for different species, we talk about isocurvature perturbations. CMB observations have definitely confirmed the adiabatic nature of primordial perturbations [131, 23, 170], with the isocurvature contribution being negligible.

The perturbed scalar field responsible for inflation can be separated into a homogeneous background plus a linear correction, as

$$\phi(t,x) = \bar{\phi}(t) + \delta\phi(t,x)\,. \tag{2.17}$$

The evolution of the $\delta\phi$ depends on if the perturbation is within or outside the

horizon. For perturbations inside the horizon, first order correction to the metric can be ignored, and the conservation of energy-momentum leads to

$$\delta\phi'' + 2aH\delta\phi' + k^2\delta\phi = 0\,. \tag{2.18}$$

This equation can be rewritten to assume the shape of a harmonic oscillator. To know the evolution of quantum fluctuations, the field needs to be quantised. This can be easily achieved in the case of a quantum harmonic oscillator. The variance of the perturbations of the field are

$$\langle\hat{\delta\phi}(\mathbf{k})\,\hat{\delta\phi}^{\dagger}(\mathbf{k}')\rangle = (2\pi)^3\delta_{\rm D}^3(\mathbf{k}-\mathbf{k}')P_{\rm s}(k)\,, \tag{2.19}$$

with † indicating the Hermitian conjugate. A quantum field $\hat{\delta\phi}$ is defined in all space, so it can be considered as an infinite collection of oscillators, each at a different spatial position, or, in Fourier space, at different values of $\mathbf{k}$. The quantum fluctuations in each of these oscillators are independent (as long as the equations are linear) so $\hat{\delta\phi}(\mathbf{k})$ is completely uncorrelated with $\hat{\delta\phi}(\mathbf{k}')$ if $\mathbf{k}\neq\mathbf{k}'$. This independence is enforced by the Dirac delta function $\delta_{\rm D}^3$. The power spectrum is derived to be (at first order)

$$P_{\rm s}(k) = \left.\frac{2\pi}{k^3}\frac{H^2}{M_P^2\,\epsilon}\right|_{aH=k} \equiv 2\pi^2 A_{\rm s}k^{-3}\left(\frac{k}{k_{\rm p}}\right)^{n_{\rm s}-1}\,, \tag{2.20}$$

where $A_{\rm s}$ is the variance (amplitude) of scalar perturbations in a logarithmic wavenumber interval anchored around the pivot scale $k_{\rm p}$, $n_{\rm s}$ is the scalar spectral index, and quantities in the last equality are evaluated at horizon crossing $aH=k$. The pivot scale is a matter of convention, often decided based on observations. The primordial scalar spectral index $n_{\rm s}$ can also be related to the slow-roll parameters $\epsilon_{\rm sr}$ and $\delta_{\rm sr}$ by taking the derivative of the power spectrum [62], providing

$$n_{\rm s} = 1 - 4\epsilon_{\rm sr} - 2\delta_{\rm sr}\,. \tag{2.21}$$

As we mentioned in section 1.3, a spectrum where $k^3P(k)$ is constant (i.e., does not depend on k) is called a scale-invariant (or scale-free) spectrum. Apart from small deviations encoded in the slow-roll parameters, scalar perturbations are scale-free. Moreover, the inflation field rolls down a potential well during the exponential expansion, meaning that the Hubble rate slowly decreases; a further generic prediction of the inflationary mechanism is that the potential is slightly red-tilted ($n_{\rm s}\lesssim 1$), with larger-scale perturbations having a slightly larger amplitude than smaller-scale perturbations. A spectrum with a small red tilt has indeed been conclusively detected

in the CMB, assuming a ΛCDM model (table 1.1, 95% CL limits),

$$\begin{aligned} A_{\rm s} &= \frac{k_{\rm p}^3}{2\pi^2} P_{\rm s}(k_{\rm p}) \simeq 2.1 \times 10^{-9} \,, \\ n_{\rm s} &= 0.968 \pm 0.007 \end{aligned} \tag{2.22}$$

The slow-roll parameters are therefore very useful to condense the predictions of an inflationary model. They can be measured through observations, but they are also directly connected with inflationary physics, and in particular to the potential V of the scalar field responsible for inflation.

2.2.2 Tensor perturbations

We described tensor perturbations to the metric in section 1.3.1, with the two polarisation states obeying the equation

$$h'' + 2\frac{a'}{a}h' + k^2 h = 0 \qquad (h = h_+, h_\times) \,. \tag{2.23}$$

This equation has exactly the same form as the expression (2.18) for the evolution of the perturbed inflation field reported in the previous section. Therefore, we can restructure it to have the shape of a harmonic oscillator, and quantise it. The variance of the quantised field is again

$$\langle \hat{h}(\mathbf{k})\, \hat{h}^\dagger(\mathbf{k}') \rangle = (2\pi)^3 \delta_{\rm D}^3(\mathbf{k} - \mathbf{k}') P_{\rm t}(k) \,, \tag{2.24}$$

where the same considerations hold. The power spectrum associated to tensor perturbation is

$$P_{\rm t}(k) = \left. \frac{32\pi}{k^3} \frac{H^2}{M_P^2} \right|_{aH=k} \equiv 2\pi^2 A_{\rm t} k^{-3} \left(\frac{k}{k_{\rm p}} \right)^{n_{\rm t}} \,, \tag{2.25}$$

where $A_{\rm t}$ is the variance (amplitude) of tensor perturbation and $n_{\rm t}$ is the tensor spectral index. Note that, compared to the scalar power spectrum, the spectral index does not have a -1 factor; here the spectrum is scale-invariant if $n_{\rm t} = 0$. In practical observations, it is common to replace $A_{\rm t}$ with the tensor-to-scalar ratio r, defined as

$$r(k) \equiv \frac{P_{\rm t}(k)}{P_{\rm s}(k)} \overset{k=k_{\rm p}}{=} \frac{A_{\rm t}}{A_{\rm s}} = 16\, \epsilon_{\rm sr}|_{aH=k} \tag{2.26}$$

Since scalar and tensor modes have different spectral index, in principle r depends on k; this dependence is negligibly small in practical applications. It is also possible

to replace $\epsilon_{\rm sr}$ with the expression for the potential V to get

$$V^{1/4} \simeq 10^{16}\,\mathrm{GeV}\left(\frac{r}{0.01}\right)^{1/4}\,. \tag{2.27}$$

The primordial spectral index $n_{\rm t}$ is connected to the slow-roll parameter $\epsilon_{\rm sr}$ and, consequently, to the tensor-to-scalar ratio as

$$n_{\rm t} = -2\epsilon_{\rm sr} \quad \rightarrow \quad r = -8n_{\rm t}\,. \tag{2.28}$$

Many proposed inflationary models offer different predictions for $\epsilon_{\rm sr}$ and $\delta_{\rm sr}$, but in almost all of them the ratio of tensor-to-scalar modes is directly related to the tensor spectral index. Therefore, extracting the values of r from data would open to the possibility of probing the potential of the field driving inflation, which scale could be on the order of 10^{16} GeV, orders of magnitude out of the reach of modern particle accelerators. Currently, tensor perturbations parameters have not been measured yet, we only have upper limits (at 95% CL):

$$\begin{aligned} &r < 0.036 \quad (\text{fixed } n_{\rm t})\,,\\ &r < 0.028 \quad (\text{free } n_{\rm t})\,,\\ &-1.37 < n_{\rm t} < 0.42\,, \end{aligned} \tag{2.29}$$

from [205] and [81], respectively. Unfortunately, there is no guarantee that gravitational waves produced during inflation will be detectable, as their amplitude depends on the energy scale of the potential. If inflation takes place at scales sufficiently smaller than the Planck scale, then primordial gravitational waves will not be detectable by any future experiment, being lower than fundamental quantum noise limits and sample variance due to observing a single Universe. The measurement of this parameter is one of the main frontiers in modern Cosmology, being the primary targets of many cosmological experiments observing the CMB. It is also the central focus of this Thesis.

2.3 Models of inflation

In the previous section we saw that the slow-roll parameters $\epsilon_{\rm sr}$ and $\delta_{\rm sr}$ quantify all the predictions of the inflationary model (at least for a single field model). Thus, different field potential V can be compared against observations. For many inflationary models, the power spectrum parameters can be linked with the number of e-folds, as $n_{\rm s} \simeq 1 - 2/N_*$ and $r \simeq 12/N_*$. In order for inflation to solve the Hot Big Bang problems, the number of e-folds should be $N_* \gtrsim 50$; this implies a small

but non zero deviation from 1 of $n_{\rm s}$ (red-tilting), as well as a small value of r. We summarise here some observational predictions induced by the existence of an inflationary epoch in the very early Universe:

- Nearly scale-invariant primordial power spectrum ($n_{\rm s} \lesssim 1$ and $n_{\rm t} = 0$), with the scalar index slightly lower than 1. This agrees very well with CMB measurements, as will be detailed in chapter 3.

- The existence of a gravitational waves background producing a tensor-mode power spectrum, usually parametrised by the tensor-to-scalar ratio r. Primordial gravitational waves have not been detected yet.

- Gaussian primordial perturbations are also expected, in agreement with the CMB observation so far. Small deviations from primordial Gaussianity in the perturbation statistics are also predicted in some inflationary models [18]. These are not detected yet and can also be investigated using the CMB.

To conclude, we report some of the most commonly cited models of inflation, either for their simplicity or for the connection with observational data, referring to [135] for more details:

► **Chaotic inflation:** one of the first working class of inflationary models, introduced in 1983; simple phenomenological models where the potential evolves as a power law:

$$V(\phi) \propto \phi^{\alpha} \,. \tag{2.30}$$

These models do not assume much on the initial conditions, and thermal equilibrium in the early Universe is not required [126]. Predictions of these models include

$$n_{\rm s} - 1 \simeq -\frac{\alpha + 2}{2N_*} \,, \qquad r \simeq \frac{4\alpha}{N_*} \,. \tag{2.31}$$

► **Axion-like inflation**: these models consider a potential typical of an axion-like particle:

$$V(\phi) = V_0 \left[1 + \cos\left(\frac{\phi}{\mu}\right)\right] \,, \tag{2.32}$$

where V_0 is the vacuum energy density driving the accelerated expansion. These models are also called "natural inflation".

► **Hilltop inflation**: generic class of inflationary models where inflation takes place near a maximum of the potential, with the inflaton field rolling away from an unstable symmetric point at $\phi = 0$ toward a true vacuum [31, 177]. The

potential is parametrised as

$$V(\phi) = V_0 \left[1 - \left(\frac{\phi}{\mu}\right)^{\alpha} + \cdots\right] . \tag{2.33}$$

These models are also called "small-field" models. They usually predict a very small tensor-to-scalar ratio, $r \lesssim 0.002$.

- R^2 **or Starobinsky inflation**: it is possible to generalise GR by adding higher-order terms in the Einstein-Hilbert (EH) action [194, 195]. One of the simplest possibility is to add a quadratic term as $\mathcal{L}_{\rm grav} = \mathcal{L}_{\rm EH} + R^2/(6m^2)$, which could arise from quantum effects in a quantum theory of gravity. Under certain conditions, this theory is formally equivalent to GR with a scalar field of mass m able to produce inflation in the early Universe, with a potential of the form

$$V(\phi) = \frac{3}{2} m^2 \left(1 - e^{-\sqrt{2/3}\phi}\right)^2 . \tag{2.34}$$

These models give predictions on primordial observables as

$$n_{\rm s} \simeq 1 - \frac{2}{N_*} , \quad r \simeq \frac{12}{N_*^2} , \tag{2.35}$$

with a predicted tensor-to-scalar ratio around $r \simeq 0.004$, given that $50 < N_* < 60$, which could be targeted by next-generation space-based telescopes. These models are also able to provide predictions about higher-order parameters like the running of the scalar spectral index or the negative tensor spectral tilt.

- **Higgs inflation**: the Standard Model of particles already contains a known scalar field, which is the Higgs boson. It can theoretically be an explanation for inflation, where the potential assumes a form as

$$V_{\rm H}(\phi) = V_0 \left(1 + e^{\frac{-2\phi}{\sqrt{6}M_P}}\right) . \tag{2.36}$$

This model could also provide a natural mechanism for reheating through the Yukawa couplings. Unfortunately, the interactions and properties of the Higgs field are sufficiently constrained by now to know that it cannot serve as the source for inflation, or at least it is very difficult to make it work [62].

Different models of inflation can be constrained or ruled out by measurements of the scalar spectral index and tensor-to-scalar ratio parameters, which can be used to build a $(n_{\rm s}, r)$ plane as illustrated in figure 2.3, where CMB and BAO data are combined together to draw contours. Some models are less compatible than

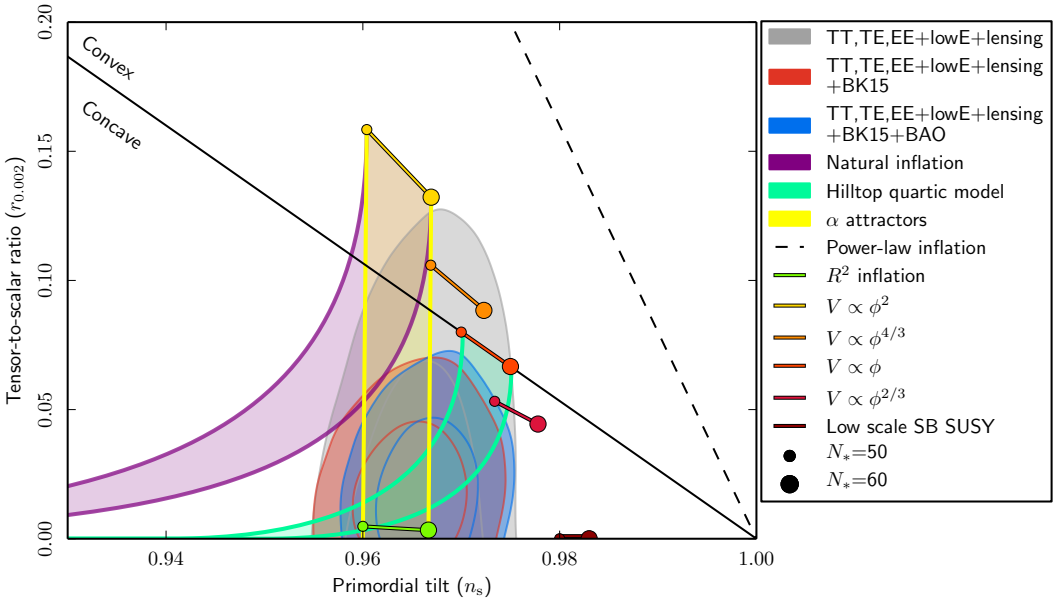


Figure 2.3. Marginalized joint 68% and 95% CL regions for n_s and r at $k_p = 0.002\,\mathrm{Mpc}^{-1}$ from *Planck* data alone and in combination with *BICEP/Keck* and BAO data, compared to the theoretical predictions of some inflationary models. Figure from [167].

others, already giving insights to what could have happen in the very early Universe. We will describe in more detail how this measurements can be obtained from CMB observations, in the next chapter.

Chapter 3

The Cosmic Microwave Background

The Cosmic Microwave Background (CMB) radiation is largely recognised as one of the main observational pillars of Cosmology, representing the foundation of our knowledge about the Universe. It is considered as the key experimental evidence for the Hot Big Bang theory, being both a prediction and a confirmation. These photons are the relic radiation of the primordial hot and dense Universe, when it was 380,000 years old. The radiation uniformly illuminates the microwave sky, permeating the entire spacetime. A radio telescope can detect the CMB photons, coming from every direction pointed by the telescope, all with the same temperature. Since the photons have travelled for ~ 13 Gyrs, they lost a lot of energy during the trip; the temperature we measure now ($T_{\rm CMB} = 2.725\,{\rm K}$) is about 1000 times lower than the one they had at emission time ($T \sim 3000\,{\rm K}$). The CMB thus offers a direct window on what was going on in the primordial Universe and beyond, providing information on the evolution of the Universe, structure formation, the interplay of sub-atomic particles, and even on inflationary physics. The CMB is probably the most important trove of information we have at our disposal in Cosmology.

As we briefly mentioned at the beginning of Part I, at the time of the discovery of the Universe expansion by Hubble and Humason [110], there were already various speculations [120] about the shape of the early Universe, with the expansion hinting toward a contracted, hot, and dense version of the cosmos. As narrated in [216] and [154], the first indirect discoveries of the CMB were actually made around 1940, when several scientists measured the rotational emission lines of molecules in the interstellar medium, finding them in the first excited status induced by a $\sim$ 3 K astrophysical background [141]. These discoveries were not recognised as such by the authors, since the first reasonable prediction of a relic radiation was made years later by Alpher and Herman in 1948 [14], which estimated the

temperature of the background to be not higher than 5 K. The first attempt to connect theory and observations was made by Russian physicists Doroshkevich and Novikov [63] in 1964, who referred to a Bell Labs paper [152] reporting about a ~ 3 K excess measured at the Crawford Hill horn telescope. They misjudged the outcome of the system characterisation, and were not able to properly connect the dots. Penzias and Wilson inherited the horn telescope excess and improved its characterisation, leading the official, "accidental" discovery of the CMB, with the cosmological interpretation given by Dicke's group [60] in Princeton, where they were trying to land the same exact experiment.

From that moment, Cosmology started to harvest the tremendous source of information that the CMB is. As of today, the majority of the knowledge we have about our Universe, its origins, content, and evolution, is founded on what we have extracted from this extremely rich probe. The CMB observable and its analysis is the main focus of this Thesis; in the remaining of this Part, I will describe the physics of the CMB, its main features, how it is observed, and what are the main experimental frontiers tackled by the community. In particular, I will focus on the topics to which I contributed the most during my PhD.

3.1 From the primordial plasma to the CMB

In section 1.2.1 we laid down the chronology of the Universe, as per the Hot Big Bang theory in the ΛCDM scenario. After the Big Bang Nucleosynthesis, the Universe is made of interacting photons, electrons, and baryons, tightly coupled in a so-called photo-baryonic plasma. The very high temperature keeps the atoms ionised, and the photons mean free path $\lambda = 1/n_{\rm e}\sigma_{\rm t}$ is small, where $n_{\rm e}$ is the electron density and $\sigma_{\rm t}$ is the Thomson scattering section. The Universe is opaque and photons are continuously interacting through Compton scattering with electrons; in other words, the interaction rate $\Gamma = c/\lambda = n_{\rm e}\sigma_{\rm t}c$ is greater than the expansion rate, described by the Hubble parameter H, i.e. $\Gamma \gg H$, and the photons mean free path is smaller than the horizon. The plasma is at thermal equilibrium, everything has the same temperature: it is a perfect thermal blackbody at a certain common temperature $T_\gamma = T_{\rm b}$.

The CMB emission is not an instantaneous process, rather a succession of events lasting some thousands of years, around $t \sim 380,000$ ($z \simeq 1100$). It can be divided in three separated events:

- **Decoupling.** As the Universe expands, the electron are more diluted over the cosmic volume and their density gets lower. The photons mean free path becomes larger, and interactions with electrons are less likely to happen.

Moreover, the temperature of the plasma lowers, as the Universe expand adiabatically ($T \propto 1/a$). When the interaction rate becomes equal and then smaller than the expansion rate ($\Gamma = H$), the photons decouple from the electrons, as neutrinos similarly did at earlier times.

- **Recombination.** After the beginning of decoupling, electrons are not scattered by photons as frequently as before. They are able to bound with baryons in order to form hydrogen atoms. The Universe goes from being fully ionised to being almost entirely neutral, becoming "transparent" to photons. This is actually the first time electrons and protons combine into hydrogen in the history of the Universe, with the "re" of recombination being misleading. Decoupling is an initially gradual process which then becomes abrupt after the beginning of recombination, since the particles available for photons' scattering are both diluted and captured by baryons, basically disappearing.

- **Last scattering.** As the last electrons recombine with protons, photons are able to interact one last time with them before becoming free to propagate through the entire Universe. This time region is called "last scattering surface". These photons will travel for ~ 13 Gyrs before reaching us on Earth, detectable as the CMB.

As we saw in the previous chapter, the physical conditions typical of these epochs are very well known, in the framework of the Hot Big Bang and cosmological principle assumptions. Physical interaction at these times are rather simple, and deviations from this picture can be classified at the linear order. Therefore, our tools are more than enough to completely characterise the origins and the observable features of the CMB, which were predicted and calculated years before the actual measurements. We will now describe some properties of this probe.

3.2 Main features of the CMB

The CMB is an isotropic emission of uniform blackbody thermal energy. This emission at present day is peaked at millimetre wavelengths, or equivalently in the microwave frequencies of the electromagnetic spectrum. Its intensity is usually measured in temperature units, estimated to be $T_{\rm CMB} = 2.72548 \pm 0.00057\,{\rm K}$ by the *FIRAS* instrument onboard the *COBE* satellite experiment [140, 219, 78, 77]. At the time of its emission, the CMB temperature would be $T(z = z_{\rm CMB}) = (1 + z_{\rm CMB})T_{\rm CMB} \simeq 3000\,{\rm K}$: without the expansion of the Universe cooling the CMB, the night sky would shine in the visible as the Sun does.

The CMB contains the vast majority of photons in the Universe, with 400 photons per cubic centimetre. The number density of CMB photons is one billion times the number density of matter in the Universe, and its present-day energy density greatly exceeds that of the photons emitted by all the stars over the history of the Universe. The CMB temperature (also called the monopole) is remarkably smooth and uniform across the sky, in accordance with the cosmological principle of homogeneity at the largest scales; small temperature variations are registered at the 10^{-3} (dipole) and 10^{-5} level. The latter variations are called anisotropies, and the main goal of many modern experiments is to accurately measure them, as they hold a wealth of information about our Universe. Their properties will be described in section 3.3. The dipole is caused by the peculiar velocity of the Sun relative to the comoving CMB rest frame, as it moves at $369.82 \pm 0.11\,\mathrm{km/s}$ towards the constellation Crater near its boundary with the constellation Leo [168]. This motion induces a relativistic Doppler shift of the CMB photons, creating a dipole across the sphere with an amplitude of $3.3666 \pm 0.0027\,\mathrm{mK}$ in our direction of motion. The monopole and dipole of the CMB are shown on the left side of figure 3.1, expressed as deviations from the mean value, in thermodynamic units ($\mathrm{K_{CMB}}$). In these units, the CMB signal is independent of the frequency.

Being the snapshot of a gas in thermal equilibrium, the brightness of the CMB radiation is extremely close to a blackbody, varying in frequency (or wavelength) accordingly to the Planck's law:

$$B(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_\mathrm{B}T} - 1}\,, \tag{3.1}$$

where k_B is the Boltzmann constant, h is the Planck constant, c is the speed of light, and the emission is irrespective of shape and composition of the body. The right side of figure 3.1 reports the CMB intensity spectrum (technically called the monopole) as firstly measured by the *FIRAS* instrument, part of the *COBE* space mission [140, 219, 78] operating in the 90s. The accuracy of this measurement is such that uncertainties in the data points are smaller than the line width. Departures from the perfect blackbody emission have not been measured yet, only upper limits have been placed. Nonetheless, faint deviations (called spectral distortions) induced by local variations of the thermal equilibrium are expected to exist, possibly hiding important information on primordial physics [50, 51].

The CMB radiation is also slightly linearly polarised at the 10% level, due to the Thomson scattering at the last scattering surface. The accurate mapping of the faint polarised signal has now become the main frontier of CMB science, as it is forecasted to contain an enormous amount of information on primordial physics and the origin of the Universe. We will describe it in more detail in the next sections.

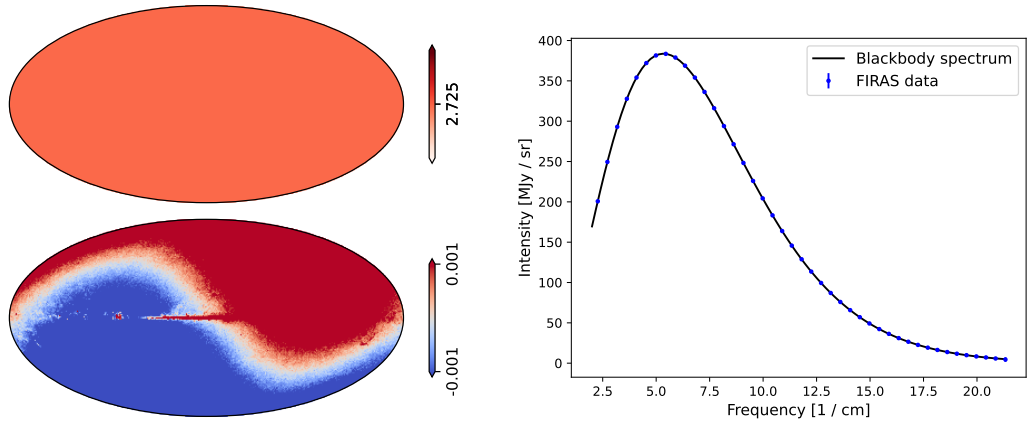


Figure 3.1. *Left:* monopole and dipole maps of the CMB intensity. Upper panel illustrate the smooth isotropic CMB temperature $T_{\rm CMB} \simeq 2.725\,{\rm K}$, almost the same in every pixel. The lower panel shows a *Planck* intensity map at a frequency of 100 GHz [173], after removal of the monopole, in $\mathrm{K}_{\rm CMB}$ units. The dipole anisotropy is evident in the hotter and colder regions, typical of a Doppler shift. *Right:* the monopole spectrum of the CMB, from the original data by the *FIRAS* instrument of the *COBE* experiment [140, 219, 78], available at this link. The vertical axis is in surface intensity units, while the horizontal axis is the reciprocal of the microwave wavelength, proportional to the frequency. The black curve shows the expected intensity from a single temperature blackbody spectrum at $T = 2.725\,{\rm K}$. The error bars are there in the plot, smaller than the line width (!).

3.3 CMB anisotropies

The CMB anisotropies are usually divided into two types: primary anisotropies, due to effects that occur at the last scattering surface and earlier; and secondary anisotropies, due to effects and interactions that occur between the last scattering surface and the observer, which can significantly alter the primary ones. Primary anisotropies are further separated according to their impact on temperature and polarisation observables. In this section we provide a brief and physically intuitive summary about the origin of these fluctuations, referring the reader to standard textbooks and papers [30, 108, 109, 35, 190, 107, 123, 62] for a more formal and complete derivation.

3.3.1 Temperature anisotropies

Figure 3.2 shows the CMB temperature sky maps as observed by the *Planck* experiment, for its nine frequency bands. The fluctuations have a similar structure across different frequencies, with the main difference being the contamination spreading from the horizontal region that is the Galactic plane of the Milky Way (detailed in chapter 4). The anisotropies we see today are a snapshot of the primordial photo-baryonic plasma at the time of recombination. Back then, photons, electrons,

and baryons were tightly coupled and continuously interacting with each other. The other key ingredient is the presence of cold dark matter over-densities; they do not interact electromagnetically with other particles, but exert gravitational attraction on the massive baryons. Therefore, the primordial plasma is in equilibrium between different interactions: baryons are gravitationally attracted by dark matter wells, but the radiation pressure from photons works in the opposite direction. This interplay generate baryonic oscillations in the photo-baryonic plasma, which propagate as sound waves in the gas (usually called "acoustic oscillations"). After recombination, photons and baryons are not coupled any more, and these oscillations are "frozen" and imprinted in the CMB anisotropies pattern. Therefore, in the standard ΛCDM model, CMB fluctuations are sourced by three different contributions: gravity, the Doppler effect, and intrinsic temperature variations. The total deviation from the average temperature $\bar{T}$ can be written as

$$\frac{\Delta T}{\overline{T}} = [\Psi + \hat{n} \cdot \mathbf{v} + \Phi]_{\tau=\tau_{\rm LS}} \ , \tag{3.2}$$

where Ψ is essentially the Newtonian gravitational potential, Φ is the perturbation to the spatial curvature, $\mathbf{v}$ is the baryon peculiar velocity field, $\hat{n}$ is the pointing unit vector indicating the direction on the sky, and $\tau_{\rm LS}$ is the time of the last scattering epoch. In more detail:

- **Gravitational shift (Sachs-Wolfe effect).** Regions of the fluid with higher density will have a stronger gravitational potential Ψ. Photons travelling in this perturbed field will be gravitationally blue-shifted or red-shifted, associated with a temperature variation of

  $$\frac{\Delta T}{\overline{T}} \simeq \frac{1}{3}\delta\Psi \, . \tag{3.3}$$

  Photons passing through a potential well (over-densities) will lose energy and become colder, while photons passing through an under-dense spot will gain energy and become hotter.

- **Doppler effect.** Local portions of the fluid moving at velocity $\mathbf{v}$ will induce a Doppler shift of the emitted CMB photons, resulting in a change of temperature

  $$\frac{\Delta T}{\overline{T}} \propto \mathbf{v} \, . \tag{3.4}$$

- **Density perturbations.** Over-densities in the primordial plasma are expected to be warmer, while under-densities are colder. These fluctuations $\delta\rho$ generated by the inflationary mechanism evolve through the perturbation equation

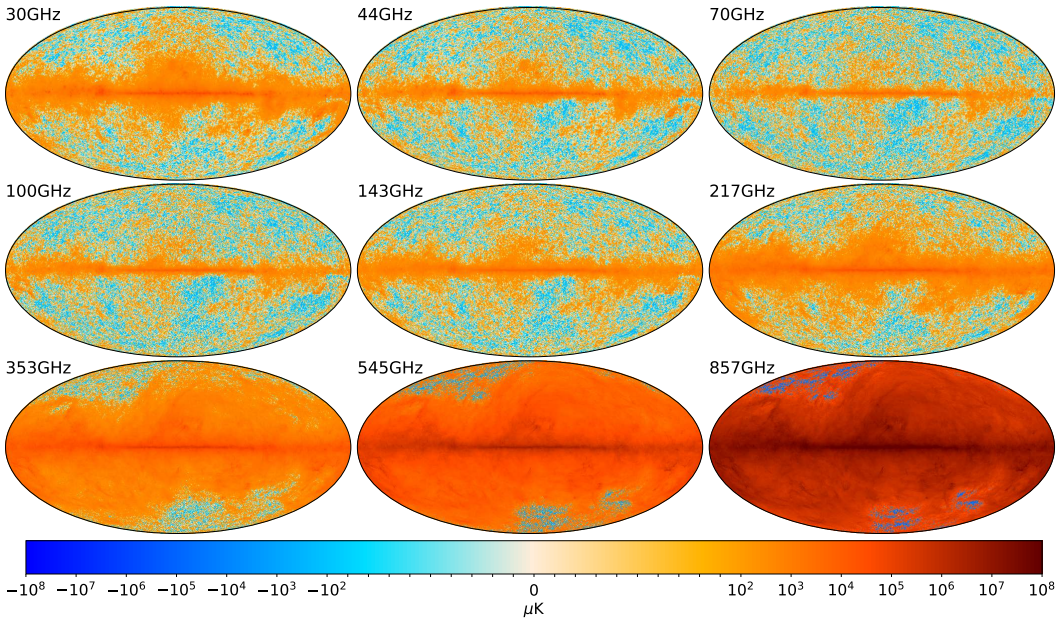


Figure 3.2. Sky emission fluctuations in each of the nine *Planck* frequency bands, after removal of the common dipole component. The fluctuations are expressed as equivalent temperature variations (in $K_{\rm CMB}$ units), so that CMB fluctuations will appear the same in each map. Figure from [173].

until decoupling, generating a temperature shift as

$$\frac{\Delta T}{\overline{T}} \simeq \frac{1}{3}\frac{\delta\rho}{\rho}\,. \tag{3.5}$$

Density perturbations and gravitational shifts are responsible for the acoustic oscillations of the photo-baryonic plasma; the Doppler shift generates the same acoustic oscillations, but out of phase with the others[1]. Therefore, the measurement of these acoustic oscillations (through correlation functions and power spectrum estimation) is a powerful way to extract valuable information on the physics and content of the early and modern Universe; in particular, anisotropies observations are able to put constraints on the matter densities $\Omega_{\rm c}h^2$ and $\Omega_{\rm b}h^2$, but also on the curvature Ω_κ, the amplitude of primordial perturbations $A_{\rm s}$, and the spectral index $n_{\rm s}$ (although the constraint on $A_{\rm s}$ is limited by the degeneracy with the reionisation optical depth $\tau_{\rm rei}$).

3.3.2 Polarisation anisotropies

The information encoded in the CMB anisotropies is not entirely exhausted by their temperature. The thermal history derived from the Big Bang theory predicts the

[1] Being out of phase is a typical feature between the positions and velocities of oscillators.

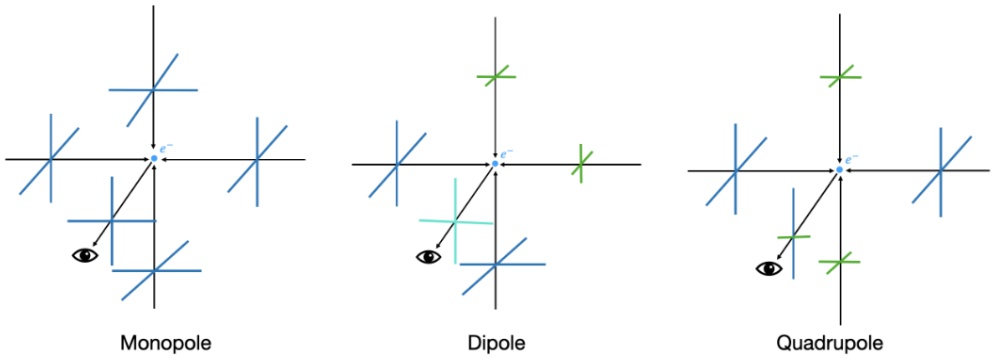


Figure 3.3. Caricature of the Thomson scattering, comparing the output linear polarisation generated by monopole, dipole and quadrupole configurations. The monopole and dipole generate unpolarised light, while the quadrupole leaves a net linear polarisation on the outgoing photons. Figure from [207].

anisotropies to be slightly polarised, at the $\sim 10\%$ level (order of μK). The origin of this polarisation is to be found in the collisional interactions between photons and electrons in the photo-baryonic gas. Before decoupling, the plasma is in thermal equilibrium; photons and electrons are continuously interacting through Thomson scattering[2] $e^- + \gamma \to e^- + \gamma$. The energy/frequency of the photon is not changed by the interaction. The Thomson scattering differential cross section depends on polarisation as

$$\frac{\mathrm{d}\sigma_\mathrm{T}}{\mathrm{d}\Omega} \propto |\hat{\varepsilon} \cdot \hat{\varepsilon}'|^2 \,, \tag{3.6}$$

where $\hat{\varepsilon}$ and $\hat{\varepsilon}'$ are the incident and scattered directions, respectively, and $\hat{\varepsilon} \cdot \hat{\varepsilon}' = \cos\varphi$, with φ being the angle between the two directions. If the incoming unpolarised radiation field is isotropic, orthogonal polarisation states from incident directions separated by 90 degrees are going to balance, so that the outgoing radiation is unpolarised. The same applies to an incoming field showing a dipolar component. Conversely, if the incident radiation field possesses a quadrupolar variation in temperature, the result is a linear polarisation of the scattered radiation. A reversal in sign of the temperature fluctuation corresponds to a 90 degrees rotation of the polarisation, which reflects the spin-2 nature of polarisation[3]. Before recombination, Thomson scattering is rapid, and the randomisation of photon directions destroys any quadrupole anisotropy and, consequently, polarisation. At the time of last scattering, the polarised emission is free to propagate to us and is not cancelled any more. Only photons that scattered in an optically thin region near last scattering

[2] which is the elastic, low-energy limit of the Compton scattering.

[3] This is usually phrased by saying that polarisation is a "headless vector", a rod with no arrow. A rotation of π (180°) brings the vector to a position equivalent to the initial one, and during a rotation of 2π (360°) the vector has come back to itself 2 times.

could have had a quadrupole anisotropy, and that is the reason why the polarisation fraction is small and dependent on the duration of last scattering.

An electromagnetic signal is generally described by the four Stokes parameters I, Q, U and V. They are defined starting from the oscillating electric field $\mathcal{E} = \mathcal{E}_x\hat{x} + \mathcal{E}_y\hat{y}$ of a plane wave propagating in the $\hat{z}$ direction in the $(\hat{x}, \hat{y}, \hat{z})$ frame. The Stokes parameter can be written as

$$\begin{aligned} I &= \langle|\mathcal{E}_x|^2\rangle + \langle|\mathcal{E}_y|^2\rangle\,, \\ Q &= \langle|\mathcal{E}_x|^2\rangle - \langle|\mathcal{E}_y|^2\rangle\,, \\ U &= 2\mathrm{Re}(\langle\mathcal{E}_x\mathcal{E}_y^*\rangle)\,, \\ V &= -2\mathrm{Im}(\langle\mathcal{E}_x\mathcal{E}_y^*\rangle)\,, \end{aligned} \tag{3.7}$$

where z^* is the complex conjugate, $|z|^2 = z^*z$, and $\langle\cdot\rangle$ is the time average. I corresponds to the energy density of the electric field, representing the total intensity of the signal which is also referred to as the temperature of the field. Q and U quantify the linear polarisation, while V account for circular polarisation. Thomson scattering is not expected to generate significant circular polarisation, and V is usually neglected in Cosmology, with Q and U fully describing the CMB polarisation. They can also be combined into different mathematical quantities, like the spin-2 complex polarisation field $\mathcal{P} = Q + iU$, and its amplitude $P^2 = Q^2 + U^2$ and phase $\alpha = \frac{1}{2}\tan^{-1}(U/Q)$, or in a symmetric, traceless polarisation tensor Π defined as

$$\Pi = \begin{pmatrix} Q & U \\ U & -Q \end{pmatrix} = Q\sigma_z + U\sigma_x\,. \tag{3.8}$$

where σ_z and σ_x are the Pauli matrices. Figure 3.4 shows a view of the polarised microwave sky as observed by the seven polarised frequency channels of the *Planck* satellite. We notice how the CMB anisotropies are evidently contaminated by the presence of the Milky Way, intervening between us and the last scattering surface.

One practical problem of the Q and U parameters is their dependence on the coordinates system $(\hat{x}, \hat{y}, \hat{z})$, meaning that a change of coordinates will modify their value. In particular, following a rotation of the axes by an angle α, the Q and U parameters will transform as:

$$\begin{aligned} Q' &= Q\cos 2\alpha + U\sin 2\alpha\,, \\ U' &= -Q\sin 2\alpha + U\cos 2\alpha\,, \end{aligned} \tag{3.9}$$

while I does not change, being a scalar quantity. It is common in CMB science to change the basis in order to work with coordinate-independent quantities. To do so,

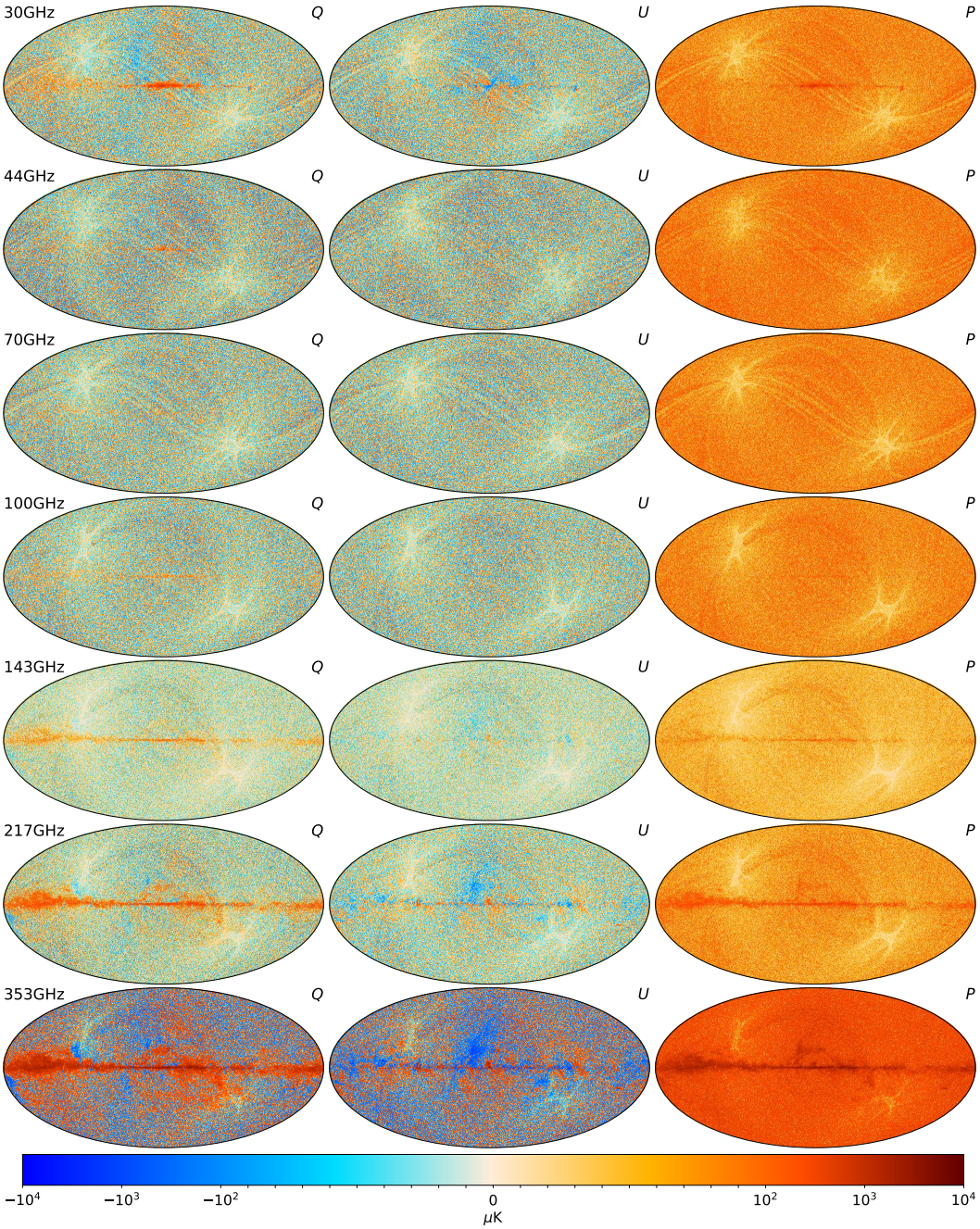


Figure 3.4. Polarised microwave sky as observed by the seven frequency bands of the *Planck* satellite. The first two columns show the Q and U Stokes parameters; the last column shows the polarised intensity, $P = \sqrt{Q^2 + U^2}$. Figure from [173].

the same scalar-vector-tensor decomposition of section 1.3.1 is exploited to split the polarisation tensor into the sum of two components, corresponding to a parity symmetric curl-free and a parity odd divergence-free quantity. Being Π symmetric and trace-free, there are two constraints over 4 components; thus, 2 real quantities are enough to fully describe it. The tensor can be decomposed as

$$\Pi = f(E) + g(B) \tag{3.10}$$

where f and g are two functions (that we do not treat here) of the parity-even E and parity-odd B quantities. Contrary to Q and U, E and B are scalar fields and thus coordinate invariant. These two quantities are usually called E-modes and B-modes of the polarisation field, from which cosmological information is extracted [221, 191, 111]. Under a parity transformation (reflection in a mirror), E-modes do not change, while B-modes change sign. This is why the B-mode field is technically a pseudo-scalar. The geometrical patterns of these two modes can be intuitively understood by considering that E-modes are curl-free ($\nabla \times E = 0$) and B-modes are divergence-free ($\nabla \cdot B = 0$); the former will show a radial shape, while the latter will have a "swirly" pattern. Figure 3.5 illustrates these two behaviours as they would be observed in CMB maps, where the length of the rods is proportional to the amplitude of the polarised signal, and the underlying colour represent the temperature anisotropies. Thus, quadrupolar asymmetries in the primordial plasma can polarise the CMB light and generate specific patterns in maps of the CMB emission. How are these quadrupoles sourced?

- **Scalar perturbations (density fluctuations).** Electrons falling into the potential well of an over-dense region will be subject to gravitational shifts due to the radial nature of the gravitational force ($\propto R^{-2}$). In its own reference frame, a single electron will see the electrons in front of and behind it move away from it, because particles in front fall faster toward the centre, while electrons behind fall slower. The electron receives red-shifted light. At the same time, the electron sees its right and left neighbours getting closer, as they are all falling towards a common point. Hence, it receives their light as blue-shifted. It experiences then a quadrupolar asymmetry, which generates linear polarisation. The same applies to under-densities, from which electrons are moving away in a decelerated way. This mechanism can be visualised in figure 3.6; the resulting polarisation will appear as radial or concentric rods surrounding the over/under-densities. This pattern corresponds to a pure E-mode signal; no B-modes can be generated by density fluctuations.

- **Tensor perturbations (gravitational waves).** in sections 1.3.1 and 2.2 we

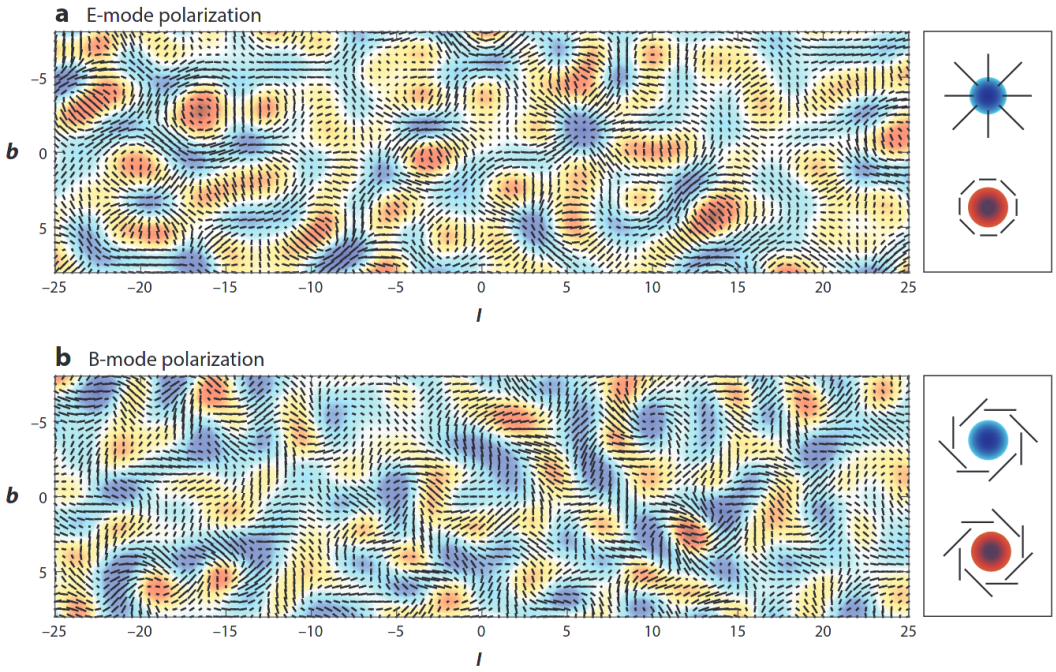


Figure 3.5. *Upper panel:* polarisation pattern composed only of E-modes. *Lower panel:* polarisation pattern composed only of B-modes. As indicated on the right, the E-mode pattern is tangential around hot spots (red) and radial around cold spots (blue). B-mode polarisation show instead a characteristic swirling pattern (with different orientations around hot and cold spots). Figure taken from [111].

have seen that the inflationary theory predicts the existence of a stochastic background of gravitational waves, propagating through spacetime as a metric perturbation. The two polarisation states h_+ and $h_\times$ create distortions in the spacetime as they travel: a circle of particles is deformed into ellipses whose axis are oriented differently for the two states. This creates a situation similar to the density perturbations one: each tensor state induces a quadrupolar asymmetry, that in turn generates linear polarisation. Since we have two deformations with an inclination of 45° between them, the total pattern will be a mixture of E- and B-modes, as it is shown in figure. Thus, tensor perturbations can generate both E- and B-modes.

The polarisation of CMB anisotropies contains a huge amount of cosmological information, which is additional and complementary to the one provided by temperature. The measurement of the E-mode pattern, generated by both scalar and tensor perturbations, contributes to breaking the degeneracy on some cosmological parameters of the ΛCDM model. It also provides indirect proofs of the inflationary paradigm. The B-mode pattern, instead, can only be sourced by primordial gravitational waves. This is extremely crucial for CMB science: it means the if we see a

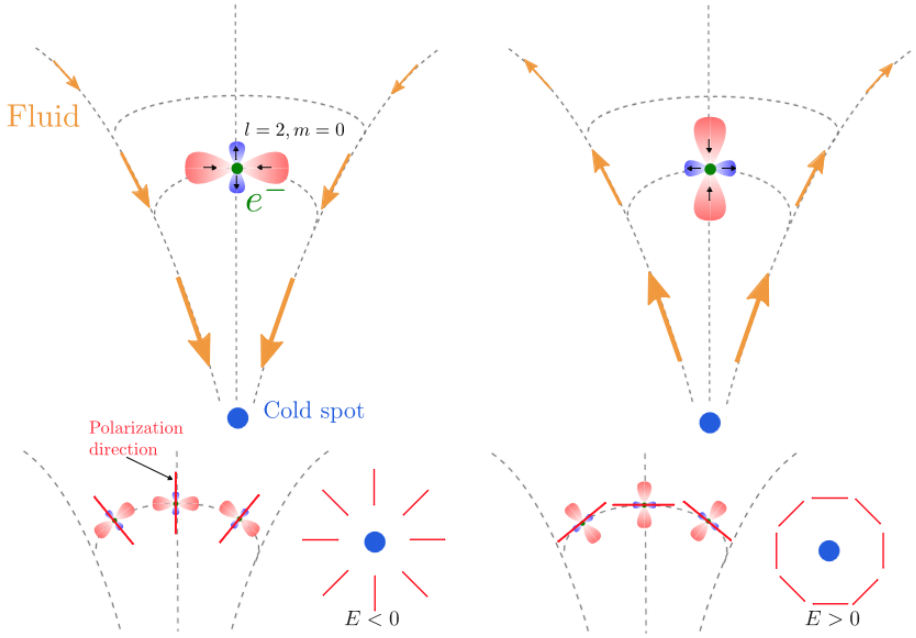


Figure 3.6. Generation of E-mode polarisation from density perturbations. The fluid either falls into the potential well while accelerating, or climbs up while decelerating (yellow arrows), creating quadrupole anisotropy in the radiation. Black arrows show the fluid motion in the electron rest frame; red lines trace the electric field direction of the radiation scattered by the electrons, resulting in positive or negative E-mode polarisation patterns. Figure from [147].

B-mode pattern in CMB polarisation, we are directly detecting a predicted product of inflation. Scalar perturbation can not produce B-modes. As we discussed in section 2.2, tensor perturbation are usually parametrised in terms of the tensor-to-scalar ratio r, which is an observable quantity in experiments:

$$r = \frac{A_{\rm t}}{A_{\rm s}} \,. \tag{3.11}$$

In turn, the tensor-to-scalar ratio can be directly linked to the potential of the scalar field driving inflation, and thus to the energy scale characteristic of the inflationary phase:

$$V^{1/4} \simeq 10^{16}\,{\rm GeV} \left(\frac{r}{0.01}\right)^{1/4} \,. \tag{3.12}$$

This is the main reason why B-modes are considered as the frontier in CMB science: a detection of this pattern in CMB polarisation would open a new window on the physics governing the first moments of the Universe, at energy scales which will not be probed by particle accelerators on Earth for many years. Currently, tensor perturbations parameters have not been measured yet, with only upper limits (at 95% CL) coming from the combination of the most advanced instruments available

(*Planck*, *BICEP/Keck*):

$$\begin{aligned} r &< 0.036 \quad \text{(fixed } n_{\rm t}\text{)}\,, \\ r &< 0.028 \quad \text{(free } n_{\rm t}\text{)}\,, \\ -1.37 &< n_{\rm t} < 0.42\,, \end{aligned} \tag{3.13}$$

from [205] and [81], respectively. We note here that a detection of primordial B-modes has been claimed in 2014 by the *BICEP* collaboration, reporting a measurement of the tensor-to-scalar ratio of $r = 0.20^{+0.07}_{-0.05}$ [4]. A later joint reanalysis with *Planck* data [27] revealed this signal to be consistent with polarisation from Galactic dust, contaminating the measurement and inducing a false detection. We will face this topic in more details in the following Parts of this Thesis.

3.3.3 CMB secondary anisotropies

After the recombination and last scattering phases, CMB photons are decoupled from other particles and are free to travel toward our planet. Even if so, they are expected to occasionally interact during their long trip, since the rest of the Universe was able to form structures in the meantime. These (quite) rare interaction can modify the original radiation emitted at the last scattering surface, in several ways:

- **Reionisation.** Before the birth of the first stars, the Universe is fully neutral and no source of light is present. The neutral hydrogen clumps up to ignite the first stars, which are very luminous and emit predominantly in the UV, and in particular they produce photons with energies of 13.6 eV. This is the amount of energy required for hydrogen ionisation, and thus the majority of the atoms present in the Universe become ionised again[4]. This creates a new scattering surface, where photons can interact with the free electrons through Thomson scattering, and thus producing additional linear polarisation. This effect is more visible on the large-scale anisotropies.

- **Integrated Sachs-Wolfe (ISW, or Rees-Sciama) effect.** We already saw this effect while describing the origins of temperature anisotropies in section 3.3.1. Photons passing through different gravitational potentials will gain or lose energy (blue- or red-shifted). The same can happen when the photons travel through gravitational wells and hills intervening between us and the last scattering surface. In particular, photons encountering superclusters or voids will experience a shift in energy. This effect is more pronounced after dark energy starts to dominate the evolution of the Universe, because the accelerated expansion makes the gravitational potential to decay over the

[4] Unlike recombination, this name is actually accurate, since atoms are really ionised *again*.

time a photon is traversing them. A photon gets a kick of energy going into a potential well (a supercluster), and it keeps some of that energy after it exits, after the well has been stretched out and shallowed. Similarly, a photon has to expend energy entering a void, but will not get all of it back upon exiting the slightly reduced potential hill [181].

- **Gravitational lensing.** CMB photons gravitationally interact with the very large structures present in the Universe, which are able to deflect their trajectories and modify the original anisotropies pattern [222, 121, 192]. This effect is particularly impactful in B-modes polarisation analysis. More details on it are reported in chapter 4.

- **Sunyaev-Zel'dovich effect.** The environment in which galaxy clusters are embedded is called the intra-cluster medium (ICM). It is mostly made of hot, ionised plasma, where electrons are heated due to gravitational interactions. Photons can interact with these electrons through inverse Compton scattering, gaining energy and shifting frequency, resulting in a spectral distortion of the CMB. This process is known as the thermal Sunyaev-Zel'dovich (tSZ) effect [198, 28]. The peculiar motion of the electrons can also imprint a Doppler shift on the interacting photons, in a process called kinetic SZ (kSZ) effect.

All these features must be taken into account when analysing CMB data. In most of the cases they represent a contamination of the original, primordial signal that we want to recover. Nonetheless, the physics they rely on is rich of astrophysical and cosmological information, and are thus of fundamental importance in branch of astrophysics which use the CMB as a backlight for the study of the rest of the Universe.

3.4 The CMB angular power spectrum

Back in section 1.3.1, we noted that it is impossible for us to derive a theory predicting the exact distribution of density perturbations at any given point of the spacetime. This is why cosmological theories prefer to rather talk about a statistical description of the field, providing some statistical quantities reflecting information on the ensemble. To do so, we usually work in Fourier space, which simplifies a lot of tasks behind a couple of assumption, as the independency of Fourier modes and the idea of our single Universe being a representative of the underlying fiducial distribution. The same thing can be done to study the statistical properties of the CMB, but now we have to work on the sphere rather than a flat patch.

In general, the CMB temperature field depends on the spacetime coordinates $(\mathbf{x}, \tau)$ and on our line-of-sight $\hat{n}$, that is the direction in the sky we are looking at:

$$T = T(\mathbf{x}, \hat{n}, \tau) \tag{3.14}$$

Although this field is defined at every point in space and time, we can observe it only here from Earth ($\mathbf{x} = \mathbf{x}_0$) and now ($\tau = \tau_0$)[5]. The only anisotropies dependence we can access is the direction of the incoming photons $\hat{n}$. Telescopes typically build sky maps, wherein the temperature is reported at a number of incoming directions, or "locations on the sky", usually labelled by polar coordinates θ and φ. It is possible to transform back and forth as $\hat{n} \rightleftharpoons (\theta, \varphi)$.

In a similar way to the Fourier decomposition, a scalar (spin-0) field defined on the 2-sphere can be expanded in terms of spherical harmonics as

$$T(\mathbf{x}, \hat{n}, \tau) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=\ell} a_{\ell m}(\mathbf{x}, \tau) Y_{\ell m}(\theta, \varphi)\,, \tag{3.15}$$

where the dependence of the temperature field have been separated in a spacetime part $a_{\ell m}(\mathbf{x}, \tau)$ and an angular part $Y_{\ell m}(\hat{n}) = Y_{\ell m}(\theta, \varphi)$. The spherical harmonics $Y_{\ell m}$ are an orthonormal basis (complete set of eigenfunctions) describing the surface of the 2-sphere. They are defined as

$$Y_{\ell m}(\theta, \varphi) = \sqrt{\frac{2\ell + 1}{4\pi} \frac{(\ell - m)!}{(\ell + m)!}} P_{\ell m}(\theta) e^{-im\varphi}\,, \tag{3.16}$$

where $P_{\ell m}$ are the Legendre polynomial. The spherical harmonics expansion can be thought of the analogous of the Fourier expansion, where ℓ is the multipole moment which, as the wavevector $\mathbf{k}$, is associated to the inverse of the angular scale $\ell \sim \pi/\theta$, and m is the azimuthal moment encapsulating the possible orientation on the sphere of a given multipole. The multipole $\ell = 0$ is the monopole (the CMB average temperature) and covers the whole sphere, $\ell = 1$ is the dipole term that divides the sphere in half, $\ell = 2, 3$ are the quadrupoles and octupoles terms, and so on and so forth. The $Y_{\ell m}$ are normalised via

$$\int \mathrm{d}\Omega\, Y_{\ell m}(\hat{n}) Y^*_{\ell' m'}(\hat{n}) = \delta_{\ell \ell'} \delta_{m m'}\,, \tag{3.17}$$

[5] Even if we can observe the effects of the CMB on systems located at different coordinates of the spacetime. At higher redshift the CMB radiation would be at higher temperature and its statistics would be different. An example is the observation of distant galaxy clusters through the Sunyaev-Zel'dovich effect.

and the spherical harmonics expansion can be inverted as

$$a_{\ell m}(\mathbf{x},\tau) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \int \mathrm{d}\Omega\, Y^*_{\ell m}(\hat{n})\, T(\mathbf{k},\hat{n}) \tag{3.18}$$

to give the spherical harmonics coefficients $a_{\ell m}$. All of the information contained in the temperature field is also contained in the $a_{\ell m}$ coefficients.

Again, we are not able to make exact predictions about any particular $a_{\ell m}$. Theories can only speculate about the statistics of these coefficients, and in particular about the underlying distribution from which they are drawn. Quantum fluctuations seeded by the inflationary mechanism suggest this distribution to be a Gaussian, which means that the average value of the $a_{\ell m}$ is zero, but with a non-zero variance. The variance of the $a_{\ell m}$ is given by the angular power spectrum C_ℓ:

$$\langle a_{\ell m}\rangle = 0\,, \qquad \langle a_{\ell m} a^*_{\ell' m'}\rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell \tag{3.19}$$

where $\langle\cdot\rangle$ denotes an ensemble average, which is the result that would be obtained in the limit of measuring over an infinite volume. Thus, for a given ℓ, each $a_{\ell m}$ has the same variance. For example, for $\ell = 100$ (angular size of $\simeq 1^\circ$) we have $2\ell + 1 = 201$ m values drawn from the same distribution, which are a good sample to accurately measure this distribution. Vice versa, for the quadrupole $\ell = 2$ we only have 5 m values to sample the distribution, which will be then measured with less statistical precision. This is a fundamental uncertainty in the knowledge of the C_ℓ, called "cosmic variance", which follows from the C_ℓ being defined as the sum in quadrature of Gaussian stochastic $a_{\ell m}$; their probability distribution is a chi-squared one with $2\ell + 1$ degrees of freedom, and variance

$$\sigma^2_{\text{cosmic}} = \frac{2}{2\ell+1} C_\ell^2\,. \tag{3.20}$$

which is more pronounced at low ℓ. Another way to say this is to notice that, unfortunately, we have only one observable Universe at our disposal to extract information from. In practice, this limit is never achieved, because even if an instrument observes the full sky (such as the satellites *COBE*, *WMAP*, *Planck*, and the future *LiteBIRD*), the large contaminating emission from the Milky Way force us to exclude some parts of the map. If we observe just a fraction f_{sky} of the sky, the measurements precision will be limited by a "sample variance" increased by roughly a factor $1/f_{\text{sky}}$ with respect to the cosmic variance:

$$\sigma^2_{\text{sample}} \simeq \frac{1}{f_{\text{sky}}}\sigma^2_{\text{cosmic}} = \frac{2}{f_{\text{sky}}(2\ell+1)} C_\ell^2\,. \tag{3.21}$$

The contamination to CMB observations will be described in more detail in chapter 4.

Turning to polarisation, the same expansion can be applied with some modifications. The complex, spin-2 polarisation field $\mathcal{P}_\pm = Q \pm iU$ can be expanded in the generalised spin-weighted spherical harmonics on the sphere as

$$\mathcal{P}_\pm(\mathbf{x}, \hat{n}, \tau) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=\ell} {}_{\pm 2}a_{\ell m}(\mathbf{x}, \tau)\, {}_{\pm 2}Y_{\ell m}(\theta, \varphi)\,, \tag{3.22}$$

where ${}_{\pm 2}a_{\ell m}$ are the spin-2 spherical harmonic coefficients. They can be separated in E and B components as

$$\begin{aligned} E_{\ell m} &= -\frac{1}{2}\left({}_2a_{\ell m} + {}_{-2}a_{\ell m} \right)\,, \\ B_{\ell m} &= -\frac{1}{2i}\left({}_2a_{\ell m} - {}_{-2}a_{\ell m} \right)\,, \end{aligned} \tag{3.23}$$

which can then be used to construct the E- and B-modes scalar (spin-0) fields as

$$\begin{aligned} E(\hat{n}) &= \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{m=\ell} E_{\ell m} Y_{\ell m}(\hat{n})\,, \\ B(\hat{n}) &= \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{m=\ell} B_{\ell m} Y_{\ell m}(\hat{n})\,. \end{aligned} \tag{3.24}$$

This is the E and B decomposition of a set of Q and U maps, for which the same angular power spectrum described in eq. (3.19) is defined. The same geometrical considerations discussed in section 3.3.2 also apply here, being the same decomposition.

The angular power spectrum thus describe the variance of the spherical harmonic coefficients. If the $a_{\ell m}$ are drawn from a Gaussian distribution (as predicted by inflation), then all the statistical information is encoded in the power spectrum, making it one of the most powerful tools in Cosmology. This the main quantity measured by CMB experiments and compared with theoretical models. Following the definition given in section 1.3.1, the angular power spectrum is simply the Fourier transform of the 2-point angular correlation function ξ; in this case, we are looking at two different directions in the sky, and evaluating the anisotropy power on those separated angular scales, parametrised by the multipole moment ℓ. The angular power spectrum between two different fields X and Y can be computed as

$$C_\ell^{XY} = \left\langle a_{\ell m}^X \left(a_{\ell m}^Y \right)^* \right\rangle\,, \tag{3.25}$$

with the cosmic variance being

$$\sigma^2_{\text{cosmic}}(XY) = \frac{1}{2\ell+1}\left[C_\ell^{XX}C_\ell^{YY} + \left(C_\ell^{XY}\right)^2\right] . \tag{3.26}$$

If the two fields are different, we talk about cross-spectra and they describe the cross-correlation of field X with field Y. If the two fields are the same (e.g., $X = Y \rightarrow C_\ell^{XX}$) we talk about auto-spectra and they provide information on the auto-correlation of the field with itself, with the cosmic variance reducing to eq. (3.20). C_ℓ^{XY} and C_ℓ^{YX} contain the same statistical information. In CMB data analysis we can compute auto- and cross-correlations of three different fields which are temperature T, and polarisation E- and B-modes, resulting in six independent power spectra: C_ℓ^{TT}, C_ℓ^{TE}, C_ℓ^{TB}, C_ℓ^{EE}, C_ℓ^{EB}, C_ℓ^{BB}, usually shortened to TT, TE, TB, EE, EB, BB. Cross-correlations with the parity-odd B-modes (TB, EB) are expected to be zero, since temperature and E-modes are parity-even fields. In the standard ΛCDM model there is no mechanism to generate cross-correlation with B-modes; a significant detection of non-zero correlation would imply a parity violation and/or rotation, and consequently the existence of new physics needed to explain it. In general, this is attributed to effects like the presence of primordial magnetic fields [160, 153] or axion-like particles inducing cosmic birefringence [134, 37, 106].

Since we only have one sky to look at, the ensemble average $\langle\cdot\rangle$ is not accessible to us, as we have only one observable set of $a_{\ell m}$ coefficients. When looking at real data, we replace eq. (3.25) with an angular power spectrum estimator built by averaging over all the $(2\ell+1)$ possible orientations m at a given multipole ℓ:

$$C_\ell^{XY} = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} a_{\ell m}^X \left(a_{\ell m}^Y\right)^* . \tag{3.27}$$

It is usual to report the power spectra in a different normalisation, as

$$D_\ell^{XY} = \frac{\ell(\ell+1)}{2\pi}C_\ell^{XY} , \tag{3.28}$$

which is approximately proportional to the total power per logarithmic interval in ℓ (analogous to the $k^3P(k)$ dimensionless quantity for the 3D Fourier power spectrum). In a flat cosmological model with an $n_{\rm s} \simeq 1$ power spectrum, the Sachs-Wolfe contribution to the power spectrum is proportional to $1/\ell(\ell+1)$, dominant on large scales. In the latter normalisation, the angular power spectrum is thus flat at low ℓ.

In practical life, the estimation of the CMB angular power spectrum is not as straightforward as it may seem. Many problems arise when working with real data,

and a particular attention is required when extracting cosmological information. Here we report some of the main challenges commonly faced in CMB data analysis:

- **Partial sky.** In general, it is not possible to access the full microwave sky. Either because we are only observing a portion of the sky from the ground, or because the full sky maps contain contamination due to galactic and extra-galactic foregrounds which needs to be masked out. Estimating angular power spectra on a partial sky has several consequences. The first one is a loss of power due to lower number of pixels observed. At first order, this power can be recovered by diving the estimator by the fraction of pixels included in the analysis $f_{\rm sky}$. But at higher order, the partial sky implies the non-orthogonality of the spherical harmonics basis, since they are not defined any more on the full sphere. This induce an effect known as mode-coupling, where different multipole moments are not independent any more, but coupled. This ultimately bias the power spectrum estimator [103, 12]. Moreover, when treating polarisation we have an additional problem due to the E- and B-modes decomposition being inherently non-local, and non-unique in the presence of boundaries [122, 36]. Since the E-modes are generally much more dominant with respect to the fainter B-modes, the net effect of this problem is a "leakage" of E into B power, able to bias the recovery of the B-mode power spectrum.

  Both of these effects (mode-coupling and E-to-B leakage), are induced by observing a partial sky, and they are usually treated with mathematical tools able to correct most of the bias. In this Thesis we will mostly use the `NaMaster` package[6] [12] to account for these problems.

- **Noise bias.** When we compute the auto-angular power spectrum of a map, we are evaluating the correlation of the underlying fields with themselves. This is very useful to highlight the statistical properties of the CMB and the correlation power at different angular scales. Unfortunately, these maps also contain a realisation of instrumental noise, which in the simplest scenario is random, white, Gaussian noise, but is in reality more complicated. The correlation of the noise field with itself will show up in the angular power spectrum as a bias, more pronounced at smaller scales where noise is expected to dominate. The simpler way to avoid the noise correlation problem is to replace the auto-spectra with the cross-spectra, correlating two maps possibly containing independent noise realisations. This is usually achieved by employing data "splits" instead of full-mission maps, where the splits have been built using a

[6] https://namaster.readthedocs.io/en/latest/

subset of the total data. For example, two common splits used in the analysis of *Planck* data are built by either dividing the mission in two half-mission chunks, or by building maps as observed by different, separate detector sets. In this way, the final frequency maps will contain different noise realisations and systematics, mitigating the impact of the noise bias.

Once the above effects have been taken into account, it is possible to proceed with the extraction of cosmological information. Each of the six angular power spectra possesses a rich structure which is sensitive to the physics of the primordial and late Universe, with the value of cosmological parameters influencing the shape and amplitude of the curves. Figure 3.7 shows a compilation of recent measurement of the CMB angular power spectra, by different microwave experiments, both space- and ground-based, and their combinations. We will now use this figure to briefly describe the main features of the different power spectra, ignoring the lower panel showing the lensing deflection power spectrum $C_\ell^{\phi\phi}$.

$\boldsymbol{TT}$ The auto-correlation of the temperature field is the spectrum with the largest amplitude. The $\ell = 0$ and $\ell = 1$ components correspond to the monopole and dipole already discussed. From $\ell = 2$ onwards, the spectrum shows the oscillatory features characteristic of the physics of the photo-baryonic plasma, namely the acoustic oscillations. At $\ell < 100$, we have large scales that were outside the causal horizon at the time of last scattering, which have never interacted with the rest of the Universe since the end of inflation. Thus, we are looking at the primordial perturbations as generated by inflation, which are close to scale-invariance ($n_\mathrm{s} \simeq 1$). This translates to the Sachs-Wolfe plateau linked to the density perturbations. Deviations from this plateau are due to the integrated Sachs-Wolfe secondary anisotropies. In the $100 < \ell < 1000$ interval we have the acoustic peaks, which encode the oscillations occurring in the primordial plasma at scales within the causal horizon. The relative amount of baryonic matter Ω_b, cold dark matter Ω_CDM, and dark energy Ω_DE establish amplitudes and positions of these peaks, together with the curvature of the Universe Ω_κ and shape of primordial perturbations A_s and n_s. The maximum peak is at $\ell \simeq 200$. At $\ell > 1000$, structures are so small that their perturbations are efficiently smoothed out by photon diffusion damping (or Silk damping) [193]: on such scales, photons are able to travel from hotter to colder regions, dragging electrons and baryons; inhomogeneities are erased. This result in a damping envelope (or tail) which smooth the amplitude of the peaks.

$\boldsymbol{TE}$ **and** $\boldsymbol{EE}$ The E-modes auto-correlation is two orders of magnitude fainter than the temperature signal, being the CMB polarised at the $\sim 10\%$ level. The

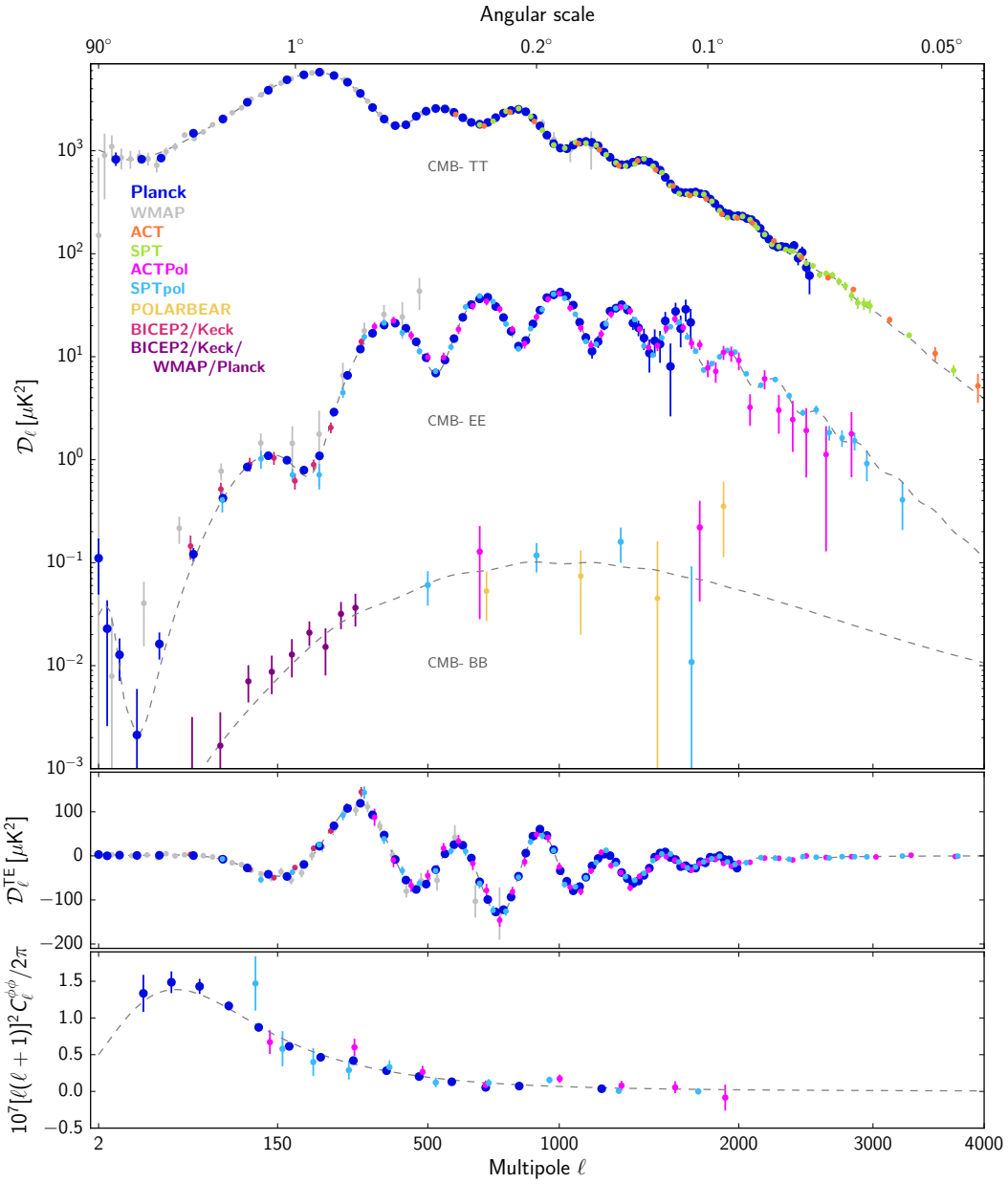


Figure 3.7. Compilation of recent CMB angular power spectrum measurements. *Upper:* temperature, E- and B-mode polarisation. *Middle:* TE cross-correlation. *Lower:* lensing deflection power spectrum. Different colours correspond to different experiments. The dashed line indicates the best-fit ΛCDM model to the *Planck* temperature, polarisation, and lensing data. Figure from [168].

wiggly structure is very similar, with the main difference being the phase shift: the EE peaks are out of phase with respect to TT ones. This stem from the fact that the polarisation anisotropies are sourced by the Doppler effect only, and not by the under- and over-densities (like the Sachs-Wolfe effect for temperature). The excess at $\ell < 10$ is known as reionisation bump, and it is connected to the additional scattering layer provided by the reionisation process, adding extra power at the largest scales.

The TE spectrum quantifies the correlation between the temperature and E-modes fields. Positive values indicate correlation, while negative values anti-correlation. The crucial thing to notice here is the negative dip in the range $30 < \ell < 200$. This anti-correlation between super-horizon scales is one of the most important indirect proofs of the inflationary paradigm, being a consequence of initial conditions setup outside the horizon. It is possible to build other theories explaining the peaks *inside* the horizon, but inflation is the only scenario able to also properly address perturbations *outside* the horizon.

$\boldsymbol{BB}$ At the time of writing of this Thesis (August 2026), the only measured CMB-related contribution to the BB spectrum is from gravitational lensing, which is a secondary anisotropies arising from the gravitational deflection of CMB photons from intervening structures in the late Universe. The E-modes pattern is distorted by this effect and some of it will be converted to "spurious" B-modes, even if the real scalar B-modes are zero, being more pronounced at intermediate and small scales. This feature has been measured by several experiment [174, 21, 192], and represents one of the main contamination to the primordial signal we are mostly interested in, induced by inflationary gravitational waves. As we said in the previous section, tensor perturbations have not been detected yet, and their measurement is one of the main goal of current and future experiments.

3.5 Observing the CMB

Years before the official discovery of the CMB, things were already in motion to hunt for this primordial signal [216]. Starting from theoretical arguments, people had a more or less precise idea on where and how to look for the faint radiation, with the CMB power spectra being drawn on paper years before its actual measurement. We report here a brief, non-exhaustive list in chronological order of past, current, and future experiments targeting the characterisation of the CMB. In the list, space-based and sub-orbital experiments are reported. We defer the discussion about the difference between these operational methods to the next chapter.

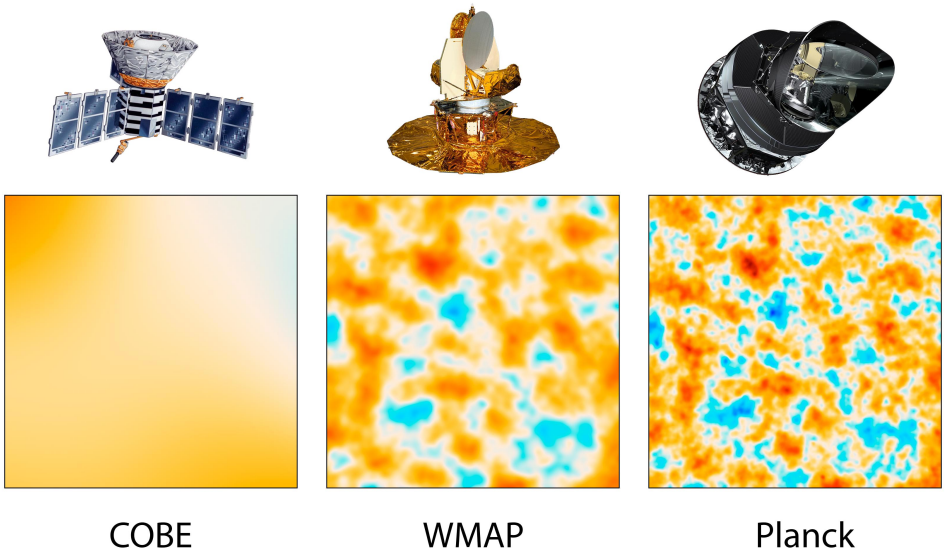


Figure 3.8. Comparison between the CMB anisotropies observed by the three satellites *COBE* (left), *WMAP* (middle), and *Planck* (right), showing the amazing improvement in sensitivity and resolution. The three panels show 10-square-degree patches of all-sky microwave maps. Source: link. Courtesy NASA/JPL-Caltech.

3.5.1 Past experiments

- 1964 – Penzias and Wilson report a measurement of the sky temperature [155], interpreted as the monopole of the CMB radiation [60].
- 1990-1992 – The *FIRAS* instrument onboard the *Cosmic Background Explorer*[7] (*COBE*) satellite measures the blackbody shape of the CMB spectrum with exquisite precision (figure 3.1), and shows that the CMB has a nearly perfect blackbody spectrum with $T \simeq 2.73\,\mathrm{K}$ [140, 219, 78, 77]. The *DMR* instrument onboard the same spacecraft discovers the CMB anisotropies.
- 1999 – Among others, the balloon-borne *BOOMERanG*[8] experiment makes the first measurements of acoustic oscillations in the CMB anisotropy angular power spectrum, also providing evidences for a flat geometry of the Universe [56, 54, 131].
- 2002 – CMB polarisation is discovered by the *DASI*[9] telescope [113].
- 2003 – The *Wilkinson Microwave Anisotropy Probe*[10] (*WMAP*) spacecraft, successor of the *COBE* mission, produces its first maps of the sky and

[7] https://lambda.gsfc.nasa.gov/product/cobe.
[8] https://en.wikipedia.org/wiki/BOOMERanG_experiment.
[9] https://en.wikipedia.org/wiki/Degree_Angular_Scale_Interferometer.
[10] https://science.nasa.gov/mission/wmap/wmap-overview.

measures the first peaks of the CMB temperature angular power spectrum, constraining cosmological parameters.

- 2009 – The *Planck*[11] space telescope is launched, and Cosmology enters its precision era. Exquisite measurements of temperature and polarisation anisotropies are reported, with the TT and EE power spectra characterised until small scales.

- 2007-2022 – The ground-based *Atacama Cosmology Telescope*[12] (*ACT*) provided sharp measurement of (seven) acoustic peaks in the CMB power spectrum, both in temperature and polarisation, improving satellites results on the smallest angular scales. It also provided the first measurement of gravitational lensing in a CMB map only, and evidences for dark energy from CMB data alone [192, 150, 130].

3.5.2 Current / active experiments

- from 2006 – The *Background Imaging of Cosmic Extragalactic Polarization* (*BICEP*)[13] and the *Keck Array* [24, 7] are a series of successive ground-based experiments targeting the measurement of primordial B-modes by looking at a relatively small, clean sky region from Antarctica. The majority of the upper limits derived on the tensor-to-scalar ratio r come from the *BICEP*/*Keck* data, for it currently being the experiment with the best sensitivity for the search of primordial B-modes. In 2014, the collaboration claimed a detection of primordial gravitational waves from inflation, reporting $r = 0.20^{+0.07}_{-0.05}$ [4]. After a joint analysis with *Planck* data [27], the claim was later revealed to be a false detection induced by the polarised emission of Galactic dust.

- from 2007 – The ground-based *South Pole Telescope*[14] (*SPT*) reports the first measurement of the B-mode component in CMB polarisation at small angular scales, sourced by gravitational lensing [21].

- from 2016 – The *Cosmology Large Angular Scale Surveyor*[15] (*CLASS*) [73] ground-based experiment targets the polarised CMB emission at large scales to extract information on reionisation and the birth of the first stars, while also searching for inflationary primordial B-modes. It also tries to constrain the circular polarisation of the CMB.

[11] https://www.esa.int/Science_Exploration/Space_Science/Planck/Planck_overview.
[12] https://act.princeton.edu.
[13] https://www.cfa.harvard.edu/facilities-technology/telescopes-instruments/bicep.
[14] https://pole.uchicago.edu/public/Home.html
[15] https://sites.krieger.jhu.edu/class/

- from 2024 – The *Simons Observatory* (*SO*) [82, 6] started operations in 2024 and it is currently collecting its initial data of the sky. The main scientific goal of the experiment is to constrain the amplitude of primordial gravitational waves by measuring r, together with gravitational lensing and Sunyaev-Zel'dovich effect (thermal and kinematic).

3.5.3 Future experiments

- 2030s – *LiteBIRD*[16] (Lite satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection) [127, 128] is a planned space mission selected by the Japanese Space Agency, to be the successor of the *Planck* observatory. Its main scientific target is the measurement of the B-mode signal due to cosmic inflation, with the current forecast predicting an experiment able to map the CMB polarised emission over the full sky with a sensitivity of $\sigma(r) \lesssim 0.002$.
- 2030s-2040s – *FOSSIL* [51] is a proposal for a space mission targeting the spectroscopy of the CMB, being the first experiment dedicated to the measurement of CMB spectral distortions after the *FIRAS* instrument.
- 2030s-2040s – *PICO* (Probe of Inflation and Cosmic Origins) [95] is a mission concept for a future space experiment with the aim of observing the full sky in intensity and polarisation over a broad range of frequencies. Scientific goals include the characterisation of inflationary physics, reionisation, structure formation.

Figure 3.8 shows a comparison between the CMB anisotropies observed by the three successive space missions, highlighting the improvement in instrumental performance between the different experiments.

[16] https://www.isas.jaxa.jp/en/missions/spacecraft/future/litebird.html.

Chapter 4

Foregrounds as a contaminant to CMB observations

Until now we have described the CMB and its observation as it being the only radiation present at the relevant experimental frequencies. In reality, the CMB is only one part of the multi-facet emission of the microwave sky. Being the CMB a radiation coming from ~ 13 Gyrs ago, we have the entire Universe which is between us and the last scattering surface. The primordial signal is a background, and everything else not coming from the instrumentation is foreground emission (or shortly, foregrounds), generated by various physical mechanisms taking place inside and outside our Galaxy. For cosmological purposes, this foreground emission is generally considered as a contamination of CMB observation, as it obstructs our view of the primordial photons and prevents us from taking a clear picture. Instead, from the astrophysical point of view these foregrounds are a trove of information; their emission provides invaluable insights about the complex interplay between small scales components of the sky. A better understanding of the physics behind the contamination is crucial for its mitigation.

4.1 Inventory of the contaminants

As it is evident from the *Planck* maps of the microwave sky reported in figure 3.2 and 3.4, the main contamination is in the form of a horizontal strip, representing the plane of our Galaxy. Being the Earth located in one of the spiral arms of the Milky Way, the thick Galactic disk will be present in any full-sky observation. But we also have a lot of bright emission in regions of the maps far from the plane, extending anisotropically in different directions. A notable feature is the visual dependence of the contamination on the frequency of observation, with some frequencies less

contaminated than others, and the morphology of the emission being different from lower to higher frequencies. Additionally to the Galactic plane, there are both localised and diffuse emissions visible and detectable on the whole sphere; here we briefly describe the most important contributions, referring to [184, 64] and *Planck* results [159, 169] for many more information:

- **Synchrotron radiation.** It is generated by relativistic electrons accelerated by the Galactic magnetic field in the ionised interstellar medium. Detailed in section 4.2.1.
- **Thermal dust emission.** Produced by dust grains in our Galaxy which thermally re-emit the absorbed starlight. More details in section 4.2.2.
- **Free-free emission.** Similar to synchrotron radiation, it is emitted by electrons deviated by positive ions in the interstellar plasma. Detailed in [61].
- **Anomalous Microwave Emission (AME).** Fairly unknown radiation contributing at frequencies around ~ 30 GHz, still under studies. Usually attributed to the emission of small, rapidly spinning dust grains [65, 47]. Its polarised emission has been constrained to be less than 1%, being thus negligible in observations of CMB polarisation.
- **Spectral lines.** Atoms and molecules produce emission lines in the microwave frequency range, with extragalactic CO molecular rotational line being the most prominent for CMB science.
- **Zodiacal light.** Generated by dust grains in the Solar system, reflecting and re-emitting Solar radiation. Visible as a diffuse and weakly polarised signal in the ecliptic plane [84].
- **Cosmic Infrared Background (CIB).** Diffuse background radiation due to the combined infrared emission of extra-Galactic sources, mainly coming from unresolved dusty galaxies.
- **Point sources.** Localised bright sources, which can be either of extra-Galactic origin like quasars or radio galaxies, or Galactic like compact object (pulsars and supernovae remnants).

All these different foregrounds contribute to the total brightness of the millimetre sky, which is illustrated in figure 4.1 as a function of the observing frequency, both in temperature and polarisation. This function is usually called spectral energy distribution (SED). Sky masks are also applied in order to reduce the contamination from the Galactic plane, retaining a certain sky fraction $f_{\rm sky}$. The impact of the sky

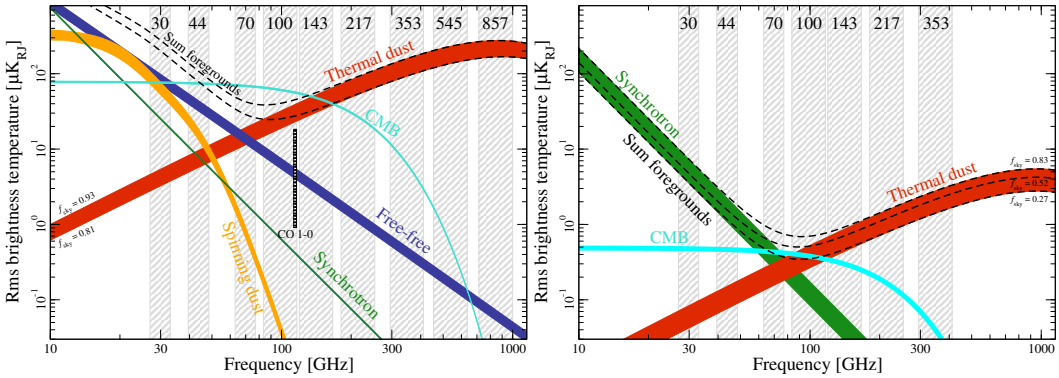


Figure 4.1. Frequency dependence of the main components of the sub-millimetre sky in temperature (*left*) and polarisation (*right*, shown in P). The vertical grey bands indicate the Planck channels, while the coloured bands show the major signal and foreground components. For temperature, the widths of the bands show the range for sky masks with 81-93% sky coverage. For polarisation, the range is 73-93%. Figure from [168].

mask can be seen from the thickness of the coloured bands. In temperature observations, the CMB signal is stronger than the total foregrounds emission minimum at frequencies around ~ 90 GHz, making it one of the best windows to look at CMB (*Planck* channels are shown as vertical bands). At lower and higher frequencies, the contamination becomes dominant over the primordial signal. Similar considerations apply to polarisation, with synchrotron and thermal dust being the only important contaminants. The CMB signal is now lower than the cumulative foregrounds for every frequency, even at the minimum of their emission. This Thesis will focus on observing the polarised signal of the CMB, and for this reason I will restrict the discussion to those foregrounds contributing the most in polarisation.

Therefore, the extraction of primordial information from the microwave sky requires a careful treatment of the contamination, which represents one of the main challenges in CMB science. Methodologies tackling the problem of retrieving the cleanest possible CMB signal are known as "component separation" methods, and they will be described in more details in section 4.3.

4.2 Polarised foregrounds

In comparison with temperature observations of the sub-millimetre sky, polarisation measurements provide good news and bad news (see figure 4.1). The good news is that the majority of the foregrounds described in the previous section do not emit significantly polarised light: synchrotron radiation, thermal dust emission, and point sources are the only contaminants we need to worry about. The bad news is that the foreground emission is now dominant with respect to the CMB signal,

over the entire frequency spectrum. In the context of observations targeting the measurement of the primordial B-mode signal, polarised foregrounds represent the major challenge to overcome. Although the level of contamination depends on the sky region, polarised foregrounds are known to have amplitudes corresponding roughly to $r = 0.01 - 0.1$ [114] at their minimum frequency (~ 80 GHz). As a function of scale, the contamination can be up to two orders of magnitude higher than the CMB B-mode power spectrum, which itself is dominated by non-primordial B-modes induced by gravitational lensing. We now describe these contaminants in more detail.

4.2.1 Synchrotron

Generally speaking, synchrotron radiation is the electromagnetic emission produced by relativistic charged particles subject to an acceleration perpendicular to their velocity. In the presence of a magnetic field $\mathcal{B}$, a free charge q will spiral around the direction of the field following from the Lorentz force

$$\frac{\mathrm{d}(\gamma m \mathbf{v})}{\mathrm{d}t} = \frac{q}{c}(\mathcal{E} + \mathbf{v} \times \mathcal{B}) \,, \tag{4.1}$$

where γ is the Lorentz factor. As illustrated in the left panel of figure 4.2, the electron will emit a focalised beam of polarised light tangential to its trajectory, where the power carried by the radiation is given by the relativistic Larmor formula:

$$\mathcal{I} = \frac{q^2 \gamma^4}{6\pi\epsilon_0 c^3} a^2 \,, \tag{4.2}$$

where ϵ_0 is the vacuum permittivity and a is the magnitude of the acceleration. Additionally, this radiation is linearly polarised. Therefore, if we consider as the free charge an electron in the interstellar medium of the Milky Way, synchrotron can be seen as a tracer of the Galactic magnetic field.

In the context of microwave sky observations, synchrotron is the dominant polarised emission at frequencies $\lesssim 70$ GHz. It is strongly linearly polarised, up to $\sim 20\%$ at intermediate and high Galactic latitudes [112]. In the frequency range $2 < \nu < 100$ GHz, the SED of synchrotron emission is usually modelled as a power-law with amplitude A_s and spectral index β_s:

$$S_\mathrm{s}(\nu, \hat{n}) = A_\mathrm{s}(\hat{n}) \left(\frac{\nu}{\nu_\mathrm{s}^*}\right)^{\beta_\mathrm{s}(\hat{n})} \,, \tag{4.3}$$

where $\hat{n}$ is the direction of observation, and ν_s^* is a reference frequency used for normalisation, usually taken to be $\nu_\mathrm{s}^* = 23$ GHz in models built upon observations

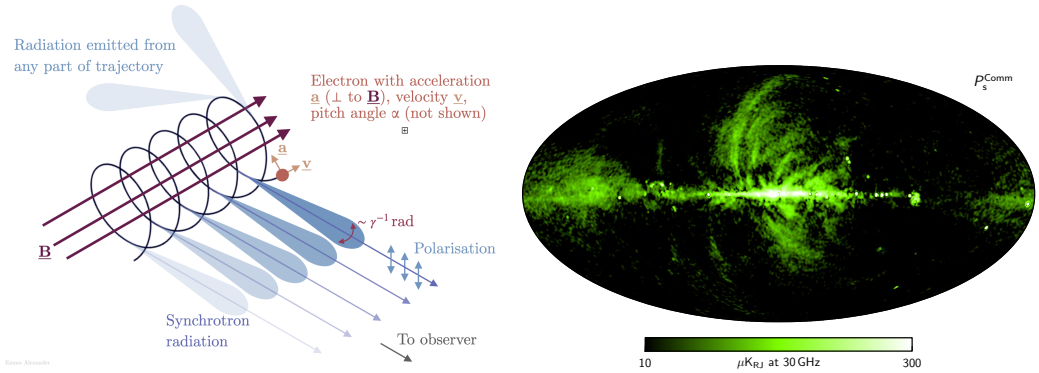


Figure 4.2. *Left:* diagram showing the mechanism of synchrotron radiation emitted by an electron moving along a helical path around an ordered magnetic field. The Lorentz force is perpendicular to the magnetic field. Picture taken from [9]. *Right:* polarised synchrotron amplitude map by *Planck*, obtained with the `Commander` algorithm, at 40' resolution, and evaluated at the reference frequency of 30 GHz. From [169].

made by the lowest frequency channel of *WMAP* [23]. The amplitude and spectral index are in principle dependent on the direction we are looking at, identified by a specific pixel in a tessellated map of the sky. The mean spectral index extracted from *Planck* and *WMAP* data is $\beta_{\rm s} \simeq -3$ [80, 115, 138].

Statistically, on the largest scales the angular power spectra of the polarised synchrotron emission is approximated with a power-law:

$$C_\ell^{XX,{\rm s}}(\nu) = A_{\rm s}^{XX} \left(\frac{\ell}{\ell_0}\right)^{\alpha_{\rm s}^{XX}} \tag{4.4}$$

where $XX = EE, BB$ and ℓ_0 is a pivot multipole used for normalisation, typically chosen to be close to the expected scale of recombination $\ell_0 = 80$. The power-law index is measured to be $\alpha_{\rm s} \simeq -3$ for both EE and BB [115, 138]. The asymmetry of E- and B-modes is generally quantified with the B-to-E ratio, evaluated as the amplitude ratio at the pivot multipole ℓ_0; depending on the frequency and the considered sky region, it assumes values of ~ 0.5 (at 2.3 GHz) [115] or ~ 0.2 (at 23 and 30 GHz) [138]. The right panel of figure 4.2 shows a map of the synchrotron polarised amplitude obtained from *Planck* data, at the reference frequency of 30 GHz [169].

Synchrotron emission represent a major challenge in the search of primordial B-modes, contaminating the lower frequency bands and hindering the recovery of the underlying CMB signal. Its characterisation is crucial for a future measurement of the tensor-to-scalar ratio [114, 115, 138, 67].

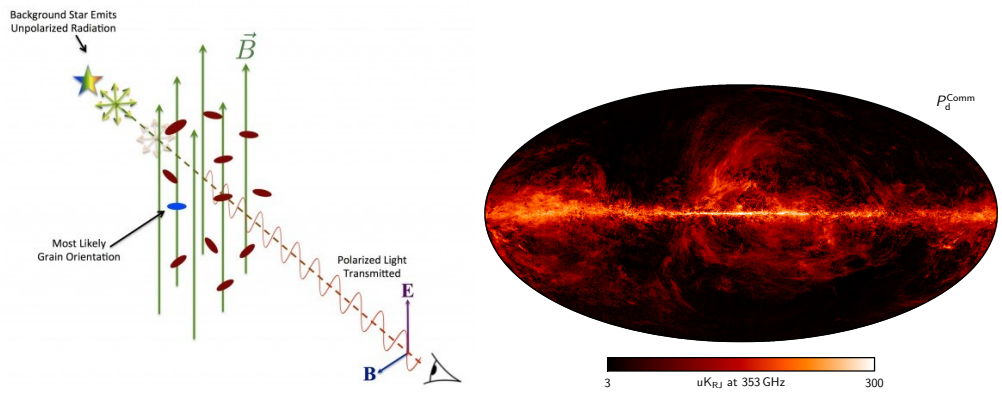


Figure 4.3. *Left:* diagram showing how dust clouds are able to polarise background starlight. Asymmetric and elongated grains tend to align with the direction of the magnetic field, emitting polarised radiation along the big axis of the grain. Picture from the *Pasiphae* website. *Right:* polarised thermal dust amplitude map by *Planck*, obtained with `Commander` algorithm, at 5' resolution, and evaluated at the reference frequency of 353 GHz. From [169].

4.2.2 Thermal dust

Interstellar dust is made of small grains of various dimensions, spanning from the size of a molecule to $\sim \mu\mathrm{m}$. The chemical composition include carbon, silicates, and graphites, with the temperature of the grains being in the 15-25 K range. With thermal emission we refer to the radiation produced by these grains: they heat up by absorbing high-energy starlight, and then re-emit it at lower energy, in the infrared and microwave regimes. Since these dust grains are not spherical, they tend to align perpendicularly to interstellar magnetic fields, inducing a certain degree of linear polarisation to the starlight that passes through dust clouds [64]. The left panel of figure 4.3 offers a sketch of this mechanism. For this reason, interstellar dust is a very good tracer of the Galactic magnetic field shape and intensity.

Concerning CMB observations, thermal dust emission is the dominant contribution to the sky polarised emission at frequencies $\gtrsim 100$ GHz, with a peak at 1 THz. The SED of thermal dust emission in the microwave is close to a modified blackbody (MBB), which is the product of a blackbody function and a power-law:

$$S_{\mathrm{d}}(\nu,\hat{n}) = A_{\mathrm{d}}(\hat{n}) \left(\frac{\nu}{\nu_{\mathrm{d}}^*}\right)^{\beta_{\mathrm{d}}(\hat{n})} \frac{B(\nu,T_{\mathrm{d}})}{B(\nu_{\mathrm{d}}^*,T_{\mathrm{d}})}\,, \tag{4.5}$$

where A_{d} and β_{d} are the amplitude and spectral index, $B(\nu,T_{\mathrm{d}})$ is the Planck's law, and ν_{d}^* is a reference frequency used for normalisation, generally taken to be the highest *Planck* polarised channel at 353 GHz. T_{d} is the temperature of the dust grains, which is taken to be ~ 20 K for the simplest scenarios assuming a single

population of grains. The modified blackbody emission is characteristic of a dust component which is almost in thermal equilibrium with the background radiation field, while the power-law contribution is induced by the shape and composition of the grains [158, 164]. *Planck* data have been used to determine the spectral index, finding $\beta_{\rm d} \simeq 1.59$ [158].

Similarly to synchrotron, the large-scales angular power spectra of thermal dust emission follow a power-law:

$$C_\ell^{XX,{\rm d}}(\nu) = A_{\rm d}^{XX} \left(\frac{\ell}{\ell_0}\right)^{\alpha_{\rm d}^{XX}} \tag{4.6}$$

where, again, $XX = EE, BB$ and $\ell_0 = 80$. The power-law index is measured to be $\alpha_{\rm s} \simeq -2.4$ for both EE and BB, while the B-to-E ratio is around 0.5 [164]. The right panel of figure 4.3 shows a map of the dust polarised amplitude obtained from *Planck* data, at the reference frequency of 353 GHz [169].

Thermal dust emission has already proven itself to be a formidable challenge for the cosmological measurement in polarisation, with early studies already advocating the need for a careful mitigation [17]. In particular, the underestimation of dust contamination led to the false detection of primordial B-modes, reported in 2014 by the *BICEP* collaboration [4]. The dust-induced B-modes excess signal was interpreted to be the signature of primordial gravitational waves generated by inflation. A successive re-analysis, joint with *Planck* data [27] confirmed the astrophysical nature of the signal, and further studies showed that contamination from Galactic dust can not be neglected even in the cleanest regions of the sky [114]. There is no possibility to extract information on primordial B-modes without a careful, proper characterisation and treatment of Galactic foregrounds.

4.2.3 Point sources

In general, extra-galactic compact objects are sufficiently distant to be unresolved by instruments targeting the mapping of the CMB, as far as their brightness is sufficiently small. Their contribution to the sky emission is a combination of two signals: the brightest sources, which can be resolved; the faintest sources (and the most numerous) which are not individually significant but form a background with fluctuations contributing to the total signal. On the low frequencies range, radio sources, such as active galactic nuclei (radio galaxies, quasars, blazars), follow a large variety of emission laws. Extrapolation of their flux to CMB wavelengths is therefore difficult. *WMAP* [220] and *Planck* [162] experiments have produced catalogues of the brightest radio sources in their observational frequencies, on the totality of the sky.

Point sources have been shown to have a fractional polarisation of a few percent [129, 20, 139]. They contribute to the polarisation angular power spectra as a Poissonian emission, meaning that they present a flat power spectrum (constant at every scale) gathering contributions from the unresolved background of sources [67]. Their impact is larger in the radio observational window, which correspond to the lower frequency bands of the usual experiments targeting CMB science, and at smaller scales where the synchrotron signal decreases. Moreover, the brightest of the polarised point sources can be responsible for ringing in the angular power spectra estimation, induced by the computational limitation of operating a spherical harmonic decomposition of a very bright and compact region [197]. The general treatment adopted for the mitigation of point sources is masking, where the compact object and their surrounding are masked at the pixel level. Other useful techniques include the inpainting of the point sources, replacing their emission with with a statistical representation of the underlying sky signal.

4.3 Component separation

In the previous sections we discussed how foreground emission is able to prevent us to have a clean view of the CMB signal, making the extraction of cosmological information much more difficult. It is crucial to develop ways to mitigate foregrounds contamination and disentangle the different contribution that are superimposed on the sky. Component separation methods are numerical tools employed to isolate and recover the CMB signal from the foregrounds, exploiting several physically justified hypothesis. Different component separation techniques are commonly used, often being complementary. They can be roughly divided as:

- **Parametric methods.** They explicitly model the frequency properties of the foregrounds through a set of parameters, which are then fitted to the data in order to extract a cleaned CMB estimate. They use prior information on the SEDs of the sky components, being thus able to provide an estimate of the CMB and foregrounds emissions. However, if the assumed underlying model of the sky is incorrect, it can lead to severe biases in the recovered maps, and consequently in the derived cosmological parameters. Examples of parametric methods include `Commander` [69, 70] and `FGbuster` [196].

- **Blind methods.** These techniques aim at recovering the CMB signal without strong assumption on the underlying model of foreground emission. In this way, the cleaning is not affected by our limited knowledge of the complex contamination, especially in polarisation. In general, blind methods linearly

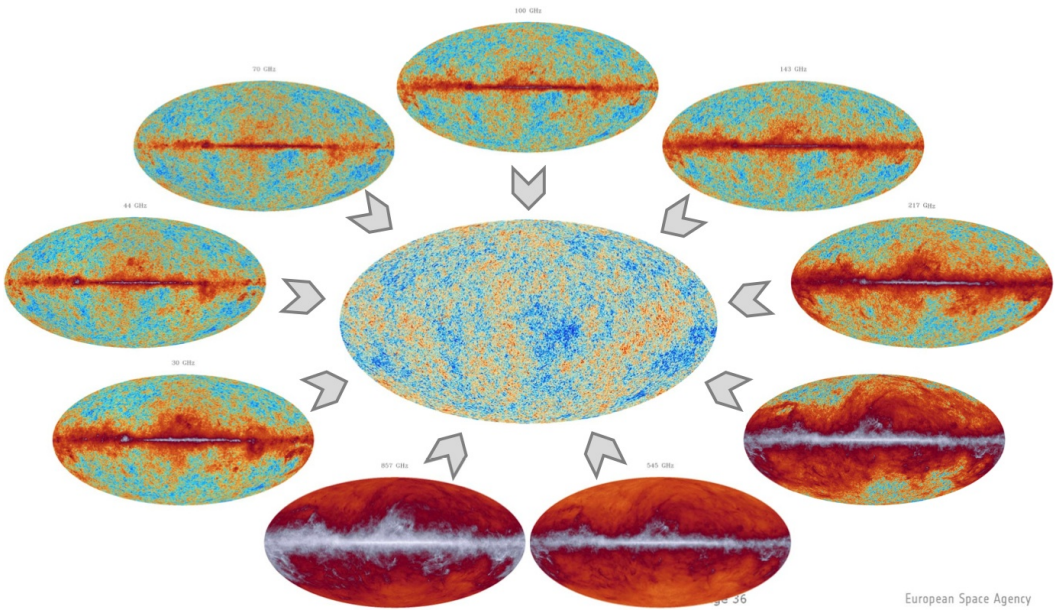


Figure 4.4. Diagram illustrating the component separation principle, combining *Planck* multi-frequency sky maps (in temperature) to obtain a foreground-cleaned CMB reconstruction. From this website.

combine the multi-frequency sky maps observed by an experiment in such a way as to minimise some meaningful statistical quantity of the final recovered map.

- **Template fitting.** Internal foregrounds templates either in pixel or harmonic space are obtained using multi-frequency observations, fitted to the data and removed. Usually, these foregrounds templates are built as difference maps between two neighbouring frequency channels, in order to remove the common CMB signal. `SEVEM` [136, 74] is an example of a template fitting algorithm which has also been applied on *WMAP* and *Planck* data [159].

For the sake of results robustness, many component separation methods are usually applied within the same experiment, to the same datasets. As an example, the *Planck* collaboration has adopted four different component separation algorithms over the years [159, 169], typically finding consistent results between methodologies. Figure shows and compares the recovered CMB maps (in Q and U) after the application of the `Commander` and `SEVEM` component separation methods on the latest *Planck* data release (PR4, or `NPIPE`) [173].

In this Thesis, we will adopt blind component separation techniques in order to disentangle the CMB signal from the foregrounds contamination and recover a clean view of the anisotropies. In particular, we exploit two implementations of the more general Internal Linear Combination (ILC) methodology: Needlet ILC (NILC)

and Multi-Clustering NILC (MC-NILC) [57, 40, 41]. Here we briefly describe the main ideas behind the ILC technique and the two extensions, assuming to work with a set of observed (real or simulated) multi-frequency sky maps, pixelated on a spherical domain.

4.3.1 ILC

The Internal Linear Combination (ILC) method is one of the most widely used approaches for the reduction of contaminants in CMB observations. Proposed in 1996 [201], it was adopted for the first time in the analysis of *WMAP* intensity data [22, 202]. The algorithm has to be applied to maps of a scalar quantity, such as temperature, E- or B-modes, which hereafter we will refer to generically as X. The ILC CMB solution is obtained by linearly combining the N_ν multi-frequency maps X_j of one or multiple CMB experiments with frequency-dependent weights w_j:

$$\hat{X}_{\rm ILC} = \sum_{j=1}^{N_\nu} w_j X_j \,, \tag{4.7}$$

where the weights are derived by minimising the variance of the final map $\hat{X}_{\rm ILC}$. Each frequency map X_j can be described as the sum of the CMB signal s, the emission of the foregrounds f_j and the instrumental noise n_j:

$$X_j = a_j s + f_j + n_j \,, \tag{4.8}$$

where a_j are the calibration coefficients. If observations are calibrated with respect to the CMB and reported in thermodynamic units (which is generally the case), then $a_j = 1 \quad \forall j$. In order to preserve the CMB signal in the output map, the weights must satisfy the constraint $\sum_j w_j = 1$. Therefore, the final ILC product will be a map containing the full CMB signal, plus some residual contamination from foregrounds and noise:

$$\hat{X}_{\rm ILC} = \sum_{j=1}^{N_\nu} w_j \cdot (s + f_j + n_j) = s + \sum_{j=1}^{N_\nu} w_j \cdot (f_j + n_j) \ . \tag{4.9}$$

If we are working with a simulated data sets, the estimation of these foregrounds and noise residuals is straightforward, as it is enough to combine the weights with pure foregrounds and pure noise maps. For real data instead, the residuals have to be estimated through Monte-Carlo simulations based on prior knowledge of the contamination in the input multi-frequency maps.

The weights that minimise the variance of the recovered CMB map can be

analytically estimated as [22]:

$$\mathbf{w} = \frac{\mathbf{A}^T\mathbf{C}^{-1}}{\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}} \quad \rightarrow \quad w_i = \frac{\sum_j C_{ij}^{-1}}{\sum_{ij} C_{ij}^{-1}}\,, \tag{4.10}$$

where $\mathbf{A}$ is the vector of calibration coefficients, $\mathbf{C}$ is the empirical covariance matrix of the input multi-frequency maps (with dimensions $N_\nu \times N_\nu$), and the relation implied by the arrow holds if observations are well calibrated. If the CMB signal, foregrounds, and noise are statistically independent, we have that:

$$\sigma^2(\hat{X}_{\rm ILC}) = \sigma^2(s) + \sigma^2\left[\sum_{j=1}^{N_\nu} w_j \cdot (f_j + n_j)\right]\,, \tag{4.11}$$

which means that the variance of the discrepancy between the recovered map and the CMB signal is minimum when the variance of the total map itself is minimum. This is the reason why these techniques are also called "minimum variance" methods.

ILC does not need to assume anything about the foreground components, since all the needed information is already contained in the observations. The only (fair) assumptions are that the CMB has a known emission law (blackbody spectrum), and that it has no correlations with foregrounds or noise. Of course, there are some negative sides: as discussed in [59], any error in the CMB calibration could have a relevant impact on the recovery of the signal; moreover, the empirical covariance matrix $\mathbf{C}$ is estimated over a sample with finite size, and this can lead to appearance of correlations between the CMB signal and residual contaminants, in the component separation output [57]. This causes a loss of CMB power, more pronounced at large scales (low multipoles) where a smaller number of modes is available[1]. It is commonly known as the ILC reconstruction bias.

4.3.2 NILC

A major drawback of the ILC method stems from the huge inhomogeneity of the foregrounds emission. The contamination at high Galactic latitudes is much different from the one near the Galactic plane, with strongly varying properties and spectral indices. A single linear combination is probably not enough to characterise foregrounds over the whole sphere. The Needlet ILC (NILC) [57] is an ILC extension which performs a linear combination of the multi-frequency maps, after filtering them with needlets, which are a particular wavelet system defined on the sphere. Needlets have been introduced by [151] and applied to CMB science by [133]; they

[1] This effect is similar to the cosmic variance one discussed in section 3.4.

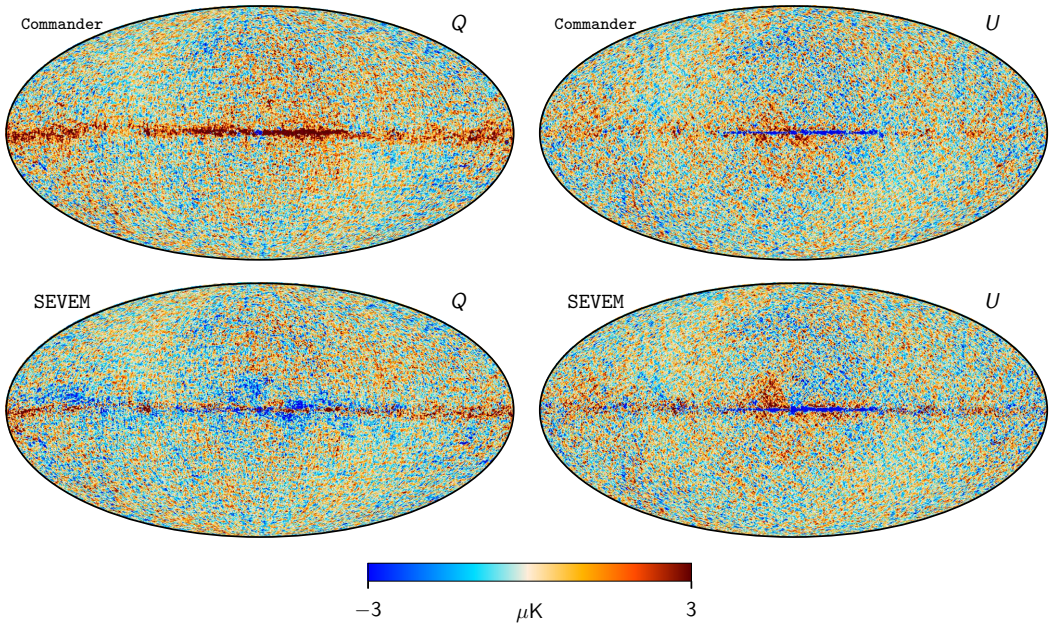


Figure 4.5. Foreground-cleaned CMB Q and U maps derived from the *Planck* `NPIPE` data set with `Commander` (top row) and `SEVEM` (bottom row), plotted with an angular resolution of 40'. From [173].

have the advantages of being simple, both in the forward and inverse transforms, and they are localised in both pixel and harmonic spaces, with the window function going to zero very quickly. It is possible to tweak the parameters of the function defining the needlets to obtain different configurations, which can be chosen depending on the nature of the problem. Standard, Mexican, and cosine needlets are the most used for component separation algorithms [85].

The NILC method consists in convolving frequency maps with a needlet function, which is equivalent to filter the same maps in harmonic space. As we saw in section 3.4, a field on the sphere can be expanded in spherical harmonics $Y_{\ell m}(\hat{n})$ with coefficients $a^j_{\ell m}$. We can filter this coefficients with a weighting function $b_i(\ell)$ for each scale i, in order to get a map β^i_j where only a specific range of scale is selected:

$$\beta^i_j(\hat{n}) = \sum_{\ell m} \left[a^j_{\ell m} \cdot b_i(\ell) \right] Y_{\ell m}(\hat{n}) \,. \tag{4.12}$$

Therefore, we have maps at several frequencies j, filtered at different scales i. In the same spirit of ILC, these maps at different frequencies are linearly combined to get a single, minimum-variance map for each scale, as

$$\beta^i(\hat{n}) = \sum_{j}^{N_\nu} w^j_i \cdot \beta^i_j \,. \tag{4.13}$$

The final map recovered by NILC is obtained by filtering again these maps in harmonic space to do the inverse transform, and summing all the resulting coefficients $a_{\ell m}^{\beta^i}$:

$$\hat{X}_{\rm NILC} = \sum_{\ell m} a_{\ell m}^{\beta^i} Y_{\ell m} \,. \tag{4.14}$$

Minimising the variance (and hence the contamination) separately on different needlet scales leads to a more effective cleaning. On large scales, diffuse Galactic foregrounds dominate over the other components and are better removed in NILC with respect to ILC. The same is expected to happen for the instrumental noise on smaller scales (larger multipoles).

4.3.3 MC-NILC

Even if the NILC method provides good performances on fields featuring a large signal-to-noise ratio (as in intensity or E-modes measurements), the leftover foreground contamination is still too high in the context of B-modes science. Future CMB missions will target the measurement of the very faint signal from primordial inflationary gravitational waves, and foregrounds mitigation is the main challenge to be faced.

Multi-Clustering NILC (MC-NILC) [41] is a recently developed extension of the NILC approach. The main idea behind it is to perform variance minimisation in needlet domain as NILC does, but doing so separately in different regions of the sky, dubbed "clusters" or "patches", which show similar spectral properties of foregrounds emission [176], in the B-modes component. To this goal, a blind tracer of the Galactic foregrounds spectral properties needs to be identified. As detailed in [41], one way to do this is to take the ratio of B-modes maps at different frequencies, in such a way that dust and synchrotron spectral dependence is encompassed through the whole frequency coverage of the considered experiment. The tracer is built on the same field on which the component separation is applied, in order to correctly trace the foregrounds spatial variations; Q and U maps show different spectral properties with respect to the B component. The reference frequencies adopted for the construction of the ratio can be calibrated with realistic simulations of the considered experiment. For example, the best choice for a space telescope similar to the *LiteBIRD* experiment is to take the ratio between a B-modes foreground template at a dust-dominated frequency, such as 402 or 337 GHz, and one at a "CMB" frequency channel, as 119 GHz. Then, pixels with similar values of this ratio are grouped together in separate sky patches, where the variance of the output solution is minimised independently for each needlet scale. Clusters should be small enough to enclose the least amount of foregrounds variability, while in the

meantime being large enough to contain enough modes and avoid a significant bias in the CMB reconstruction.

The MC-NILC algorithm provides improved performance on the reduction of foregrounds contamination when compared with NILC, especially at large scales (see [41] for a detailed study). This component separation technique is one of the most promising methodologies for the mitigation of foreground contamination in CMB observations, and it will be adopted for the data analysis of future space mission targeting the mapping of the microwave sky.

4.4 Simulating the contaminated sky

In order to test our knowledge of foreground emission and to develop and validate new methodologies for component separation, we need to be able to simulate some views of the sky. Nowadays in CMB science, this is usually achieved thanks to the Python Sky Model[2] (`PySM`) software [204, 224, 203], a `python` package which encapsulates almost all of our knowledge about foreground emission. With this software, the user can create sky maps containing different foreground components over the whole celestial sphere using the `HEALPix` pixelisation scheme [88, 223]. The components are approximated following our best knowledge of the foregrounds emission provided by the *WMAP* and *Planck* space missions, plus some information extracted by ground-based observatories. The intensity and polarisation models are built at a given resolution and frequency from a reference map, generally taken from *Planck* or *WMAP* data, and are extrapolated at any other frequencies using different models of SED. It is possible to generate simulations starting from different hypothesis, and injecting several levels of complexity in order to create optimistic or pessimistic scenarios approximating the real nature of the unknown Galactic signal. Moreover, new models are continuously added to the library, and they can be used to simulate and compare different microwave experiments. Component separation methods can be tested and benchmarked against different sky models.

In the `PySM` language, each model of foreground emission is identified by a label indicating the specific components and the level of complexity; for example, `a0b3` stands for a model including a component `a` with complexity 0 and a component `b` with complexity 3. Here we describe how some of these models are built, focusing on the dust (`d`) and synchrotron (`s`) components, as they are the predominant contamination in polarisation, and referring to the `PySM` papers [204, 224, 203] for more details.

[2] https://pysm3.readthedocs.io/en/latest/.

4.4.1 Dust

- `d1`: thermal dust is modelled as a single-component modified blackbody (MBB). Dust templates for the polarised emission come from the 353 GHz data of the *Planck* 2015 analysis. They are scaled to different frequencies adopting a MBB spectrum, with spatially varying temperature and spectral index obtained from the *Planck* data using the `Commander` code [159]. The polarisation templates are smoothed with a Gaussian kernel of 2.6 degrees, and small scales are added.

- `d0`: simplified version of the `d1` model, with a fixed spectral index of $\beta_{\rm d} = 1.54$ and a blackbody temperature of $T_{\rm d} = 20$ K.

- `d4`: a generalisation of the `d1` model to multiple dust populations [76], arising from the fact that a two component model is still a good fit to the *Planck* data.

- `d6`: this model implements dust frequency decorrelation, modelling the impact of averaging over spatially varying dust spectral indices, both unresolved and along the line of sight.

- `d10`: single-component MBB model based on templates from the needlet-based analysis of *Planck* data, with reduced contamination from CIB and point sources compared to `d1`. Small-scale fluctuations are added using the `logpoltens` (Logarithm of the Polarisation Fraction Tensor) formalism.

- `d9`: simplified version of `d10` with a fixed spectral index of $\beta_{\rm d} = 1.48$ and a fixed dust blackbody temperature of $T_{\rm d} = 19.6$ K all over the sky, based on *Planck* 2018 results [169].

- `d12`: 3D model of polarised dust emission with 6 layers, based on [137].

4.4.2 Synchrotron

- `s1`: Synchrotron emission is scaled as a power law, with a spatially varying spectral index. The emission templates are the *WMAP* 9-year 23 GHz Q/U maps [23]. The polarisation maps have been smoothed with a Gaussian kernel of 5 degrees, and small scales are added.

- `s0`: simplified version of the `s1` model, with a fixed spectral index of $\beta_{\rm s} = -3.1$ all over the sky.

- `s5`: power law model based on the same *WMAP* templates of `s1`. Small-scale fluctuations have been added to the templates in the `logpoltens` formalism.

The spectral index map from `s1` is rescaled based on *S-PASS* data, and had small scales added.

- `s4`: simplified version of the `s5` model, with a fixed spectral index of $\beta_{\rm s} = -3.1$ all over the sky.
- `s7`: power law model with a curved index. The model uses the same templates and spectral index map of `s5`.

All these models can be combined to build views of the sky with particular features and different levels of complexity. In the following Parts, we will employ the `PySM` package to simulate foreground emission.

Part II

Looking at the sky with the *Simons Observatory*

Until 20 years ago, the only way to properly observe the CMB anisotropies was to launch detectors on space satellites or sub-orbital balloons, in order to overcome the obstacle represented by the Earth atmosphere. A good part of the microwave radiation coming from space is absorbed or scattered by water molecules. The technological progress enabled the construction of ground-based observatories which can now compete with satellites in terms of sensitivity and extraction of cosmological information. All the three approaches have advantages and disadvantages, which we briefly describe. Even if we will not cover it here, we also note the existence of pioneering CMB observations made with high-altitude rockets, before the advent of satellites.

- **Space-based telescopes** can access the full sky from privileged positions, as the Sun-Earth Lagrangian L2 spot. They can observe the largest angular scales, without Earth-related contaminations as the atmosphere or ground emission. The observing conditions are extremely stable, with negligible thermal variations in a cool environment. However, the dimensions and weight of the telescopes are limited by logistic requirements, as they must fit into carriers of a certain size. Moreover, satellite observatories can not be modified or fully characterised after launch, as they will be in deep space; in case of malfunctioning, there is not much we can do from the ground to make it work again. Well-known and reliable technologies must be adopted, meaning that recent, state-of-the-art instrumentation can not be used. Ultimately, space mission are extremely costly and time-scales for them being operational are very long, requiring large investments with a non-negligible risk on the success rate.

- **Ground-based telescopes** are instead not able to observe the whole celestial sphere, since they are limited by their location on the planet. This implies that they cannot access the largest angular scales. The presence of the atmosphere and ground emission makes observations more complicated; the construction site must be carefully chosen, preferring very dry places at high altitudes. The Atacama desert in Chile and the South Pole are usually considered as the sites with the best observational conditions for CMB science. The main advantages of ground-based observatories include: the freedom to build large and heavy telescopes with a huge amount of detectors, increasing the sensitivity; the possibility to modify, fix, upgrade, and calibrate the instrumentation even after construction, adding more detectors or the most recent technology. They are much cheaper than their space counterpart, and the operation phase can be achieved in a way shorter amount of time.

- **Balloon-borne telescopes** represent a middle point between the space and ground approaches. Microwave telescopes can be mounted on balloons and launched for short flights into the stratosphere, gathering sky data for days or weeks at most. They can then be recovered after landing and reused. These sub-orbital missions allow for a cleaner view of the sky with respect to ground observations, since most of the atmosphere is not present. Moreover, they are much cheaper than satellites, and require a shorter timescale for operations. There are of course several disadvantages: balloon-borne telescopes can only observe very small portions of the sky from particular Earth sites, which guarantee a stable environment for the flights; the integration time is very short, making it very difficult to extract faint signals from the dominant noise; the risk linked to environmental factors is not negligible and must be taken into account. As we saw in section 3.5, many trailblazing measurements of the CMB anisotropies were realised with balloons; they are now mostly used as pathfinders for satellite missions.

This Part of the manuscript is dedicated to some of the CMB science that is possible to do with ground-based experiments, focusing in particular on the recently built *Simons Observatory* (*SO*), on which I have worked since the start of my PhD. The *SO* is a continuation of ground-based experimental efforts which have provided a number of exciting discoveries over the last fifteen years, including the first-ever detections of the CMB gravitational lensing power spectrum in both temperature [55, 211] and polarisation [97], lensing by galaxy clusters [21], the kinematic SZ effect [96], and the thermal SZ effect associated with radio galaxies [90]. They have compiled catalogues of SZ-selected galaxy clusters [29, 102] comparable in size to that extracted from the full-sky maps made by *Planck* [163], including many of the most extreme-mass and highest redshift clusters known [143]. Cosmological parameter constraints from ground-based temperature and polarisation power spectra are close to and competitive with the definitive ones produced by *Planck* [166]. The *Planck* full-sky maps still provide essential information at the largest angular scales and in frequency bands which are inaccessible from the ground [168].

Chapter 5

The *Observatory*

The *Simons Observatory*[1] (*SO*) is a brand new experimental cosmology facility located in the Atacama desert, in Chile's Parque Astronomico, at an altitude of 5,200 meters (17,000 ft). The *Atacama Cosmology Telescope*[2] (*ACT*) [150, 130] and the *Simons Array* [199] were located at the same site, and they have now been decommissioned and replaced by the *SO*. The *Cosmology Large Angular Scale Surveyor*[3] (*CLASS*) [73] experiment is located nearby.

The *SO* project is sustained by a collaboration of over 300 scientists from more than 50 institutions, constituted in 2016. The collaboration efforts led to the construction of the experiment, which currently consists of a Large Aperture Telescope (LAT) with a 6-meter primary mirror similar in size to *ACT*, and three 0.5-meter refracting Small Aperture Telescopes (SATs) similar in size to *BICEP3*. The LAT provides high angular resolution to measure small angular scales, surveying 40% of the sky; the small telescopes have a larger field of view and are able to better control atmospheric contamination in order to measure larger angular scales, focusing on deeply observing one of the cleanest regions of the sky, representing a 10% fraction. The LAT optical designs provides much larger focal planes with respect to previous experiments. The initial plan is to deploy a total of 60,000 detectors, approximately evenly split between the LAT and the set of SATs. Each detector pixel is sensitive to two orthogonal linear polarisations and two frequency bands. This number of detectors represents an order of magnitude increase over the size of previous microwave detector arrays, and is more total detectors than have been deployed by all previous CMB experiments combined. Plans to expand and enhance the facility are already underway: three more SATs are being built, two funded by UK collaborators and the other by Japanese ones, and a program

[1] https://simonsobservatory.org/
[2] https://act.princeton.edu/
[3] https://sites.krieger.jhu.edu/class/

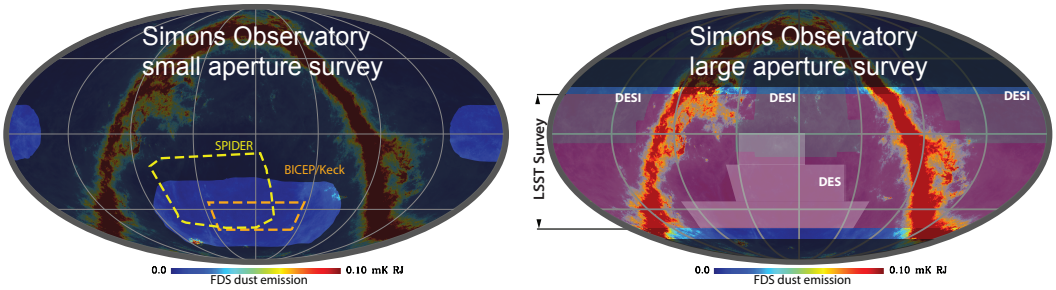


Figure 5.1. Anticipated coverage (lighter region) of the SATs (*left*) and LAT (*right*) in equatorial coordinates, overlaid on a map of Galactic dust emission. The survey regions of other experiments are also indicated. Figure from [6].

to double the number of LAT detectors is in place. The *SO* surveys will cover a sky region (show in figure 5.1) which overlaps many astronomical surveys at other wavelengths, including the *Vera Rubin Observatory*[4], as well as the *Dark Energy Survey*[5] (*DES*), the *Dark Energy Spectroscopic Instrument* (*DESI*)[6], and the *Euclid*[7] satellite.

The *SO* achieved first light in 2024 with one of the SATs, while the LAT achieved first light in February 2025, marking the end of the *SO* nominal project construction phase, and the beginning of science operations. At the time of writing (August 2026), the two mid-frequency (MF) SATs are observing the sky at 93 and 145 GHz, collecting early science data. The third SAT is taking commissioning data, at 225 and 280 GHz. Table 5.1 reports the forecasted properties of the *SO* surveys, both for the SATs and LAT, specifying their frequency channels, the resolution, and the noise levels for two different scenarios, one more conservative (baseline) and the other more optimistic (goal), to be reached after 5 years of observations.

The *SO* targets several scientific goals; here we briefly describe some of them, referring to the forecast papers for the full list and more details [6, 1, 3]:

- **Tensor-to-scalar ratio.** The nominal *SO* configuration aims to measure the tensor-to-scalar ratio r, with a sensitivity of $\sigma(r) = 0.003$ for a model without primordial B-modes ($r = 0$). This will enable at least a 3σ measurement of primordial gravitational waves if $r \geq 0.01$. If case of no signal detected, the *SO* would exclude models with $r \geq 0.01$ at more than 99% confidence. Lowering the current upper limit will also better constrain non-inflation models. This constraint on r will come from the large-scale B-modes measured by the *SO* SATs, targeting the range of scales $30 < \ell < 300$ with particular attention

[4] formerly known as the *Large Synoptic Survey Telescope* (*LSST*), https://rubinobservatory.org/.
[5] https://www.darkenergysurvey.org/.
[6] https://www.desi.lbl.gov/.
[7] https://sci.esa.int/web/euclid/.

Table 5.1. Properties of the planned *SO* surveys, from the 2019 forecast paper [6]. Some of these values will obviously be different in the current operational phase. Noise columns give anticipated white noise levels for temperature, with polarisation noise $\sqrt{2}$ higher as both Q and U Stokes parameters are measured. Noise levels are quoted for a homogeneous hits map, and are reported in $\mu\mathrm{K}_{\mathrm{CMB}}\cdot$ arcmin units.

	SAT ($f_{\mathrm{sky}} = 0.1$)			LAT ($f_{\mathrm{sky}} = 0.4$)		
frequency [GHz]	FWHM [arcmin]	noise (baseline)	noise (goal)	FWHM [arcmin]	noise (baseline)	noise (goal)
27	91	35	25	7.4	71	52
39	63	21	17	5.1	36	27
93	30	2.6	1.9	2.2	8.0	5.8
145	17	3.3	2.1	1.4	10	6.3
225	11	6.3	4.2	1.0	22	15
280	9	16	10	0.9	54	37

near the recombination bump at $\ell = 80$. The forecasted constraints on r and the spectral index of primordial perturbations n_{s} are shown in figure 5.2. The expanded version of the *SO* with six SATs will improve the sensitivity to $\sigma(r) = 0.001$.

- **Non-Gaussian perturbations.** The Observatory targets a measurement of the non-Gaussianity of the primordial perturbations, improving current constraints. This will be obtained via two methods: (i) by correlating the CMB kinematic SZ effect (from LAT temperature maps) with the galaxy distribution in tomographic redshift bins from *Rubin*; and (ii) by correlating the CMB lensing field (from LAT temperature and polarisation maps) with the *Rubin* galaxy distribution. Additional information will come from the bispectrum estimated from the LAT temperature and polarisation maps.

- **Effective number of relativistic species.** A Universe with three neutrino species and no additional light species provides $N_{\mathrm{eff}} = 3.046$ effective relativistic species. The *SO* aims to measure $\sigma(N_{\mathrm{eff}}) = 0.07$, using the primary CMB temperature and polarisation power spectrum measured from the LAT, more than halving the current limit from the *Planck* satellite.

- **Neutrino mass.** The *SO* targets the measurement of the total mass in the three neutrino species with $\sigma\left(\sum m_\nu\right) = 0.04$ eV, possibly providing a clear indication of a non-zero mass sum if $\sum m_\nu \geq 0.1$ eV.

- **Derived Hubble constant.** The *SO* aims to reduce the current uncertainty on the Hubble constant H_0 derived from the primary CMB within the ΛCDM model, reaching a half-percent measurement on H_0 from the LAT temperature and

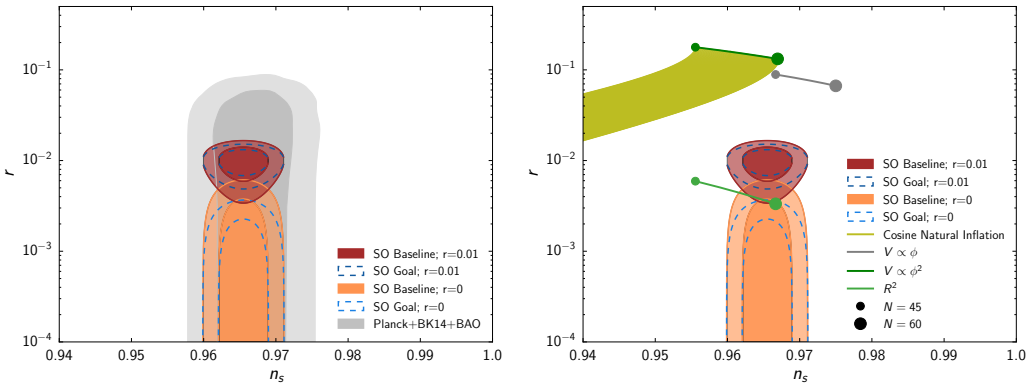


Figure 5.2. *Left*: summary of *SO* forecasts for the primordial power spectrum parameters $n_{\rm s}$ and r for two example cases: vanishingly small primordial tensor modes and primordial perturbations with $r = 0.01$. The contours correspond to 68% and 95% confidence levels. The *SO* baseline (goal) forecasts are shown as filled (dashed) contours. Constraint in this parameter space from *Planck*, *BICEP2/Keck*, and BAO are shown in grey. *Right:* same as above, with predictions for $n_{\rm s}$ and r from some inflationary models.

polarisation power spectra. This will provide information of any discrepancy found between different measurements.

- **Reionisation.** If the duration of the reionisation phase is $\Delta z > 1$, the *SO* aims to measure the average duration with a significance between 2σ and 3σ, thus constraining models for the ionising process and probing the properties of the first galaxies, quasars, and the intergalactic medium in the reionisation epoch.

With these scientific questions in mind, the *SO* will represent the state-of-the-art of CMB science for years to come, producing maps with unprecedented sensitivity. An example of the preliminary maps coming from the initial science observations of the *SO* are reported in figure 5.3, where the E-mode signal is evident from the + and $\times$ patterns in Q and U Stokes parameters, respectively. These maps are just a first demonstration of the *SO* capabilities but they are already much more sensitive than the *Planck* final data products. They are expected to greatly improve in the near future, with a longer integration time included in the analysis.

During the last few years, I have been a member of the *SO* collaboration, with my work focusing on the SATs data analysis and preparation. In particular, I contributed within the BB analysis working group, whose main task is to produce constraints on the amplitude of the primordial B-mode signal starting from multi-frequency sky maps observed by the SATs. This huge task encompasses several steps and a variety of challenges: first, foregrounds contamination must be mitigated through component separation techniques; then, the tensor-to-scalar ratio must be extracted; finally, robustness tests need to be performed to validate the eventual detection and quantify the residual foregrounds contamination. In *SO*, the first two steps are

performed using three different component separation pipelines, which we briefly summarise here, referring to [218] for a complete description:

- **cross-C_ℓ.** This method is based on a parametric multi-frequency power spectrum-based component separation [26, 24]. Auto- and cross-power spectra $C_\ell(\nu \times \nu')$ of multi-frequency sky maps are fitted through a Gaussian likelihood[8] which compares the data vector against a theoretical model of CMB and foregrounds emission, and instrumental features. Power spectra are generally evaluated outside of a mask covering the most contaminated regions of the sky. Thus, the full sky is not available. As we mentioned in section 3.4, the incomplete sky coverage is the source of two effects [122, 36]: mode-coupling, coupling different ℓ and biasing the estimator, and E-to-B leakage causing contamination of the larger E-modes into B-modes. These effects are taken into account through the `NaMaster` package [12], which corrects for the mode-coupling by providing an analytical estimate of it, and mitigates the E-to-B contamination with a map-level purification. Moreover, noise correlation in computed auto-spectra is addressed by exploiting data splits, that is, map built by different chunks of data which contain the same sky signal, but different and possibly independent noise properties. In the case of real data, splits may be formed by subdividing data among observation periods, sets of detectors, sky patches, etc.

- **NILC.** The second method is based on the needlet version of the Internal Linear Combination (ILC) technique described in section 4.3. It performs blind, minimum-variance cleaning on multi-frequency sky maps, providing a single reconstruction of the CMB signal. From this recovered map, the angular power spectrum is computed and cosmological parameters extracted employing a Gaussian likelihood which compares data with theoretical models.

- **Map-based.** The last method is a map-based component separation method derived upon the `fgbuster` code [71, 175], which separates frequency maps into component maps, through a spectral likelihood optimisation. Angular power spectra can then be computed and cosmological parameters inferred, adopting a Gaussian likelihood and comparing with theory.

The possibility to run the cosmological analysis with three independent pipelines allows for strong robustness checks of the final results. In principle, all the analysis should agree on the final outcome and converge to similar values, in order for the results to be properly validated. This is a powerful way to avoid blind biases coming

[8] In general, the Hamimeche-Lewis likelihood [93, 94] is a better approximation. For a large number of sampled modes, the two likelihoods provide the same results.

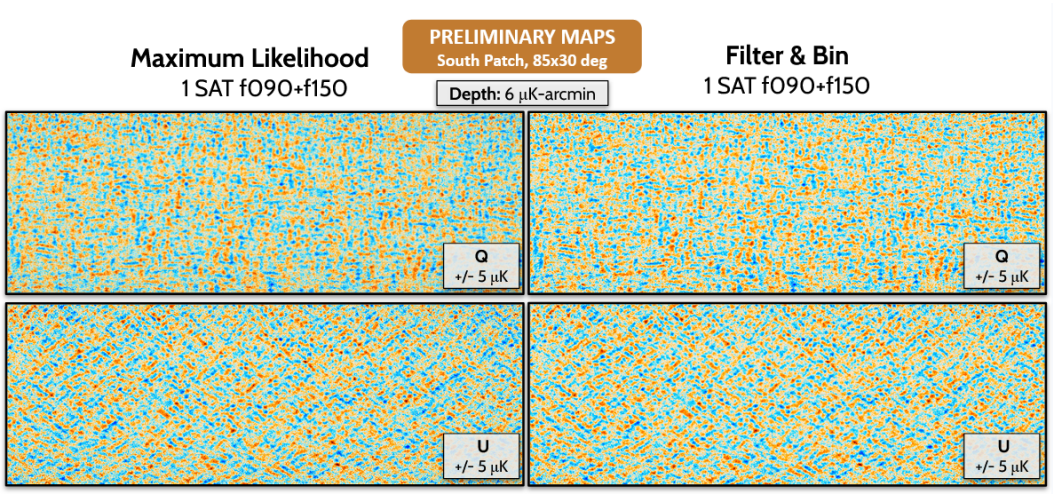


Figure 5.3. Small cutout of the preliminary *SO* CMB maps in polarisation, as observed by one of the SAT, co-added from the 90 and 150 GHz bands, and including data from ~ 14 months of observations. *Left* and *right* columns show the maps produced with two different, independent map-making algorithms, namely maximum likelihood and filter and bin, giving consistent results. Figure credit: *SO* Collaboration.

from choosing a single analysis suite. I will spend the rest of this Part describing how I contributed to these data analysis pipelines, in the context of the *SO* project.

Chapter 6

External data for the *SO*

As I mentioned in the previous chapter, all the nominal *SO* instruments are currently on-sky. While writing this manuscript, the only two telescopes collecting science data are the mid-frequency (MF) SATs, observing at the so-called "CMB frequencies" of ~ 90 and ~ 150 GHz. On one side, foregrounds contamination at these frequencies is minimum, representing the best observational windows for CMB mapping. On the other hand, these two channels alone are not enough to fully characterise the spectral and spatial properties of synchrotron and dust emissions, which is of paramount importance for their subtraction and mitigation. Component separation techniques necessarily require measurements of the SED of foregrounds emission at different channels, and possibly at frequencies which are dominated by a single component, in order to properly disentangle their impact.

Therefore, early science SATs observations will miss crucial information on foregrounds; the first constraints on the tensor-to-scalar ratio cannot rely solely on SATs maps, at least until the high- and low-frequency instrumentation will be available. The most immediate way to compensate the missing foregrounds information is to integrate SATs observations with external data, that is, data coming from other microwave experiments targeting the measurement of the CMB. Datasets from different observatories can be combined together in order to make up for their respective deficiencies or flaws. Figure 6.1 shows the SED of synchrotron and thermal dust polarised emission, with the SATs frequency channels indicated by the vertical bands. Foreground-dominated channels are currently not available.

In order to characterise foregrounds emission, we decided to exploit data coming from three separate experiments:

- The latest ***Planck*** data release[1] (PR4, or `NPIPE`) [173] represent the state-of-the-art concerning polarised observations of the full microwave sky. For the

[1] available on the *Planck* Legacy Archive (PLA) and at the National Energy Research Scientific Computing Center (NERSC).

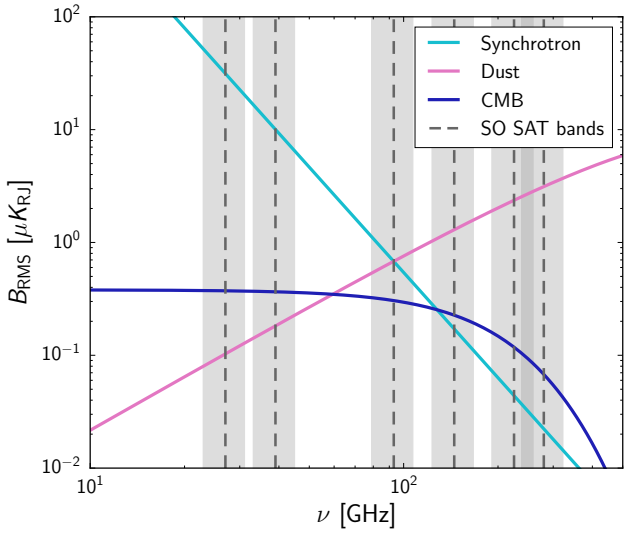


Figure 6.1. Frequency dependence, in Rayleigh-Jeans (RJ) brightness temperature, of the synchrotron and thermal dust emission at degree scales within the SATs footprint, compared to the CMB lensing B-mode signal. The turnover of the modified blackbody law for dust lies above this frequency range. The SATs channels are shown as vertical bands. Figure from [6].

high-frequency regime, we employ polarised maps at 217 and 353 GHz, being the best available tracers of Galactic dust, which is the dominant contribution to the sky emission at these frequencies. For the lower frequencies, we adopt polarised maps at 30 GHz, which are instead good tracers of synchrotron emission and point sources. We also use *Planck* data at "CMB frequencies" (100 and 143 GHz) to cross-correlate and enhance the CMB signal recovered by SATs initial science observations, which will eventually be dropped once the *SO* sensitivity gets deeper. We employ both the *Planck* PR4 full-mission maps and the two data splits denoted A and B, built by separating the focal plane in two independent detectors sets, and performing the map-making procedure independently. The systematics between these splits are expected to be uncorrelated. The detector-set splitting was not possible for the 30 GHz channel. Instead, for this frequency the split is done time-wise, with set A comprising operational years 1 and 3, while set B has years 2 and 4.

- Even if generally outperformed by its successor *Planck*, ***WMAP*** [23] low-frequency polarised data[2] are still holding as very good tracers of synchrotron and point sources emissions. Maps at 23 and 33 GHz (called K and Ka bands, respectively) are adopted, even if the latter channel should not carry many more information with respect to the *Planck* 30 GHz data. We use

[2] available on the LAMBDA website.

both the 9-years full mission maps and the 9 single-year data, adopted as time-wise splits.

- The *S-band Polarisation All Sky Survey* (***S-PASS***) [44, 115] is a ground-based project aiming to map the polarised emission of the entire southern sky (declination $< 1^\circ$) at 2.3 GHz, with the Parkes radio telescope, located in NSW, Australia, and operated by CSIRO-CASS. The very good signal-to-noise ratio of *S-PASS* maps[3] allows them to be optimal tracers for synchrotron and point sources emissions, at a frequency not usually sampled by microwave experiment targeting CMB science. The survey does not sample the full sky, but it encloses the deep patch targeted by the SATs.

The upper row of figure 6.2 reports some examples of the adopted external data, showing Stokes Q maps for *Planck* 353 GHz, *WMAP* 23 GHz, and *S-PASS* 2.3 GHz data. We notice the differences in the morphology of the *Planck* and *WMAP* maps, with the former dominated by thermal dust emission and the latter by synchrotron radiation. We can also see how small scale structures in the *Planck* maps are lost with the downgrading, as expected. We are not interested in them for our purposes. The *WMAP* data are not downgraded as they are natively at the same pixel resolution of the SATs products.

Regarding *S-PASS*, it is worth noticing how the amplitude of the diffuse synchrotron radiation at 2.3 GHz is a factor of about 10^3 larger than the *WMAP* 23 GHz channel (and a similar factor for *Planck* 30 GHz). This difference alone already suggests the benefit of using very low frequency data to characterise the Galactic synchrotron emission, given the high signal-to-noise ratio (> 3 in 94% of the pixels) which can be reached even in regions where the emission is fainter. On the other hand, *S-PASS* data show an evident granularity of the emission close to the Galactic plane (for Galactic latitudes $|b| \lesssim 20^\circ$). This is indicative of depolarisation effects induced by the Faraday rotation of the polarisation angle of light propagating through the ionised Galactic medium in the presence of the magnetic field, resulting in different rotations affecting different regions across the entire Galactic disc [44].

6.1 External data preprocessing

Operatively speaking, combining different experiments to extract cosmological information is not really a straightforward task. Each telescope collects sky data in its own way and configuration, resulting in final datasets which have different properties. In order to compare apples with apples, data from several experiments

[3] available on the *S-PASS* website.

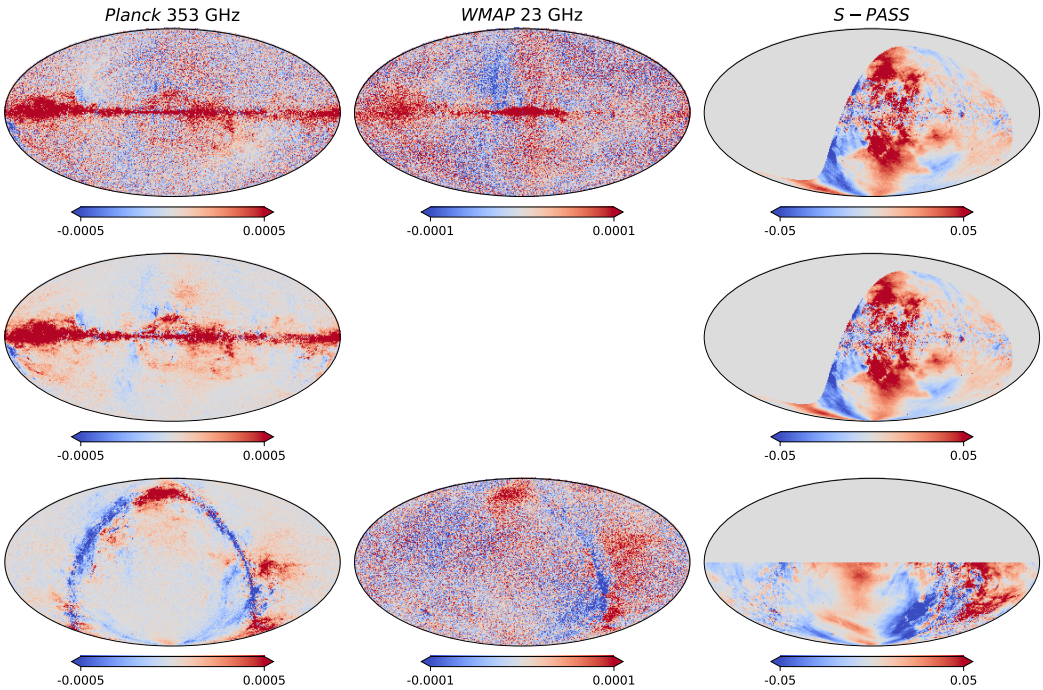


Figure 6.2. Examples of the external datasets adopted to integrate *SO* SATs observations, in Stokes Q maps. Units are $\mu\mathrm{K_{CMB}}$. *From left to right: Planck* `NPIPE` 353 GHz channel, *WMAP* 23 GHz channel, and *S-PASS* 2.3 GHz. *From top to bottom:* input maps, then downgraded to $N_{\mathrm{side}} = 512$, and rotated from galactic to equatorial coordinates. *WMAP* data have not been downgraded as their are natively at the same resolution of the SATs.

must be manipulated and brought to a common configuration as close as possible to the one they would have if observed by the same telescopes. We refer to these procedure as the "preprocessing" of external data, with the idea of preparing them for the combination with *SO*.

Some of the features changing from one experiment to the other include:

- **Pixel resolution.** In general, a typical experiment realises sky maps which are projected onto a sphere, tessellated with pixels. The number and dimensions of these pixels are different for each experiment, depending on the technology and on the science target. The higher the number of available pixels, the smaller the angular scale in the sky that is possible to sample. In a commonly used pixelation named `HEALPix`[4] [88], the celestial sphere is subdivided in pixels of equal area, where the resolution is defined through the N_{side} parameter which provides the total number of pixels as $N_{\mathrm{pix}} = 12 \times N_{\mathrm{side}}^2$. Obviously, a map with more pixels carries more numerical information and requires a larger amount of disk space to be stored. The *SO* SATs sky data are being mapped with a pixel resolution of $N_{\mathrm{side}} = 512$, which is the same resolution of the

[4] https://healpix.sourceforge.io/.

WMAP products. The *Planck* 30 GHz maps have a resolution of $N_{\text{side}} = 1024$, same as *S-PASS* data, while for the 100, 143, 217, and 353 GHz channels it is $N_{\text{side}} = 2048$.

- **Coordinates system.** Sky maps can be expressed in different astronomical coordinates systems. The most commonly used in microwave surveys are the equatorial (`C`) and galactic (`G`) systems: in the former, the fundamental plane is the celestial equator, with the two coordinates being declination and right ascension; in the latter, the fundamental plane is the Galactic one, and the two coordinates are the Galactic latitude and longitude. The *SO* SATs sky data are reported in equatorial coordinates, while *Planck*, *WMAP*, and *S-PASS* are in Galactic coordinates.

- **Observed sky fraction.** Ground-based telescopes do not have the possibility to observe the full celestial sphere, like satellite can instead do. They can only access a fraction of the sky $f_{\text{sky}} < 1$. The *SO* SATs are targeting deep observations of a small, clean region which amounts to about 10% of the sky, while *S-PASS* observed the entire southern sky, covering half of the sphere. To be combined, data from different experiments must be focused on the same sky patches.

- **Filtering.** Space-borne instruments have the great advantage of operating in a low-contamination environment. Ground-based experiment have instead to deal with many undesired effects, as the presence of the atmosphere or the ground emission. These features are usually treated at the time-streams level, by filtering them out of the data. The impact of such filtering must be also included in the external data in order to proper combine different experiments. This topic will be discussed in more details in section 7.2.

Table 6.1 summarises the different features of the considered experiments. Thus, in order to combine different datasets, they have to be first be brought to a common reference configuration, which in our case is the *SO* SATs one. To do so, external data must be preprocessed through a series of operations, namely downgrading, coordinate transformation, masking, and filtering. I will now describe how I carried out these steps. The majority of the maps manipulation is done with functions included in the `healpy`[5] [223] package.

[5] https://healpy.readthedocs.io/en/latest/

Table 6.1. Properties of the different datasets considered in this Part, including the *SO* SATs and external data (*Planck*, *WMAP*, *S-PASS*) configurations.

Feature	*SO* SATs	*Planck*	*WMAP*	*S-PASS*
frequencies [GHz]	90/150	30/100/143/217/353	23/33	2.3
N_{side}	512	1024/2048	512	1024
Pixel size [arcmin]	~ 6.9	$\sim 3.4/1.7$	~ 6.9	~ 3.4
Coord. system (G/C)	C	G	G	G
Sky fraction f_{sky}	~ 0.1	1	1	~ 0.5
Filtering	yes	no	no	yes

6.1.1 Downgrading

To properly combine the signal coming from different experiments, the sky maps must be pixelated on the same number of pixels. In other words, they must have the same resolution, defined by the N_{side} parameter. The simplest and fastest approach would be to operate in pixel space, taking the higher-resolution maps and just average the value of groups of sub-pixels, to merge them into a single super-pixel, a procedure usually adopted by general-purposes softwares to downgrade images. This is what the `ud_grade` function of the `healpy` package does. Even if computationally cheap, this procedure is not ideal to get lower-resolution maps for scientific analyses: the small scale information contained in the higher multipoles is not properly treated, inducing aliasing features showing up as unphysical artifacts in the downgraded maps (see [197] for some examples). Moreover, performing pixel averages on non-scalar maps like the polarisation Q and U do not take into account the vectorial nature of the Stokes parameters. Additionally, there are other instrumental properties that are not considered by the simple pixel average, namely the beam and the pixel window function. The instrumental beam of the telescope is responsible for the smoothed picture provided by the detectors. The optics bring the infinite resolution sky to a finite resolution image, averaging out the information smaller than the angular resolution. This resolution is usually modelled as a Gaussian, characterised by a full-width at half-maximum (FWHM) size typically quoted as the telescope's angular resolution. The pixel window function is a model of the pixelation operated by the instrument, where observations of close but different line-of-sights are averaged into a single pixel for the sake of map-making. The impact of the beam and pixel window function are very similar, resulting in a smoothing of the smaller scales. It is possible that during downgrading one needs to apply a different beam to smooth more scales, or to convolve with a different pixel window function, since downgrading changes the pixel number.

All these issues are taken into account in the correct procedure to downgrade a sky map, which consists in operating in harmonic space and exploiting the spherical

harmonics decomposition described in section 1.3.1 and 3.4. Any scalar field $X \in [T, E, B]$, can be expanded in a spherical harmonics basis, with coefficients given by

$$a_{\ell m} = \int_0^{2\pi} \int_0^{\pi} Y^*_{\ell m}(\theta, \varphi) X(\theta, \varphi) \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \,, \tag{6.1}$$

where in map space, the θ and φ coordinates are replaced by the particular pixel, and the integrals become discrete summations. The maximum multipole moment correctly described by a pixelated map is usually taken to be $\ell_{\mathsf{max}} = 3N_{\mathsf{side}} - 1$. The coefficients encode all the information contained in the sky maps, and can be used to go back and forth between the two spaces. In harmonic space, downgrading a map simply implies to cut the expansion at the maximum multipole $\ell^{\mathsf{out}}_{\mathsf{max}} = 3N^{\mathsf{out}}_{\mathsf{side}} - 1$ we are interested in, and removing all the information at smaller scales $\ell > \ell^{\mathsf{out}}_{\mathsf{max}}$. Different beams or pixel window functions can be incorporated in harmonic space, deconvolving by the ones of the original maps and re-convolving with new ones, to provide the final modified, "downgraded" coefficients:

$$a^{\mathsf{dg}}_{\ell m} = \frac{p^{\mathsf{out}}_\ell}{p^{\mathsf{in}}_\ell} \frac{b^{\mathsf{out}}_\ell}{b^{\mathsf{in}}_\ell} a^{\mathsf{in}}_{\ell m} \Big|_0^{\ell^{\mathsf{out}}_{\mathsf{max}}} \,, \tag{6.2}$$

where p^{in}_ℓ, p^{out}_ℓ and b^{in}_ℓ, b^{out}_ℓ are the spherical harmonics transforms of the input and output pixel window functions and instrumental beams, respectively, and $a^{\mathsf{in}}_{\ell m}\Big|_0^{\ell^{\mathsf{out}}_{\mathsf{max}}}$ are the coefficients of the original maps clipped to $\ell^{\mathsf{out}}_{\mathsf{max}}$. These modified coefficients can then be reprojected to sky maps with the inverse transform

$$\begin{aligned} T(\theta, \varphi) &= \sum_{\ell=0}^{\ell^{\mathsf{out}}_{\mathsf{max}}} \sum_{m=-\ell}^{m=\ell} a_{\ell m} Y_{\ell m}(\theta, \varphi) \,, \\ (Q \pm iU)(\theta, \varphi) &= \sum_{\ell=0}^{\ell^{\mathsf{out}}_{\mathsf{max}}} \sum_{m=-\ell}^{m=\ell} {}_{\pm 2}a_{\ell m} \, {}_{\pm 2}Y_{\ell m}(\theta, \varphi) \,, \end{aligned} \tag{6.3}$$

in order to obtain a downgraded version of the original maps.

6.1.2 Coordinates transformation

Sky maps from different experiments need to be converted to the same coordinates system, in order to combine data looking at the same astronomical objects. It is possible to convert ("rotate") between different coordinates systems using standard trigonometric relations; equations transforming from equatorial to Galactic

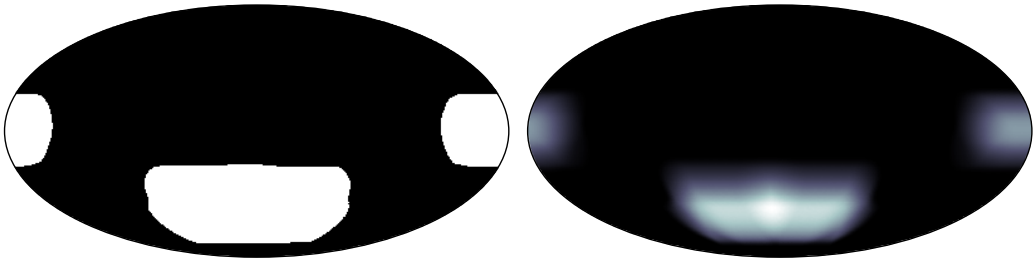


Figure 6.3. Nominal sky regions observed by the *SO* SATs [218], in equatorial coordinates. Unobserved pixels are in black. *Left:* binary mask, covering $f_{\text{sky}} = 10\%$ of the sky. *Right:* apodised mask weighted by the forecasted pixel observing time. Mask edges are apodised using a C1-type kernel with an 10 degrees apodisation scale.

coordinates are:

$$\begin{aligned} \sin(\alpha - \alpha_G)\cos(\delta) &= \cos(b)\sin(l_{\text{NCP}} - l)\,, \\ \cos(\alpha - \alpha_G)\cos(\delta) &= \sin(b)\cos(\delta_G) - \cos(b)\sin(\delta_G)\cos(l_{\text{NCP}} - l) \\ \sin(\delta) &= \sin(b)\sin(\delta_G) + \cos(b)\cos(\delta_G)\cos(l_{\text{NCP}} - l)\,, \end{aligned} \tag{6.4}$$

where α and δ are the equatorial coordinates (right ascension and declination), l and b are the Galactic coordinates (latitude and longitude), α_G and δ_G are the equatorial coordinates of the North Galactic Pole, and l_{NCP} is the Galactic longitude of the North Celestial Pole (NCP).

In `healpy` routines, this rotation can be done directly on the (downgraded) spherical harmonics coefficients $a^{\text{dg}}_{\ell m}$, using the Wigner rotation matrices defined via the above equations. Coordinates-transformed maps can then be obtained by reprojecting the rotated coefficients, following eq. (6.3).

6.1.3 Masking

All the maps must obviously be combined while observing the same sky regions, and in particular the same pixels. Since the *SO* SATs survey the smaller patches ($f_{\text{sky}} = 10\%$), external data have to be masked to exclude the pixels unobserved by the SATs. This is achieved by building a binary mask where the unmapped pixels are set to zero, and observed pixels are equal to 1. Pixels close to the mask edges have to be smoothed out to avoid artifacts due to sharp transitions. This smoothing procedure is called apodisation and it is performed through the `mask_apodization` routine of the `NaMaster`[6] package [12]. Pixels are multiplied by some functions as Gaussian or sinusoidal in order to smoothly connect pixel values to zero. Figure 6.3 shows one of the forecasted sky masks for the SATs survey, covering 10%

[6] https://namaster.readthedocs.io/en/latest/.

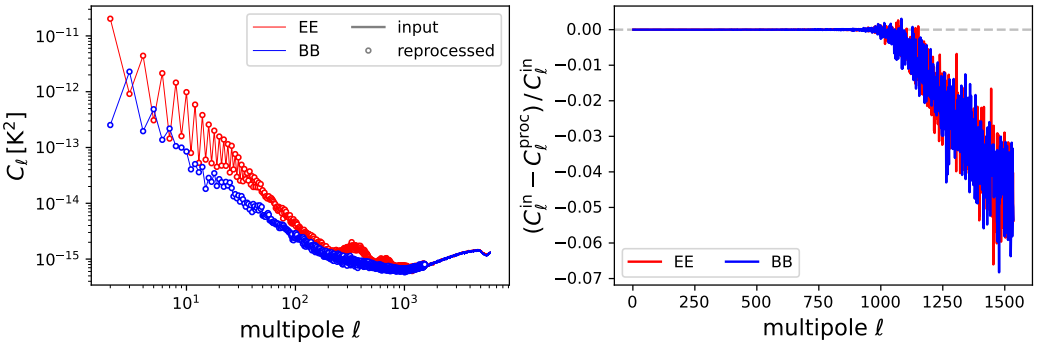


Figure 6.4. Validation of the external data pre-processing, comparing angular power spectra before (solid lines) and after (circles) the downgrading and rotation steps carried out in harmonic space (*left panel*). The spectral features are preserved, with differences (*right panel*) being at the 5% level at very small scales $\ell > 1200$. Data shown are from the *Planck* 100 GHz channel, in EE (red) and BB (blue).

of the sky, in binary and apodised versions (10° apodisation scale, using a "C1" apodisation kernel [89]).

6.1.4 Additional splits building for *WMAP*

WMAP data are further processed through an additional step. In the previous section we listed the external data we adopt for combination with *SO* ones, and we cited the 9 single-year *WMAP* maps. In order to have data splits with a larger signal-to-noise ratio, we co-add the single-year maps to form two "half-mission" (`hm1` and `hm2`) data splits, with the first comprising operational years 1 to 4, while the second has years 5 to 9. Splits are built through a weighted average of the single-year maps m_i, taking into account the amount of observational time $N_{\rm obs}$ spent on the pixels each year:

$$
\begin{aligned}
m_{\texttt{hm1}} &= \frac{\sum_{i=1}^{4} N_{\rm obs,i} m_i}{\sum_{i=1}^{4} N_{\rm obs,i}}\,, \\
m_{\texttt{hm2}} &= \frac{\sum_{i=5}^{9} N_{\rm obs,i} m_i}{\sum_{i=5}^{9} N_{\rm obs,i}}\,.
\end{aligned}
\tag{6.5}
$$

Also, in this way we have the same number of splits for the two satellite missions, *Planck* and *WMAP*, and we also reduce the computational cost of the cross-splits angular power spectra evaluation, with respect to using 9-years splits. Unfortunately, *S-PASS* data do not include splits, as only the full-mission maps are available.

Thus, external data must undergo these reprocessing steps in order to be combined with SATs data. The above passages are summarised in figure 6.2, where we show some examples of the original external maps, and their downgraded,

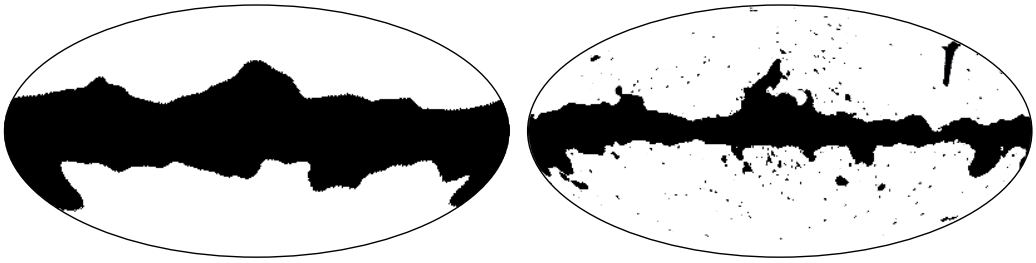


Figure 6.5. Binary version of the masks adopted for the characterisation of external data, in galactic coordinates. *Left:* mask covering the Galactic plane. *Right:* official mask used for the *Planck* analysis of polarised data, covering the Galactic plane, extra-galactic point sources, and other polarised structures.

rotated and masked versions. While developing this process, we have to be sure that all the important features of the maps are preserved, with a particular attention to the harmonics properties. It is thus extremely important to validate the entire process, and check that no distortions or artifacts are introduced at any step. Figure 6.4 shows a comparison between the auto angular power spectra of the original and preprocessed maps, in polarisation (EE, BB). Spectra are evaluated over the full sky with the `anafast` routine of the `healpy` package. The right panel reports their difference, which is at the few percent level at very large scales, above $\ell > 1000$. This implies the correct preservation of the power spectra in the processed maps, at least at the angular scales interesting for B-modes searches ($\ell < 300$).

6.2 External data characterisation

After the reprocessing phase, sky maps from external experiments are ready to be combined with *SO* data. But before being able to extract cosmological information from this data union, we need to properly characterise external data, in order for us to understand their features and look for the presence of eventual issues, which may cause problems in the successive interpretation of the results. Of course, *Planck*, *WMAP* and *S-PASS* data have been thoroughly investigated during the years, but each situation is different from the others, and some features negligible in other studies can be crucial for us. If these data contain any issue in some sky regions or in the patch observed by the *SO*, we have to be aware and understand them before the combination.

We will characterise the external data over masked skies retaining a large fraction of pixels, as it is usually done for space-based telescopes, and we will also focus on the smaller sky regions observed by the *SO* SATs (figure 6.3). In the former, the contamination induced by Galactic foregrounds is mitigated through

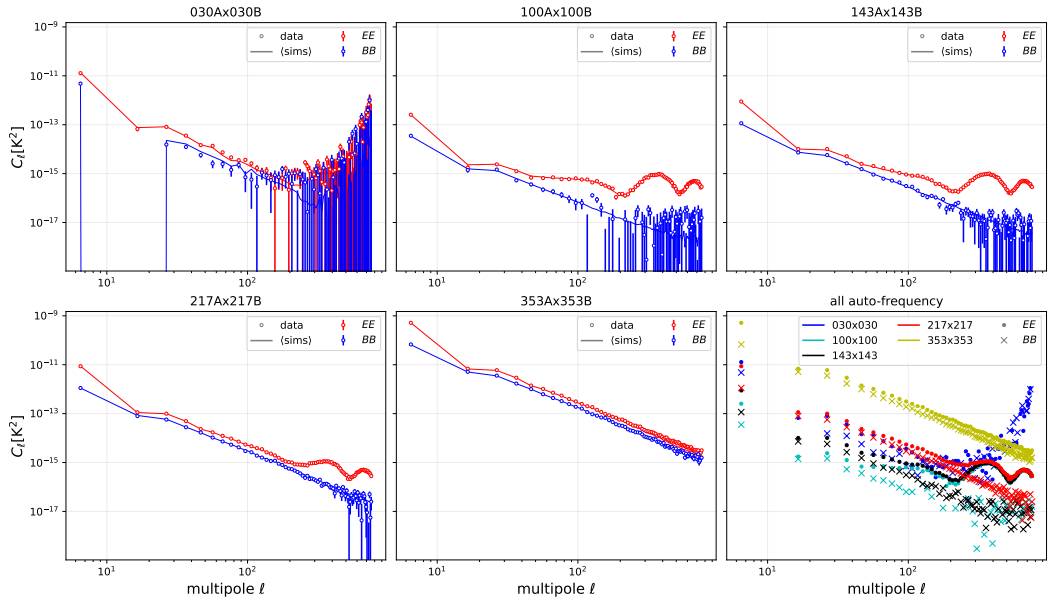


Figure 6.6. EE and BB auto-frequency angular power spectra of *Planck* NPIPE data, evaluated adopting the mask covering the Galactic plane, point sources, and other polarised structures (right panel of figure 6.5). At each frequency, cross-splits (A$\times$B) spectra are computed to avoid noise correlation bias.

the application of two *Planck* sky mask, shown in figure 6.5: one is a simple mask covering the Galactic plane and retaining $\sim 70\%$ of the sky, while the other is a more complex mask covering the Galactic plane, extra-galactic point sources, and other sky structures contaminating polarised observations. Both masks are available on the *Planck* Legacy Archive (PLA). In the following we will not show all the results but just a selection, to favour a smoother and more understandable discussion.

***Planck* data and simulations** The *Planck* NPIPE data release [173] comes with 600 high fidelity Monte Carlo simulations that include signals from CMB, foregrounds, noise, and systematics. These simulations are extremely useful for characterising the NPIPE data and also for testing analysis tools and choices with realistic frequency and detector-set maps, with known inputs. Along with frequency maps, the NPIPE simulations products also include so-called "residuals" maps, built by subtracting the input sky signal (CMB and foregrounds emission) from the total simulated maps. They thus represent the noise and the majority of systematics contained in the NPIPE maps.

In order to characterise the *Planck* data and simulations, we reprocess the latter in the same way as we did for the former, thus downgrading the maps to a pixel resolution of $N_{\rm side} = 512$ and rotating them from galactic to equatorial coordinates. Then, polarised auto- and cross-frequency angular power spectra are computed for

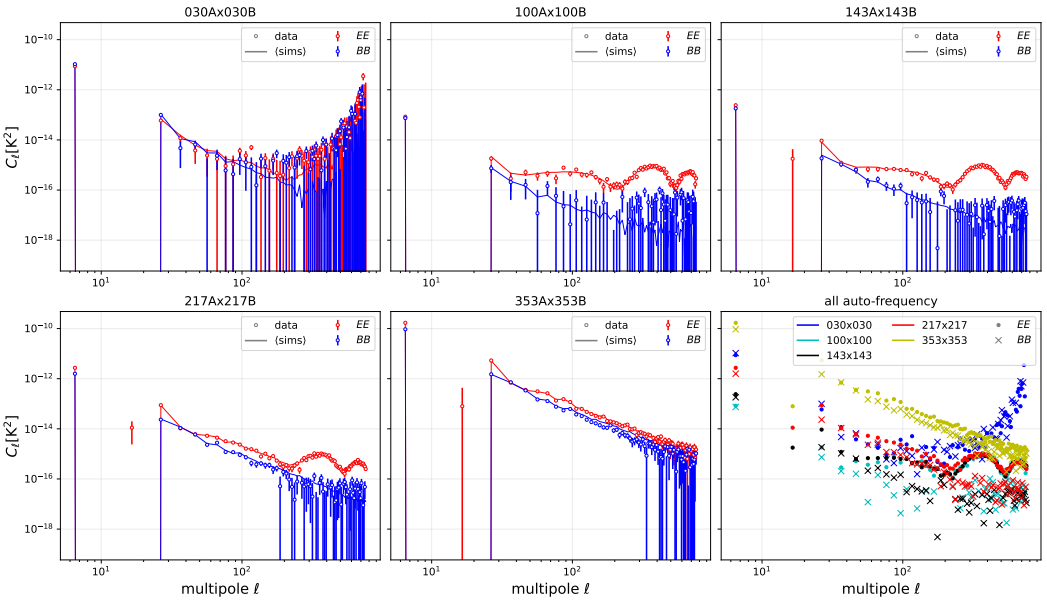


Figure 6.7. Same as figure 6.6, but here spectra are evaluated on the *SO* SATs sky patch of figure 6.3.

the data and for the 600 simulations, for the different frequency channels considered (30, 100, 143, 217, 353 GHz). Each channels has two splits A and B, built with two independent detector sets; we thus have 10 cross-frequency (considering the full-mission maps) and 5 auto-frequency spectra, for a total of 15. We then compute the average of the 600 simulations power spectra, and compare it with the data, to which we have associated an errorbar as the standard deviation of the simulations. To investigate the possible presence of systematics or noise features in the simulations, we also run a similar analysis on the "residuals" maps, which should only contain noise and known systematics. Their power spectra should be compatible with zero if no spurious correlations are present.

Figure 6.6 and 6.7 show the auto-frequency angular power spectra for *Planck* NPIPE data and simulations, both in EE and BB, evaluated on the mask covering the Galactic plane and point sources and on the SATs patch, respectively. Cross-splits spectra are computed in order to avoid the noise correlation which would be present in the full-mission auto-correlation, as described in section 3.4. Power spectra are in general compatible between data and simulation, with the BB data at low frequency being more dominated by noise and by the effect of the instrumental beam at smaller scales, which is way larger than the other channels. The first two peaks of the CMB E-mode signal are evident in the 100 and 143 GHz channels, where foregrounds emission is minimum, while the lower and higher frequencies are dominated by the power-law shape typical of synchrotron and dust, respectively.

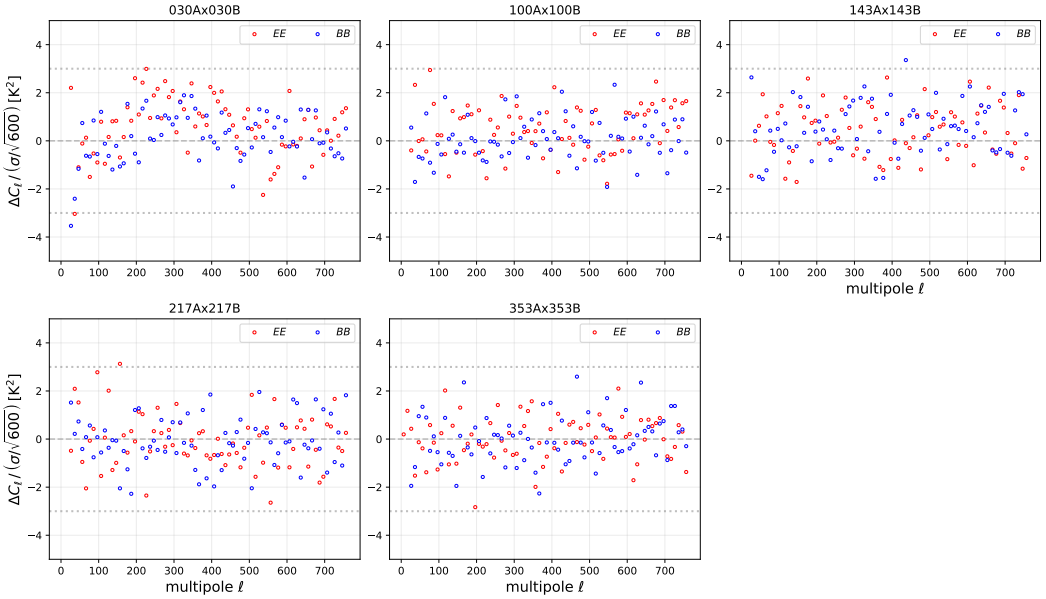


Figure 6.8. EE and BB auto-frequency angular power spectra of *Planck* NPIPE residuals simulation, evaluated adopting the mask covering the Galactic plane (left panel of figure 6.5). At each frequency, cross-splits (A$\times$B) spectra are computed to avoid noise correlation bias. Results are reported normalised by the error on the simulations.

Figures 6.8 and 6.9 report the auto-frequency power spectra of the residuals maps, normalised by the error on the simulations computed from their standard deviation as $\sigma/\sqrt{600}$. Power spectra are clearly compatible with zero, with points being enclosed in the 3σ regions. This demonstrate that there are no spurious correlation indicating systematics, or at least they are not able to be impactful within *Planck* sensitivity, both when considering large sky fractions or when looking into the *SO* smaller patches.

Incidentally, the errorbars derived here for the mask covering the Galactic plane have also been used by Léo Vacher (one of my co-supervisors) and collaborators in one of his works [208] of which I am co-author, where he used the framework of moment expansion to produce new 3D models of thermal dust emission, containing the maximal amount of complexity allowed by the scatter of the *Planck* data.

***WMAP* data** As we already mentioned, for *WMAP* we have data at two frequency channels: 23 and 33 GHz (K and Ka bands). No end-to-end simulations are available. Figure 6.10 reports the auto- and cross-frequency angular power spectra of the *WMAP* data, computed on the right mask of figure 6.5, both in EE and BB. The noise correlation bias is evident in the auto-spectra, while it greatly improves in the cross-frequency ones, allowing to reach higher multipoles and clearly see the synchrotron contribution as a power-law, until beam effects start to dominates.

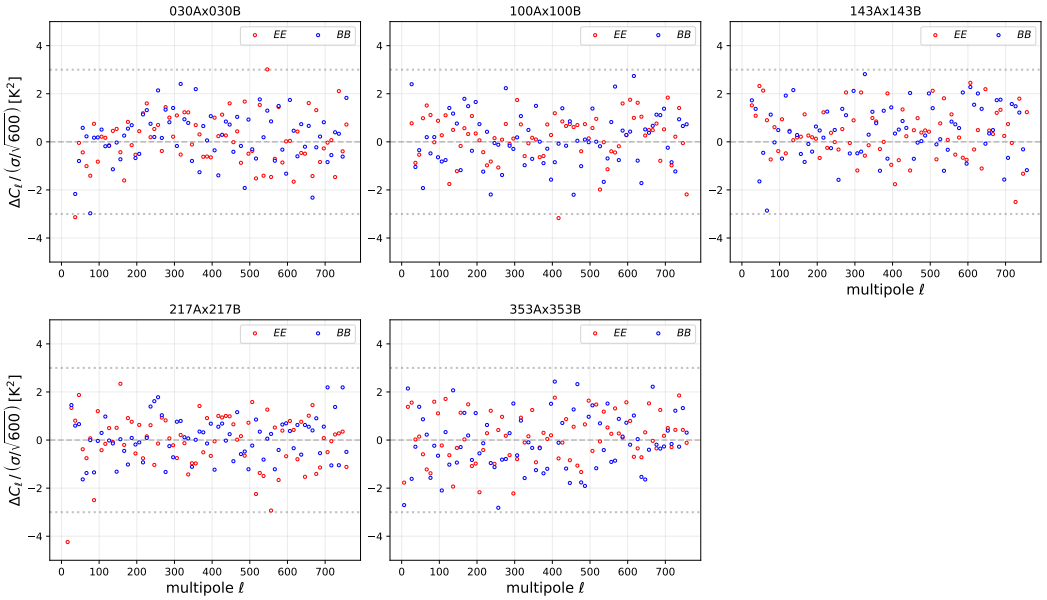


Figure 6.9. Same as figure 6.8, but here spectra are evaluated on the *SO* SATs sky patch of figure 6.3.

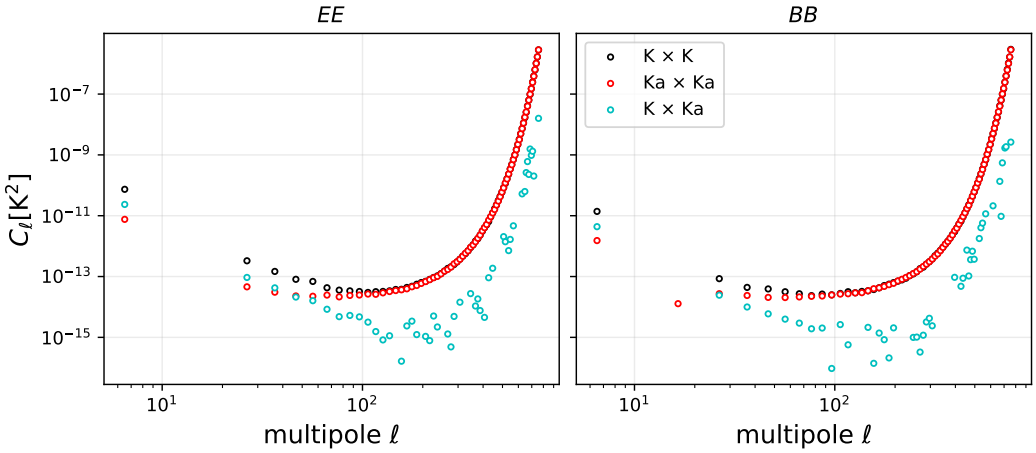


Figure 6.10. EE (*left*) and BB (*right)* auto- (black and red) and cross-frequency (cyan) angular power spectra of *WMAP* data, evaluated adopting the mask covering the Galactic plane, point sources, and other polarised structures (right panel of figure 6.5).

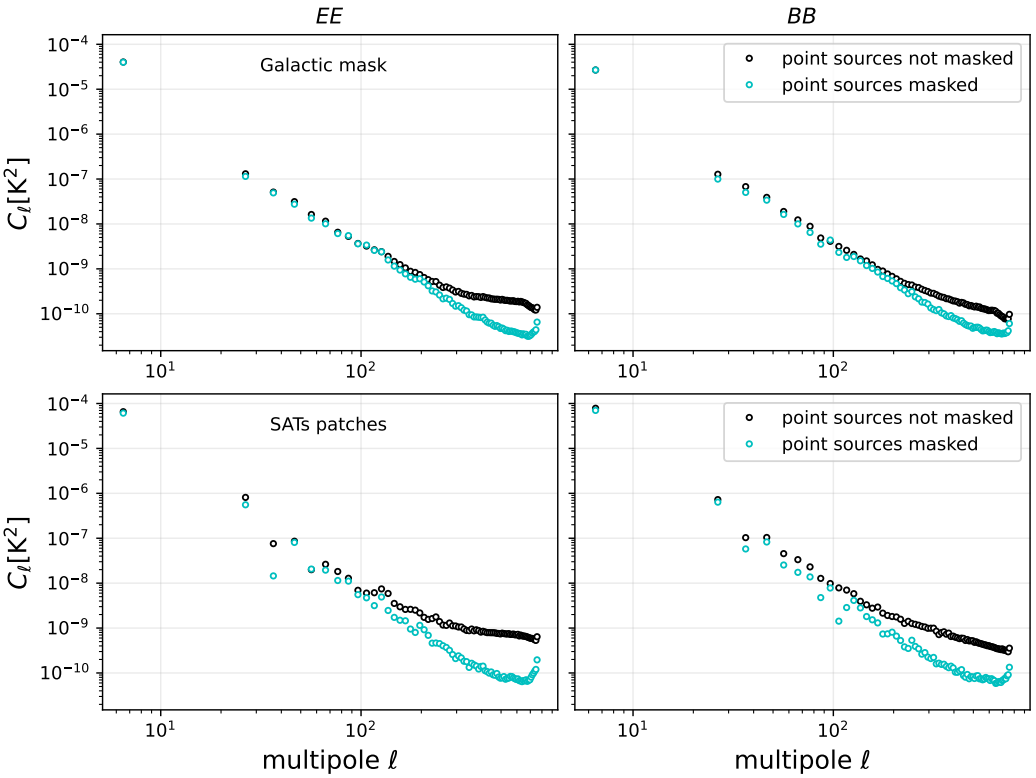


Figure 6.11. EE (*left column*) and BB (*right column*) auto-angular power spectra of *S-PASS* data. Spectra in top panels are evaluated on the mask covering the Galactic plane, while the ones in the lower panels are computed on the *SO* SATs sky patches. For black circles, point sources are included in the analysis, while for cyan circles point sources are masked out. The improvement in removing the flattening due to polarised point sources is evident.

***S-PASS* data** We have a single frequency channel for *S-PASS*, with no data splits and no simulations available. But these data have an extremely good signal-to-noise ratio at a frequency where synchrotron and polarised point sources are the dominant emissions. Figure 6.11 shows the auto-angular power spectra of *S-PASS* data for the different sky regions considered in this work. The power-law behaviour is clearly indicative of the synchrotron emission, and the spectral flattening due to presence of polarised point sources is also evident. The latter is largely mitigated by the application of a point sources mask, which restores the power-law shape of the spectra until the instrumental beam starts to be important at $\ell \simeq 700$. We also notice how the noise correlation bias is not really impactful when the considered data are signal-dominated. Power spectra computed on the SATs patches show to be noisier (more "wobbling"), as they are evaluated on a smaller sky fraction and on regions which are cleaner from foregrounds with respect to others.

Chapter 7

Simulations

Simulations are a crucial part of any experiments, both during design, construction, and operational phases. They can be used for forecasting purposes but also for the characterisation of the instruments, improving our knowledge of the telescopes. During the last few years, I participated to this effort in the *SO* collaboration, providing several sets of simulations to be used for different purposes. I am going to describe them in this chapter.

7.1 Covariance estimation

As we noted a couple of times before, we as human beings have only one sky and one Universe to look at. One of the consequences of this fact is the cosmic variance we described in section 3.4. But this also means that we have just one opportunity to use our data analysis techniques to extract information from the observed maps. In an ideal world, we would have several realisations of the target observable, and we would perform our analysis on each realisation and then extract the average outcome with a certain error associated to it. Instead, we do not have the possibility of quantifying this uncertainty on real data.

Moreover, power spectrum-based component separation methods (as the nominal algorithm for the *SO* SATs survey) rely on a likelihood analysis of the power spectra data vector [86, 15, 16, 101]. The computation of the likelihood requires an estimation of the power spectrum covariance matrix, containing correlations between different scales of the sky components. A purely Gaussian CMB signal would give rise to a diagonal covariance matrix, where all the scales are independent from one to another. This is obviously not the case for foregrounds, which are highly non-Gaussian containing anisotropic features and correlations between scales, translating into non-zero off-diagonal elements in the covariance matrix [2].

Thus, it is crucial to have a way to evaluate a set of power spectra on different realisations of the sky.

A viable solution is represented by simulations, which I tackled during my contribution time in the *SO*. Artificial skies can be generated by modelling the sky emission, while relying on some simplifying assumptions. We can then perform the data analysis on the simulations and associate an uncertainty to the information obtained from real data. In particular, if we want to attach an errorbar to the measurement of the tensor-to-scalar ratio, we should evaluate the variance linked to the analysis steps required for extracting the r value, as the component separation technique and the power spectra estimator. This can be achieved by running the same r inference pipeline on data and simulations, in the same patch of the sky, from maps to power spectra. A numerical covariance matrix can then be estimated, containing information on the pipeline correlations. The same methodology can be used to build the power spectrum covariance matrix, by calculating power spectra in the same way both for the data and the simulations.

These covariance matrix simulations should be representative of the sky signal contained in real data, namely CMB and foregrounds emission. Fortunately, there is no need for the simulated foregrounds to be as complex as the ones populating the real sky: their amplitude in the *SO* SATs region needs to be close to the real one, but their spatial variability can be much simpler. As shown in [2], Gaussian realisations are enough for a proper estimation of the covariance; the authors have shown that this assumption has negligible impact on the recovered cosmological parameters, even when off-diagonal elements due to foregrounds non-Gaussianity are considered.

The production of simulations for the computation of the covariance matrix can be separated in two steps: first, foregrounds templates based on real data (synchrotron and thermal dust) from `PySM` are used to extract the amplitude and slope parameters of the power-law governing their behaviour in harmonic space; 2) then, multi-frequency Gaussian simulations are generated from CMB and power-law power spectra, using the fitted parameters. In more details:

1. `PySM` temperature and polarisation templates are adopted, at the original pixel resolution of $N_{\text{side}} = 2048$, for synchrotron (pivot frequency $\nu_0^{\text{s}} = 23$ GHz) and thermal dust (pivot frequency $\nu_0^{\text{d}} = 353$ GHz), with the `d9s4` complexity (see section 4.4 for details). Templates are processed into a SAT-like configuration exactly as described in chapter 6. They are downgraded to a lower pixel resolution ($N_{\text{side}} = 512$) and rotated from galactic to equatorial coordinates, carrying out both operations in harmonic space. For each template, TT, EE, BB power spectra on the SATs sky patch (figure 6.3) are computed with the

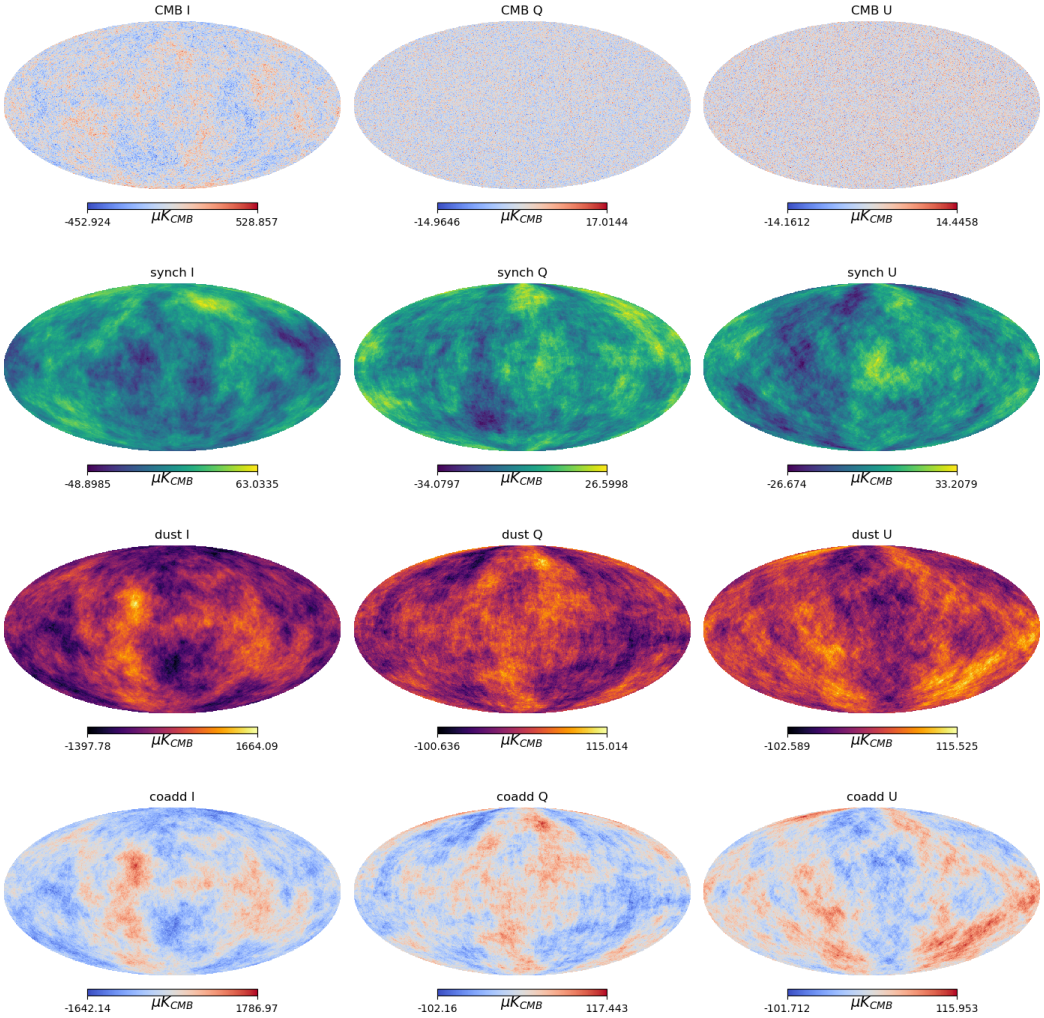


Figure 7.1. Example of sky maps simulated for the purpose of covariance matrix estimation, at a frequency of 353 GHz. *From left to right:* total intensity I, and Stokes Q and U maps. *From top to bottom:* CMB, synchrotron radiation, dust emission, and all the above co-added maps. At the selected frequency, thermal dust is the dominant component. Units are $\mu\mathrm{K}_{\mathrm{CMB}}$.

`NaMaster` package [12]. Then, power-law parameters (amplitudes and slopes) are extracted by fitting the spectra as

$$C_\ell^{XX} = A_{\rm c}^{XX} \left(\frac{\ell}{\ell_0}\right)^{\alpha_{\rm c}^{XX}} , \tag{7.1}$$

where $\mathrm{c} \in (\mathrm{s}, \mathrm{d})$ identifies the component (synchrotron or dust), $\ell_0 = 80$ is the pivot multipole, and $X \in (T, E, B)$.

2. $N = 500$ sets of Gaussianly distributed spherical harmonics coefficients are generated from the power spectra of each component (CMB, synchrotron and dust). CMB spectra are based on the cosmological parameters derived by *Planck* [166], while foregrounds spectra are computed as power laws where the spectral parameters are the ones fitted in the previous stage. In order to generate multi-frequency realisations, each $a_{\ell m}$ is frequency-scaled with the SED ($S_{\rm c}^\nu$) of the corresponding component, at frequencies of 23, 93, 145, 353 GHz, covering the foregrounds-dominated *WMAP* and *Planck* channels, and the nominal SATs CMB bands. Synchrotron radiation is rescaled with a simple power-law with index $\beta_{\rm s} = -3$, while thermal dust is rescaled adopting a modified blackbody emission with index $\beta_{\rm d} = 1.59$ and temperature $T_{\rm d} = 19.6\,\mathrm{K}$ (see section 4.2 for details):

$$\begin{aligned} S_{\rm s}^\nu &= \left(\frac{\nu}{\nu_0}\right)^{\beta_s} , \\ S_{\rm d}^\nu &= \left(\frac{\nu}{\nu_0}\right)^{\beta_d} \frac{B(\nu, T_{\rm d})}{B(\nu_0, T_{\rm d})} . \end{aligned} \tag{7.2}$$

The co-added spherical harmonics coefficients representing the total sky signal are then obtained as

$$a_{\ell m}^\nu = \sum_{\rm c} S_{\rm c}^\nu a_{\ell m}^{\rm c} = a_{\ell m}^{\rm CMB} + S_{\rm s}^\nu a_{\ell m}^{\rm s} + S_{\rm d}^\nu a_{\ell m}^{\rm d} . \tag{7.3}$$

The final spherical harmonics coefficients are then converted to 500 sets of temperature and Stokes Q and U at a resolution of $N_{\rm side} = 512$, from which the covariance matrices for different purposes can be estimated. Figure 7.1 illustrates an example of these simulations at the single dust-dominated frequency of 353 GHz, for each generated component and their sum. We notice that in these maps there is no structure associated to the presence of the Galactic plane, since these are Gaussian realisations. The simulated fields are homogeneous and isotropic. They are made for the estimation of a covariance matrix for the *SO* SATs, which observe a small sky region far from the Galactic plane. Thus, there is no reason to properly simulate

it.

7.2 Transfer function estimation

Ground-based observatories have to deal with multiple sources of low-frequency noise that can affect measurements of the primordial B-mode signal. A non exhaustive list of these factors includes: temporal and spatial fluctuations in the emission of the atmosphere due to turbulence and uneven distribution of water vapour, creating non-trivial correlations between detectors [72]; intensity-to-polarisation (I-to-P) leakage due to thermal drifts, optical mismatch, and other instrumental effects; bandpass or beam mismatches [25].

To address these difficulties, the *SO* SATs modulate the polarisation signal at frequencies where the detectors white noise dominates, above atmospheric fluctuations and thermal drifts, using a cryogenic half-wave plate (HWP) rotating with a frequency around $f \approx 2$ Hz [117, 200]. Additional applications of low- and high-pass filters to the polarised timestreams are part of standard procedures adopted to suppress low-frequency $1/f$ noise and recover the polarised signal [180]. The impact of all this filtering operations must be accounted for when generating maps from the filtered timestreams, or alternatively, the associated map-level impact must be taken into account and corrected for when analysing them [117, 174, 5].

Generally speaking then, timestreams registered by a ground-based telescope contain low-frequency contamination that is tackled through the application of some filters. These filters induce biases in the derived sky maps, which usually take the form of a general loss of power, more pronounced at the largest scales, as well as the presence of mixing and statistical coupling between E- and B-modes (E/B leakage) [149, 101]. This obviously has a huge impact on the recovery of cosmological parameters, and filtering must be carefully taken into account. Different methods exist to tackle this problem. One of the most straightforward, and the one adopted at the moment in the *SO* collaboration and on which we focus, is the transfer function method [117, 101], which allows to approximately estimate and correct for the filtering impact at the angular power spectrum level. Exact methods do exist, as the "per-ℓ" method [4, 24], which reconstructs the filtering effects on the full range of multipoles. This technique is much more computationally expensive, becoming prohibitive for the forecasted large data volume of *SO*.

In its simplest approximation, the power spectrum transfer function is assumed to be a multipole-dependent rescaling factor $T(\ell)$ between the power spectra of

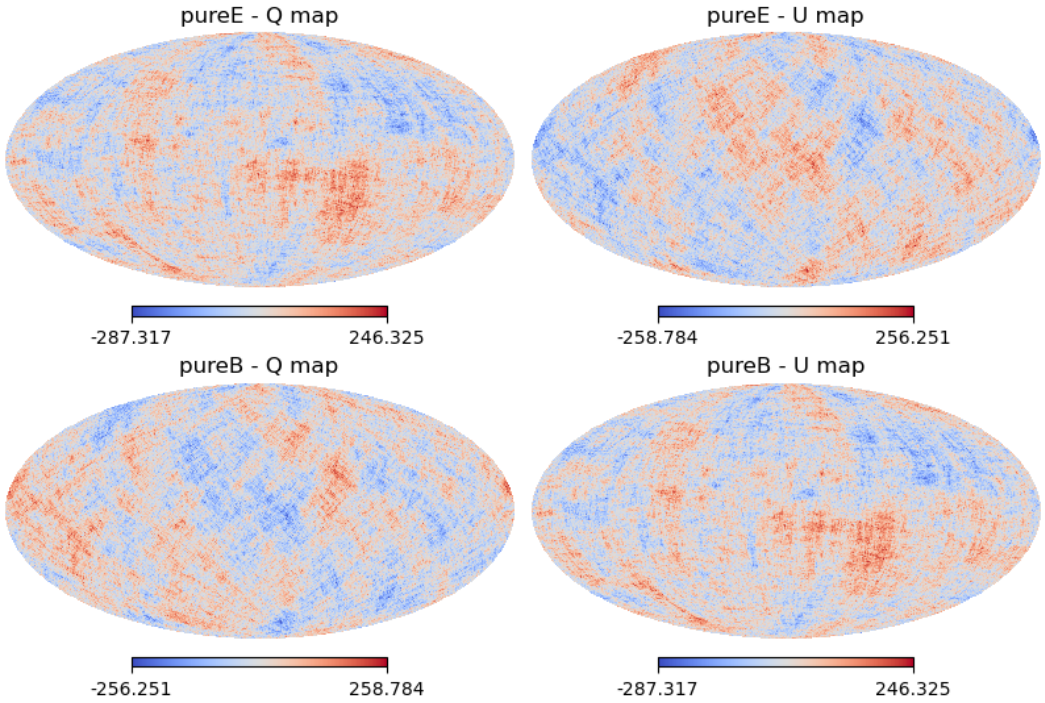


Figure 7.2. An example of Q and U maps containing E- and B-modes only. The characteristic $+$ and $\times$ patterns in the Q_E and Q_B maps indicate the dominance of E- and B-modes, respectively.

filtered and unfiltered scalar maps (e.g., temperature-only):

$$C_\ell^{\text{filter}} = T(\ell) C_\ell^{\text{unfilter}} \,. \tag{7.4}$$

In the case of spin-2 Q and U polarisation maps, the transfer function assumes the form of a 4×4 matrix, containing the suppression and leakage terms in EE, EB, BE, BB. Through the use of simulations, the transfer function can be estimated by computing the power spectra of simulated maps before and after filtering, and taking their ratio. In this way we can empirically quantify the effect of filtering on real data, assuming that its impact is always the same. We are not intrinsically interested in the particular simulated field; the only need is for it to have spectral features similar to the sky component targeted by the telescope, as the CMB and Galactic foregrounds. Thus, simple and computationally cheap Gaussian simulations are enough to do the job (see [103, 101] and reference therein for more details on the transfer function formalism).

In particular, I generated $N = 100$ sets of spherical harmonics coefficients $a_{\ell m}$, starting from an input power-law spectrum of the form

$$C_\ell = A \left(\frac{\ell}{\ell_0} \right)^\alpha \,, \tag{7.5}$$

with $\alpha = -2$, where values $\alpha \in [-3, 0]$ cover the typical steepness of CMB and foregrounds power spectra. Gaussian simulations are the generated by drawing the m values as Gaussianly distributed random numbers, with zero mean and variance given by the power spectrum (see section 3.4 for details):

$$\langle a_{\ell m} \rangle = 0 \,, \qquad \langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell \,. \tag{7.6}$$

In the above equation, A is the arbitrary amplitude, α is the slope, and $\ell_0 = 80$ is a pivot multipole chosen to be close to the recombination bump. From these generated E- and B-modes coefficients ($a^E_{\ell m}$ and $a^B_{\ell m}$), two sets of Q and U polarisation maps are built, one containing E-modes only (`pureE`) and the other containing B-modes only (`pureB`):

$$\begin{aligned} Q_E + iU_E &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=\ell} a^E_{\ell m} \, {}_{\pm 2}Y_{\ell m} \quad (\texttt{pureE}) \,, \\ Q_B + iU_B &= \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{m=\ell} a^B_{\ell m} \, {}_{\pm 2}Y_{\ell m} \quad (\texttt{pureB}) \,. \end{aligned} \tag{7.7}$$

With these particular sets of maps, we are able to properly characterise the transfer function, since by construction the Q_E/U_E fields have no BB component, and the Q_B/U_B have no EE component. The leakage of both E-to-B and B-to-E induced by filtering can be quantified, together with the large-scale power suppression. Figure 7.2 shows an example of such simulated maps, while figure 7.3 illustrate an example of a polarisation transfer function computed for a filtering pipeline similar to the one forecasted for the *SO* SATs surveys, with power spectra evaluated over the SATs sky mask of figure 6.3, over 100 simulations of maps at a frequency of 90 GHz. The auto-terms $XX \to XX$ on the diagonal show the suppression at larger scales due to filtering, followed by the plateau. The off-diagonal terms quantify the statistical coupling between E- and B-modes; for example, the $EE \to BB$ panel represent the EE power leaking into the BB spectra. The estimated transfer function can then be used to correct for filtering effects at the angular power spectrum level, allowing the extraction of cosmological information.

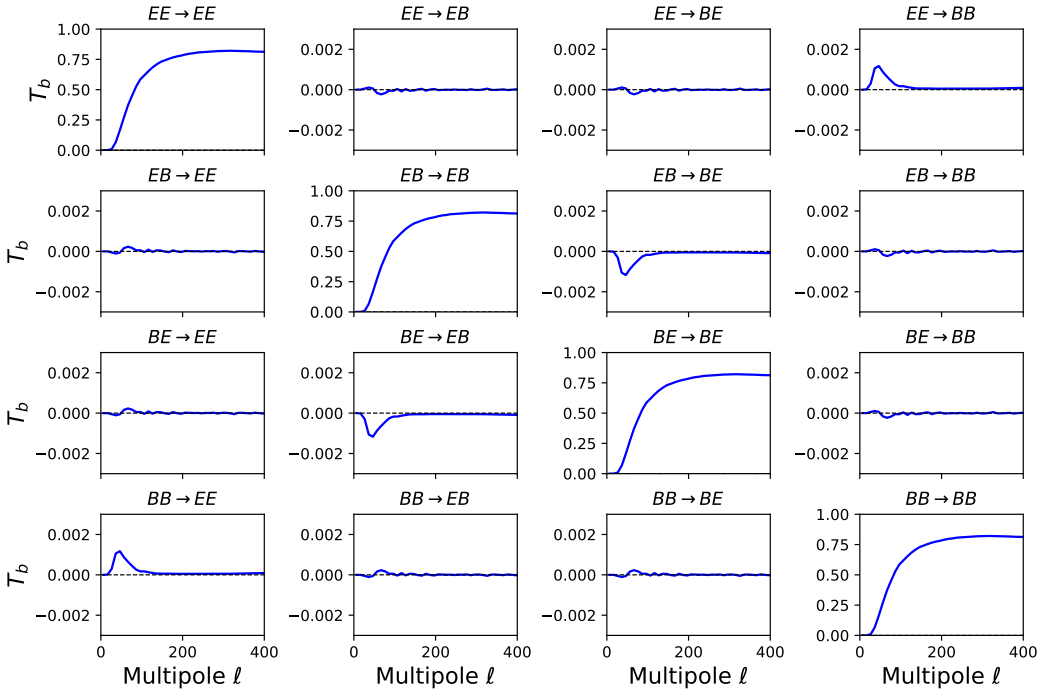


Figure 7.3. An example of transfer function for a 90 GHz SATs-like channel, showing the 16 $X'Y' \to XY$ polarisations, averaged over $N = 100$ realisations of the power-law simulations shown in figure 7.2. The (very thin) shaded area represents the standard deviation divided by $\sqrt{N}$. This transfer function is estimated using more aggressive filtering than what expected in the real *SO* data analysis. Figure from [101].

Part III

The *B*-sure project: robustness tests for primordial *B*-modes

Chapter 8

The need for robustness tests in CMB science

As we already discussed multiple times in this manuscript, the eventual measurement of a B-mode signal in CMB polarisation is one of the major target of the CMB community, representing a target of paramount importance as it is believed to be induced by the presence of a stochastic background of primordial gravitational waves generated by inflationary physics happening in the infant Universe. The detection of the swirly pattern in CMB maps would then open an observational window into the very first moments of our Universe after the Big Bang, granting us access to energy scales unreachable by particle accelerators.

In Part I, I briefly introduced the physical concepts connecting inflation with CMB B-modes. In Part II, I described the *Simons Observatory* (SO), a brand new microwave experiment with the goal (among others) of placing stringent constraints on the amplitude of primordial B-modes. We also discussed the importance of combining data between different experiments to try and mitigate the contamination from Galactic foregrounds, which heavily hinder the hunt for the underlying primordial signal. The hope of the community is for current and future observatories to finally detect this faint signal to dive into new, non-standard physics.

But once we start to approach this goal, crucial questions are going to arise: how can we be certain that the measured signal is really due to primordial gravitational waves? In other words, are we able to recognise if an eventual detection is contaminated by instrumental systematics or astrophysical foregrounds? Is the detection robust against contaminants, and are we able to reject it?

In order to confidently answers these questions, and thus be able to accept or reject a potential measurement, we need to develop some robustness tests apt at the validation of a tensor-to-scalar ratio detection from CMB observations. At the moment, this particular field has not been extensively investigated, since a

detection of primordial B-modes was not an imminent event. During the years, instrumental systematics, Galactic foregrounds, and gravitational lensing have proven to be formidable challenges in the search of the much fainter primordial signal, with experiments only being able to place upper limits on the amplitude of tensor perturbations.

In this Part, I will describe the work I carried out in the last two years on the topic of robustness tests for primordial B-modes. In particular, I focused on the contamination represented by Galactic foregrounds and on which statistical tools can be used to identify a false detection of the tensor-to-scalar ratio. I explored Minkowski functionals (MFs) and scattering transforms (ST) statistics for this scope. The former are well-known tools in the context of CMB data analysis, being used for more than 20 years now, while the latter are relatively novel in the field. The majority of what is discussed in the remainder of this Part is based on two of the published works I produced during the PhD, [178] and [179], describing the B-sure project.

8.1 The *BICEP* accident

In 2014, the *BICEP2* collaboration published results coming from three seasons of deep observations of relatively small sky patches from the South Pole [4]. They reported the detection of a B-mode signal at a degree angular scale, attributed to the presence of an inflationary gravitational waves background with a tensor-to-scalar ratio of $r = 0.20^{+0.07}_{-0.05}$, representing a $> 5\sigma$ excess with respect to the standard ΛCDM scenario where $r = 0$. The theoretical implications suggested by this measurement would be of extreme proportions. Of course, a detection of r would be a smoking gun for the inflationary mechanism, but it is worth noticing that such a (unexpected) large value would also be difficult to accommodate within theories of high energy physics motivated by other experimental data. In the weeks and months following the *BICEP* publication, online archives were flooded with papers from the communities of particle physics and theoretical cosmology, in the tentative to explain such a result with the most exotic scenarios.

Unfortunately, a subsequent re-analysis of the same *BICEP* data combined with *Planck* maps [27] revealed the claim to be a false detection, with the measured B-mode signal heavily induced by the polarised emission of Galactic dust. At the time, the foregrounds contamination to polarised observations of the CMB was severely underestimated, mainly because of the lack of sensitive data and consequently adequate modelling. Using the new maps produced by the *Planck* experiment, the community was able to assess the amount of contamination in the context of B-modes measurements. Foregrounds emission is now considered as the primary

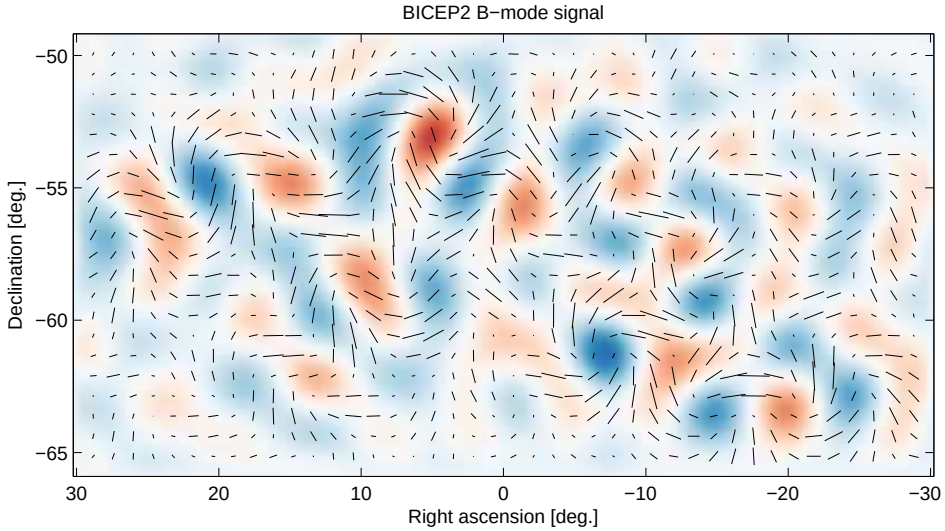


Figure 8.1. B-mode signal reported by the *BICEP2* collaboration in 2014, with the black rods indicating the polarisation amplitude in different spots of the sky region observed by the telescope. Figure taken from the *BICEP2* 2014 Release Image Gallery, realised using data from [4].

challenge for the estimation of the tensor-to-scalar ratio, with new experiments and data analysis techniques designed to mitigate contamination at the best possible level.

Nonetheless, we are not able to completely clean our maps of the sky: component separation algorithms will always leave a certain amount of foregrounds and noise residuals in the final CMB product [40, 41]. Depending on the amplitude of the signal targeted by the analysis, this residual contamination can be negligible or not. In the case of B-modes searches, the signal is so faint that a low amount of residuals is still comparable with the amplitude of tensor perturbations, possibly leading to a biased measurement of r and a consequent false detection. This is similar to what happened in the *BICEP* accident. Of course we can give an estimate of this bias and validate component separation techniques with sky simulations. Still, the exact complexity of the foregrounds emission is not replicable, and the real sky observed by an experiment will thus be unknown. Moreover, having a single Universe does not allow us to estimate the sky signal from a statistical sample: if component separation does not work properly for our particular realisation of the sky (maybe coming from the tail of the distribution), we have no way to know it.

The *BICEP* case is just one example of the more general need for robustness tests in Cosmology, and in particular for the tests tailored to the detection of primordial B-modes from CMB observations. In this context, we denote as robustness tests all those "a posteriori" analysis that can be done on data in order to validate

a potential detection of the tensor-to-scalar ratio. An ideal robustness test is able to reject a r measurement if the estimate is biased (test failed) or to accept it if it is not biased (test passed). In the remainder of this Part, we will focus on the construction of such robustness tests for the measurement of CMB B-modes, with a particular attention to false detections induced by foregrounds contamination. We will also focus on B-modes measurements achieved through the use of map-based component separation techniques, which aim to recover a cleaned estimate of the CMB field. We will not treat robustness tests for power spectrum-based algorithms, whose investigation is left for future works.

Chapter 9

How to build a robustness test

This chapter reports the methodology for the B-sure project, namely the development and characterisation of the robustness tests described in two of the published works associated with this Thesis, [178] and [179]. The procedure tries to closely replicate what would happen in a real experiment targeting the observation of CMB polarisation, where multi-frequency sky maps are combined in a map-based component separation algorithm to recover the CMB signal. Its BB angular power spectrum is then used to measure a tensor-to-scalar ratio, eventually biased by foreground residuals. The goal of the test is to validate this detection, rejecting it if contaminated.

In the previous Part, we focused our attention on a ground-based experiment, the *SO*. In this Part instead, we will work with a telescope mounted on a satellite, observing from space. The latter configuration provides several simplifications to the analysis: the full-sky is now available, while it is just a small fraction for a telescope on Earth; the angular power spectrum can be also estimated at the largest angular scales, which can be accessed from space; atmospheric emission is not a problem any more, as we do not need to worry for the consequent filtering process and for the estimation of a filtering transfer function.

The core idea behind the tests is to look for residual foregrounds contamination in the CMB map recovered by the component separation algorithm. In particular, the proposed tests aim to exploit the key morphological differences between CMB and foregrounds signals. As we mentioned in section 2.3, the theory of inflation predicts the CMB to be a Gaussian random field [18], with primordial perturbations drawn from a Gaussian distribution with zero mean and variance given by the power spectrum. This prediction agrees very well with observational data. In simpler terms, Gaussianity means that if we look at different regions of the sky, the measured anisotropies will statistically be the same. On the other hand, foregrounds emission is observed to be highly non-Gaussian, with anisotropic structures all over the sky

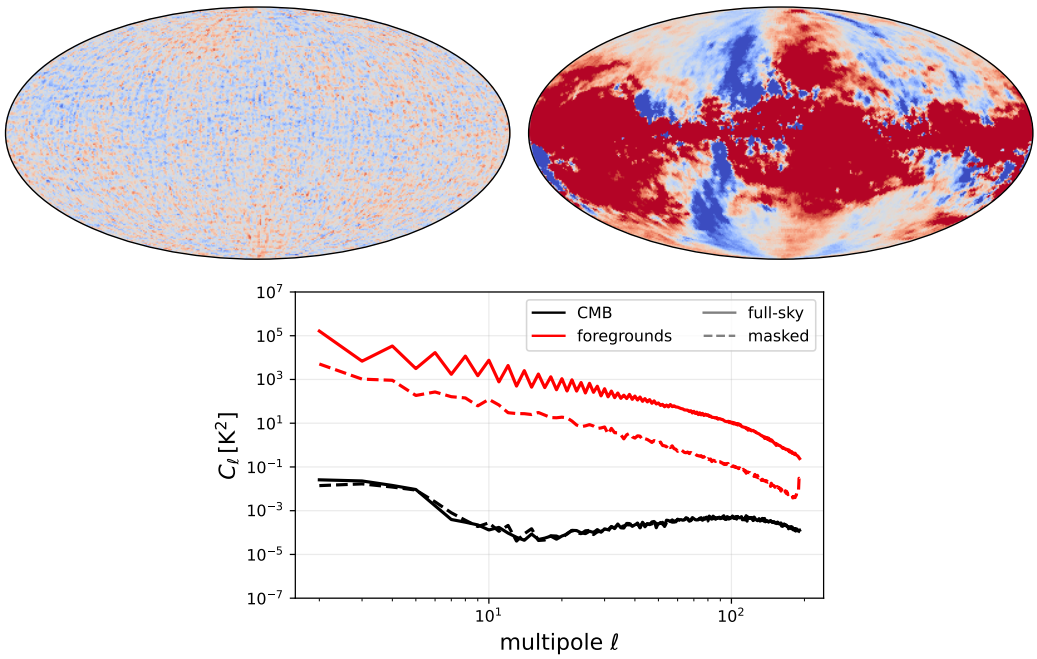


Figure 9.1. *Top panels:* Gaussian CMB field on the left, foregrounds map on the right, both in Stokes parameter Q. Visually, they clearly have a different morphology. *Bottom panel:* EE angular power spectrum of the above maps, over the full-sky or adopting a mask covering the Galactic plane. For a Gaussian field, the statistics is the same, when corrected for the sky fraction.

tracing the local Universe. If we look at different patches, the measured emission will in general have different statistical properties [114, 169]. Figure 9.1 illustrates such difference, which is clearly visible at the map level but can also be seen in the power spectra: evaluating the statistics over a full or masked sky does not impact the CMB spectrum (when corrected for the partial sky), while it is different for foregrounds.

Thus, in order to build such a robustness test, we need two key ingredients: 1) a set of simulated Gaussian maps, to adopt them as references for comparison, and 2) some summary statistics sensitive to the presence of non-Gaussianity and anisotropic structures in sky maps, able to identify morphological deviations from a Gaussian random field. The first item is necessary to have some reference Gaussian fields from which we can quantify the deviation. To have a fair comparison, the signal amplitude in these simulations should be similar to the one of component-separated CMB map. Concerning the second item, several statistical descriptors of non-Gaussianity have already been employed in CMB data analysis [52, 66, 83, 161, 165, 171]; in this Thesis, we explored Minkowski functionals (MFs) and scattering transforms (ST), adopted as summary statistics for the robustness test. If the MFs or the ST of the component-separated CMB map are compatible with the ones evaluated on reference Gaussian maps, the test is passed and the detection

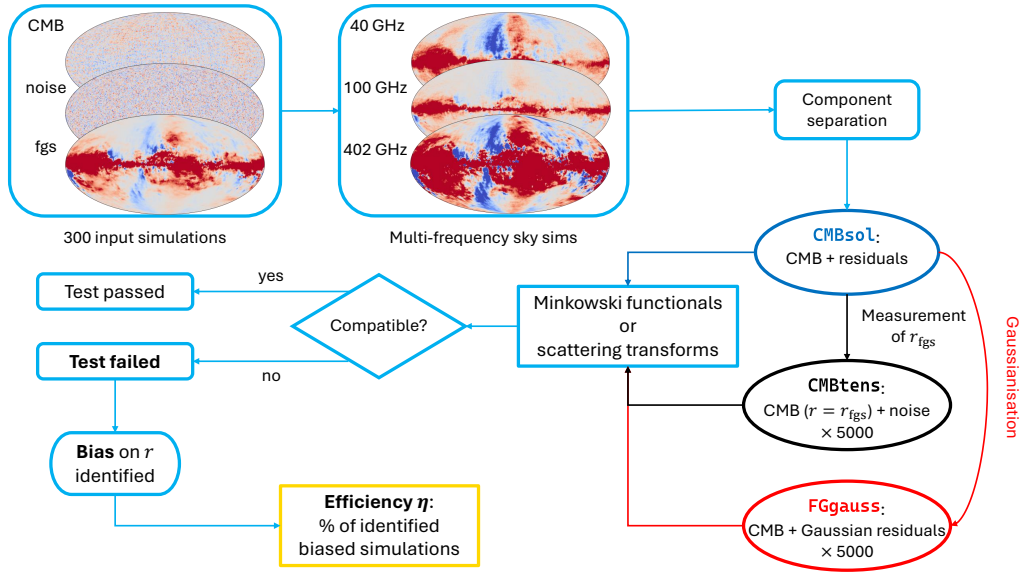


Figure 9.2. Flowchart illustrating the robustness test methodology and showing the various steps followed for the test characterisation and validation.

of the tensor-to-scalar ratio is accepted. Vice versa, if the summary statistics are not compatible, the test is failed and the detection is rejected. To properly assess if these statistics can be trusted, the robustness test is repeated and characterised over hundreds of simulations. Figure 9.2 visually illustrates the process described above; each step of the pipeline is detailed in the following dedicated sections.

9.1 Input sky simulations

In this section, we describe the sky simulations used as the starting point for the robustness test characterisation. Throughout this Part, maps generation, manipulation, and analysis is carried out adopting the `HEALPix`[1] [88] pixelisation scheme by means of the `healpy`[2] package [223]. For all our simulations, the pixel resolution is defined by $N_{\rm side} = 64$, corresponding to an angular resolution of about 55 arcmin, with the maximum multipole considered being $\ell_{\rm max} = 3N_{\rm side} - 1 = 191$.

We generate simulations of the polarised microwave full sky in Q and U Stokes parameters, as observed by a *LiteBIRD*-like experiment. *LiteBIRD*[3] [127] is a future space mission selected by the Japan Aerospace Exploration Agency (JAXA) in 2019 as a strategic large-class (L-class) mission. The scope of *LiteBIRD* is to produce

[1] https://healpix.jpl.nasa.gov/
[2] https://healpy.readthedocs.io/en/latest/
[3] Lite (Light) satellite for the study of B-mode polarisation and Inflation from cosmic background Radiation Detection, https://www.isas.jaxa.jp/en/missions/spacecraft/future/litebird.html.

maps of the microwave sky with three years of observations; among other scientific goals, its primary target is the B-mode signal from cosmic inflation, either making a discovery or ruling out well-motivated inflationary models. The instrumental properties are taken from the 2022 forecast of the experiment [127], which planned to deploy 3 telescopes covering 15 different frequencies between 40 and 402 GHz. The *LiteBIRD* mission has recently undergone a phase of reformation with the scope of simplifying the payload configuration while preserving the scientific objectives. In its updated version, published during the writing of this Thesis, the satellite will be launched around 2036, and it will host a single telescope observing the sky in 12 distinct frequencies between 40 and 402 GHz, with a polarisation sensitivity of $2.5\,\mu\mathrm{K_{CMB}}\cdot\mathrm{arcmin}$. The targeted total uncertainty on the tensor-to-scalar ratio is $\sigma(r) < 0.002$ (68% C.L.), assuming the absence of a primordial signal $r = 0$. For a fiducial value of $r = 0.01$, the targeted sensitivity enables independent 3σ detections at both the reionisation and recombination bumps. In this way, the experiment will be able to test representative inflationary models, including Starobinsky-like and Higgs-inflation scenarios, at meaningful statistical significance, probing the origin of primordial gravitational waves [37]. Therefore, the *LiteBIRD* configuration is the most useful to use for our input simulations, in order to characterise the robustness test, given its forecasted performances in constraining r. Table 9.1 reports the specifics adopted for generating the sky maps. The works described in this manuscript were already completed before the *LiteBIRD* reformation; thus, the frequencies, angular resolutions, and noise sensitivities listed in the table are the one from the previous configuration of the experiment, described in [127].

The simulated maps contain three polarised sky components, namely CMB, instrumental noise, and Galactic foregrounds. CMB maps are generated as Gaussian realisations from the angular power spectra of the best-fitting *Planck* 2018 parameters [166]. Realisations are obtained with the `synfast` routine of the `healpy` package, which draws $a_{\ell m}$ as random numbers from a Gaussian distribution having a variance equal to the power spectra. We produce two sets of $N = 300$ simulations. In one set, B-modes are sourced only by gravitational lensing, where the large-scale structure of the Universe deflect CMB photons and convert E-modes into B-modes, and no primordial signal is present ($r_{\mathrm{in}} = 0$). In the other set, an additional realistic amplitude of primordial B-modes is injected, sourced by inflationary gravitational waves ($r_{\mathrm{in}} = 5\times10^{-3}$). Figure 9.3 shows the input power spectra used for generating the Gaussian realisations, both in C_ℓ and D_ℓ. All CMB maps are smoothed to a common angular resolution of $\mathrm{FWHM_{out}} = 70.5\,\mathrm{arcmin}$, which is the largest beam between the considered frequency channels (corresponding to the 40 GHz band). We notice that technically this angular resolution is not sufficiently large compared

Table 9.1. *LiteBIRD* specifics used to generate the simulations, with values related to 15 observational bands (15 frequencies, 22 effective channels). From left to right the columns are: telescope covering the band (Low-Frequency Telescope LFT, Medium-Frequency Telescope MFT, High-Frequency Telescope HFT); band identification number (ID); band centre frequency in GHz; main beam size (FWHM) in arcmin; sensitivity in $\mu\mathrm{K_{CMB}} \cdot \mathrm{arcmin}$.

Telescope	ID	ν [GHz]	FWHM [arcmin]	sensitivity [$\mu\mathrm{K_{CMB}} \cdot \mathrm{arcmin}$]
LFT	1	40	70.5	37.42
LFT	2	50	58.5	33.46
LFT	3	60	51.1	21.31
LFT	4	68	41.6	19.91
LFT			47.1	31.77
LFT	5	78	36.9	15.55
LFT			43.8	19.13
LFT	6	89	33.0	12.28
LFT			41.5	28.77
LFT	7	100	30.2	10.34
MFT			37.8	8.48
LFT	8	119	26.3	7.69
MFT			33.6	5.70
LFT	9	140	23.7	7.25
MFT			30.8	6.38
MFT	10	166	28.9	5.57
MFT	11	195	28.0	7.05
HFT			28.6	10.50
HFT	12	235	24.7	10.79
HFT	13	280	22.5	13.80
HFT	14	337	20.9	21.95
HFT	15	402	17.9	47.45

to the pixel size at the adopted N_{side} (55 arcmin). In general, the angular resolution should be at least three times larger than the pixel resolution, while here is instead only 1.5 larger. However, we found this effect to have negligible impact, which in any case is impacting the smaller scales only. In this work we will mainly focus our attention to the larger scales, which are the most contaminated by foregrounds emission.

We also notice here that the CMB field can contain a certain level of non-Gaussianity induced by the gravitational lensing operated by large-scale structures. It is totally possible to simulate non-Gaussian CMB maps and run the same robustness tests on them. However, since the proposed robustness tests target the identification of non-Gaussianity, the presence of realistic lensing can create some ambiguity with the non-Gaussianity induced by Galactic foregrounds, which is the

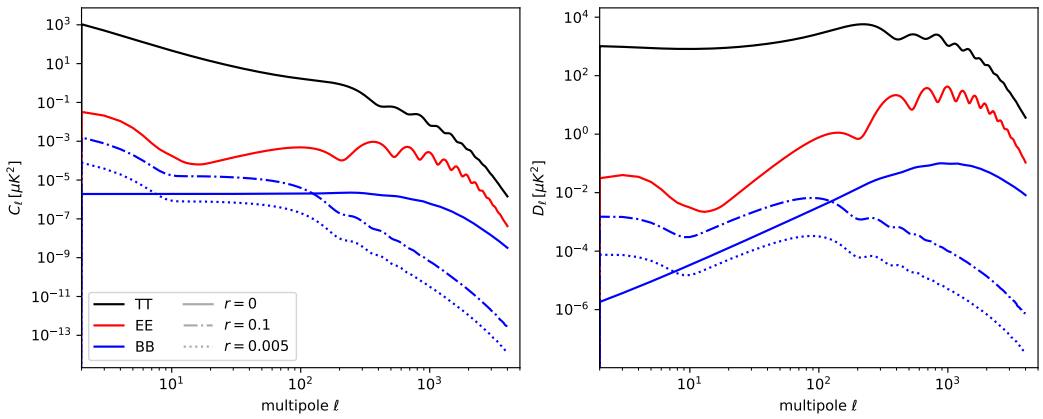


Figure 9.3. Input angular power spectra used to generate the Gaussian CMB simulations, in C_ℓ (*left*) and D_ℓ (*right*), adopting the best-fitting *Planck* 2018 parameters [166]. TT in black, EE in red, BB in blue. Solid lines represents spectra where B-modes arise from lensing only ($r_{\rm in} = 0$), while the dashed line show the contribution due to primordial gravitational waves for $r_{\rm in} = 5 \times 10^{-3}$.

main topic of the study. Moreover, the summary statistics employed in this work have demonstrated to not be sensitive enough to detect the non-Gaussian impact of gravitational lensing, which is anyway expected to be important at smaller scales than the ones considered in our analysis. We thus decide to work with Gaussian realisations of CMB lensing and leave the small non-Gaussian effects for future works and more sensitive statistics.

Noise maps are simulated by generating Gaussian realisations for each frequency channel, starting from the following noise power spectrum:

$$\begin{aligned} N_\ell &= \left(\sigma_{QU}\frac{\pi}{180\cdot 60}\right)^2 \left[1+\left(\frac{\ell}{\ell_{\rm knee}}\right)^{\alpha_{\rm knee}}\right] \\ &\simeq \left(\sigma_{QU}\frac{\pi}{180\cdot 60}\right)^2 , \end{aligned} \tag{9.1}$$

where σ_{QU} is the polarisation sensitivity in $\mu{\rm K}_{\rm CMB}\cdot{\rm arcmin}$, reported in table 9.1 and then converted to radians. $\alpha_{\rm knee}$ and $\ell_{\rm knee}$ are the two parameters defining the instrumental $1/f$ noise. In the second line of the above equation, we do not include this $1/f$ term in the noise power spectrum, as in the *LiteBIRD* experiment its contribution is expected to be fully suppressed by the modulation of the half-wave plate. We also neglect here the scanning strategy of the telescope, thus assuming the noise statistical properties to be constant across the sky. As discussed in [127], this is a fair assumption for *LiteBIRD* since the scanning strategy parameters are chosen to ensure a hit counts distribution as uniform as possible, and they are expected to have a negligible impact in Q and U maps. The sensitivity levels adopted

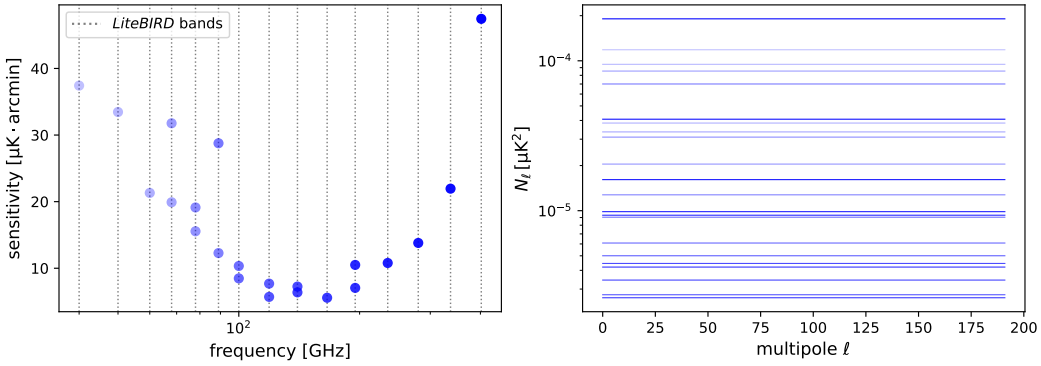


Figure 9.4. Visualisation of the noise properties of the *LiteBIRD* configuration adopted in this work. *Left*: sensitivity levels used for generating the simulations, in $\mu\mathrm{K}_{\mathrm{CMB}} \cdot \mathrm{arcmin}$ units, with frequency bands denoted by vertical lines. They are also reported in table 9.1. *Right*: corresponding white noise power spectra, obtained from eq. (9.1).

and the relative noise spectra are also visualised in figure 9.4, for each considered band. The generated noise maps at each frequency channel are smoothed to a common resolution by considering an effective beam as

$$\mathrm{FWHM}_{\mathrm{eff}}(\nu) = \sqrt{\mathrm{FWHM}_{\mathrm{out}}^2 - \mathrm{FWHM}(\nu)^2}\,, \tag{9.2}$$

where $\mathrm{FWHM}_{\mathrm{out}}$ is again the beam of the lowest frequency channel, at 40 GHz.

Galactic foregrounds are simulated with the Python Sky Model (`PySM`)[4] [204, 224, 203], a `python` package that generates full-sky simulations of Galactic foregrounds in intensity and polarisation relevant for CMB experiments, already introduced in section 4.4. Since we are focusing on polarised observations, we are only interested in thermal dust (`d`) and synchrotron (`s`) emission, as they are the main contaminants of observations. In this work, we adopt three different models for synchrotron and thermal dust emission, of increasing complexity: `d0s0`, `d1s1`, and `d10s5`. The former is a simplified view of the sky that we use for testing and validation of our pipeline, where dust is modelled as a modified blackbody, synchrotron is modelled as a power law, and dust temperature and spectral indexes are fixed across the sky. `d1s1` is a slightly more complex model than `d0s0`, where dust temperature and spectral indexes are spatially varying. The `d10s5` is the model closer to real observations, and it is the most complex foregrounds emission we consider here. We refer to the `PySM` papers and documentation for more details and full descriptions of the foregrounds templates.

We generate Q and U full-sky simulations for each of the three models `dxsy`, one per frequency channel; each map is smoothed to $\mathrm{FWHM}_{\mathrm{out}} = 70.5$ arcmin and

[4] https://pysm3.readthedocs.io/en/latest/index.html

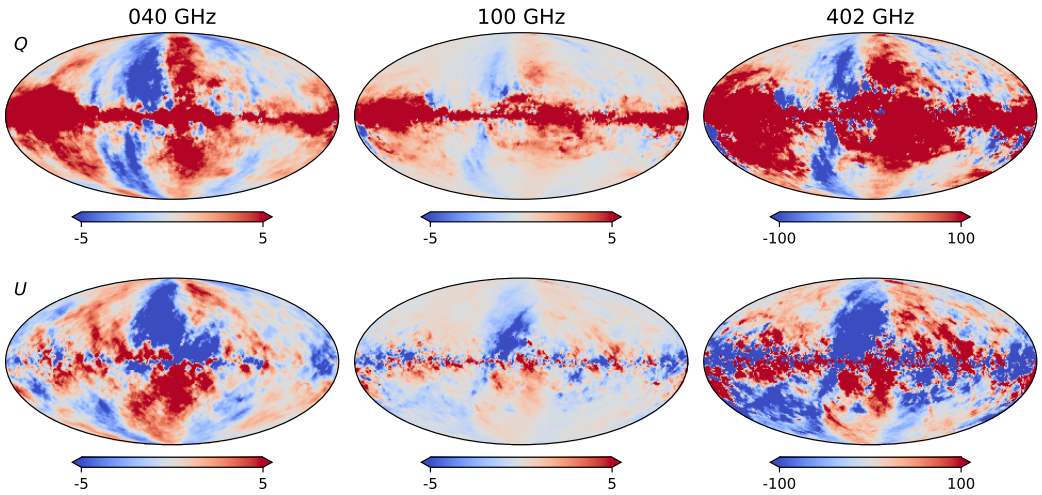


Figure 9.5. Simulated polarised foregrounds emission maps, in Q (top row) and U (bottom row) Stokes parameters, for the `d10s5` model, at synchrotron-dominated (40 GHz, *left column*), foregrounds-minimum (100 GHz, *middle column*), and dust-dominated (402 GHz, *right column*) frequencies. All maps are smoothed to 70.5 arcmin, and reported in $\mu\mathrm{K}_{\mathrm{CMB}}$ units.

then co-added to the CMB ($r_{\mathrm{in}} = 0$ or 5×10^{-3}) and noise simulations:

$$QU_{\mathrm{tot}}(r_{\mathrm{in}}, \mathtt{dxsy}) = QU_{\mathrm{CMB}}(r_{\mathrm{in}}) + QU_{\mathrm{noise}} + QU_{\mathrm{fg}}(\mathtt{dxsy})\,. \tag{9.3}$$

We end up with Q and U sky simulations which include CMB, noise, and foregrounds, for different combinations of input tensor-to-scalar ratio and `PySM` foregrounds model. For each combination we have 300 different realisations of CMB and noise. Figure 9.5 illustrates the simulated polarised foregrounds for the `d10s5` model. Emission at three different frequency channels is shown, to highlight the separate synchrotron and dust domination.

We then perform a full-sky harmonic transformation to convert Q/U simulations to E/B-mode maps, with the latter being our input dataset for all the subsequent analyses. It is important to notice here that the $Q/U \rightarrow E/B$ passage is a non-local transformation which in general is going to create artifacts in the final B-mode maps, due to the leakage of sky regions with very bright foregrounds radiation outside the Galactic plane. A more correct approach would be to first mask the most contaminated areas of the Q/U maps and then operate the transformation to B-modes, while correcting for the E-to-B leakage induced to presence of a mask (see [40, 186] for more details on the topic and for a MFs analysis of the problem). In the following we will not account for this issue, placing ourselves in the worst possible scenario, with no mitigation of this effect in the maps. The impact of this assumption will be discussed in the presentation of the results.

9.2 Component separation

We already introduced the concept of component separation in the first Part of the Thesis. Some sort of foregrounds-cleaning techniques must be incorporated in whatever data analysis pipeline targeting the measurement of the underlying CMB signal, moreover if the final goal is constraining the faint primordial B-modes, as we are doing here. As we mentioned and detailed in section 4.3, in this work we adopt component separation techniques usually referred to as "blind" algorithms, in the sense that there are no strong assumptions on the models governing the foregrounds emission. They offer a valuable alternative to parametric approaches, since they are not affected by possible wrong spectral modelling of the Galactic polarised emission, which may significantly bias the final estimate of the tensor-to-scalar ratio. These methods are also called "minimum-variance" techniques, as they linearly combine multi-frequency sky maps with the aim of minimising the variance of the recovered signal.

In particular, to recover the CMB B-mode signal from our simulated multi-frequency sky maps, we use two implementations of minimum-variance techniques: Needlet Internal Linear Combination (NILC) and Multi-Clustering NILC (MC-NILC), both of them summarised in section 4.3. NILC has already been applied in the past in CMB data analysis (e.g., *WMAP* [19], *Planck* [169]) and both NILC and MC-NILC will be used for current and future CMB surveys (*SO* [6, 218], *LiteBIRD* [127]). We use these two methods as implemented in the `BROOM`[5] package [43]. Component separation is applied directly on the input scalar B-modes maps we simulated (section 9.1).

We recall that the NILC algorithm aims at recovering a clean view of the CMB signal by linearly combining multi-frequency sky maps, performing variance minimisation at different harmonic scales [57, 40]. Input maps are decomposed with needlets filtering, allowing to separate the contribution of the different emitting components depending on their angular scales. In this way, the contaminants mitigation is able to take into account the inhomogeneity of the foregrounds emission, improving the reconstruction of the CMB signal with respect to the simpler ILC.

If the foregrounds emission is very complex, the NILC method may present some limitations or be suboptimal, since it still does not fully handle the local spectral variations of foregrounds emission. MC-NILC [176, 41] is a NILC extension which performs component separation independently in different sky regions, where the polarised B-mode Galactic emission shows similar spectral properties. The spatial variability of dust and synchrotron spectral parameters is assessed by identifying

[5] https://github.com/alecarones/broom

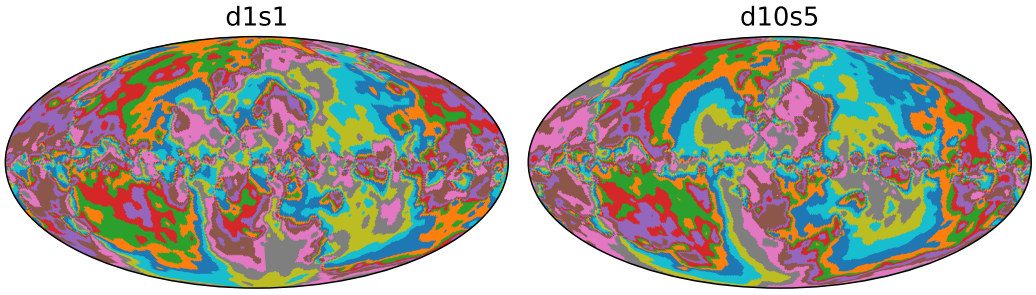


Figure 9.6. Sky partitions adopted in the MC-NILC component separation for `d1s1` (*left*) and `d10s5` (*right*) foreground models.

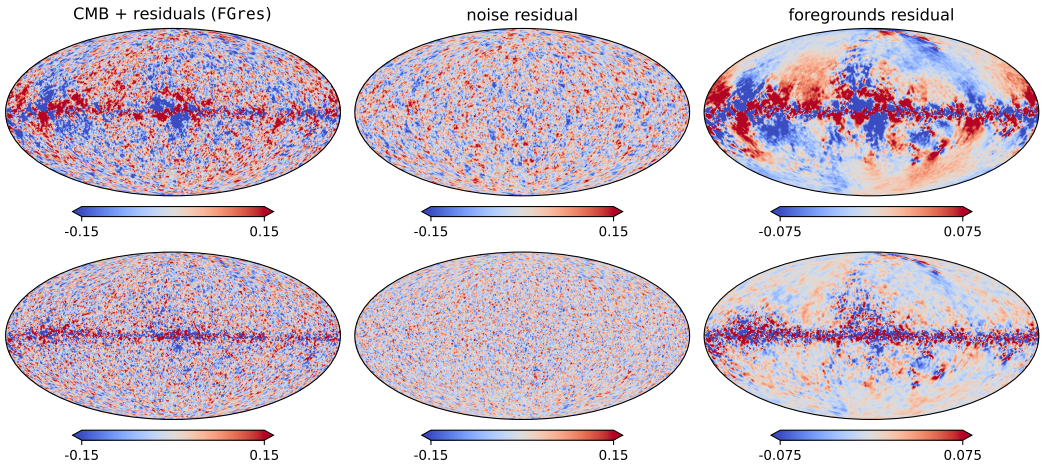


Figure 9.7. An example of the component separation outputs, for one simulation. *Left column:* reconstructed CMB B-mode signal (CMB solution), containing the CMB field, and noise and foreground residuals. *Middle column:* noise residuals. *Right column:* foreground residuals. Outputs are shown for NILC (*upper row*) and MC-NILC (*lower row*) methods, applied on the `d10s5` model with $r_{\rm in} = 0$. Units are $\mu{\rm K}_{\rm CMB}$.

a blind tracer of their distribution across the sky. In the implementation used in this work, the tracer is obtained from the simulated dataset by applying a model-independent and realistic de-noising technique [41]. Figure 9.6 illustrates how the full sky is separated into 50 patches for the `d1s1` and `d10s5` models, where regions with the same colour have similar values of the tracer and thus similar foregrounds spectral properties. Component separation is independently performed in each of these partitions.

Both NILC and MC-NILC algorithms relies on the notion of needlets. As we briefly mentioned in section 4.3, needlets are a wavelet system that guarantees simultaneous localisation in harmonic and pixel space. The shape of the needlet bands is defined by the choice of the harmonic function $b_j(\ell)$, whose width is set by the parameter B_n: lower values of B_n correspond to a tighter localisation

in harmonic space (less multipoles entering into any needlet coefficient). Here we adopt the Mexican needlet construction [151] as our harmonic function, with $B_n = 1.3$. The needlet filters have been obtained through the `MTNeedlet`[6] package [45], and they are shown in the left panel of figure 9.8. Since the first needlet bands would select only very few large-scale modes, we merge together the first 14 bands following

$$b_0^{\text{merged}} = \sqrt{\sum_{j=0}^{14} b_j^2(\ell)}\,. \tag{9.4}$$

An example of the performances of the two algorithms is reported in figures 9.7 and 9.8 (right panel), for the $r_{\text{in}} = 0$ case. The amount of leftover contamination in the recovered CMB field can be visually compared both at the map and at the power spectrum level, and the improvement achieved by MC-NILC can be clearly seen in the reduced amplitude of the noise and foreground residuals, in particular on the largest scales. The leftover contamination is noticeable in the foregrounds residual maps, with anisotropic structures present even at higher latitudes, far from the Galactic plane. But it is not so evident in the recovered CMB map, where the noise and CMB lensing contributions are present.

In general, a blind component separation algorithm returns a single CMB "solution" as final product, containing the reconstructed CMB signal plus some leftover contamination from noise and foregrounds. In our experimental setup, we are working with simulated datasets, thus having complete control over the separated sky components, as we independently generated CMB, noise, and foregrounds realisations. Component separation weights can be applied to the input noise and foregrounds maps to obtain the residuals maps shown in figure 9.7, in order to have an exact estimate of the leftover contamination in the recovered CMB field. Moreover, we also have the possibility to retrieve the exact input CMB field by removing the aforementioned noise and foreground residuals from the CMB solution. This is very useful for manipulating the CMB map without touching the other components; in particular, we will also characterise the robustness test in a scenario where the gravitational lensing contribution to the B-mode power spectrum is reduced by 50%, parametrised by a factor $A_{\text{lens}} = 0.5$. The mitigation of this kind of contamination is usually called "delensing"; modern experiments are able to achieve up to a 70% reduction of the CMB lensing power by means of correlations with other surveys of the large-scale structure of the Universe, responsible for the deflection of the CMB photons [6, 100].

In a real experiment, these single contributing maps are not accessible, as isolated noise and foregrounds residuals are not available. Nevertheless, it is useful

[6] https://javicarron.github.io/mtneedlet/index.html

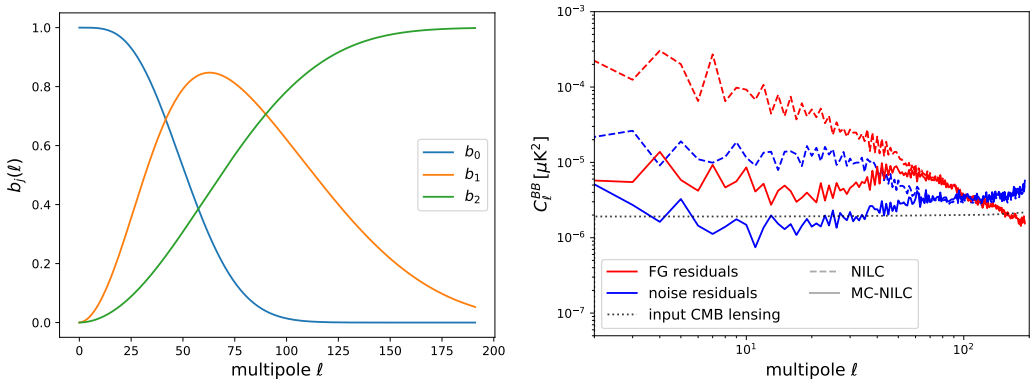


Figure 9.8. *Left panel*: needlet bands Mexican configuration adopted in this work for component separation, where the first bands have been merged. *Right panel*: full-sky angular power spectra evaluated on foreground (red) and noise (blue) residuals, for NILC (dashed lines) and MC-NILC (solid lines), applied to the `d10s5` model. Black dotted line is the input CMB spectrum (lensing-only).

to use them here for the full validation of the component separation pipeline and also for the characterisation of the robustness tests.

9.3 Likelihood

The application of component separation methods provided a reconstruction of the CMB signal, where foregrounds contamination is mitigated. In order to infer the tensor-to-scalar ratio from these maps, we compare their BB angular power spectrum with the theory through the evaluation of a likelihood function, which in this work is adopted to be an inverse-Wishart distribution [94, 86]:

$$-\ln\mathcal{L}\left(C_\ell^{\rm obs}|r\right) = \frac{1}{2}\sum_\ell (2\ell+1) f_{\rm sky}\left[\frac{C_\ell^{\rm obs}}{C_\ell^{\rm th}(r)} + \ln C_\ell^{\rm th}(r) - \frac{2\ell-1}{2\ell+1}\ln C_\ell^{\rm obs}\right] . \quad (9.5)$$

where $C_\ell^{\rm obs}$ is the "observed" power spectrum evaluated on the CMB solution map, containing CMB lensing, noise and foreground residuals, and $C_\ell^{\rm th}$ is the theoretical BB spectrum built as

$$C_\ell^{\rm th}(r) = C_\ell^{\rm lensing} + \langle C_\ell^{\rm nres}\rangle + r\cdot C_\ell^{r=1} . \quad (9.6)$$

In the above equation, $C_\ell^{\rm lensing}$ is the B-mode contribution induced by gravitational lensing, $C_\ell^{r=1}$ is the theoretical B-mode power spectrum sourced solely by tensor perturbations with $r=1$, and $\langle C_\ell^{\rm nres}\rangle$ is the average of the spectra computed on all the $N-1=299$ noise residuals maps obtained from the application of the component separation on the input simulations.

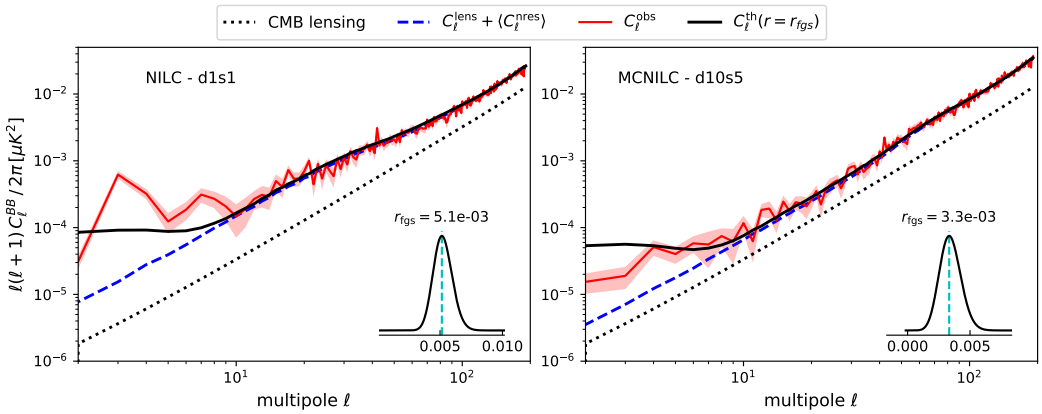


Figure 9.9. Angular power spectra going into the likelihood evaluation for a single simulation, for the NILC - `d1s1` (*left*) and MC-NILC - `d10s5` (*right*) scenarios. *Main plot*: spectra, with $C_\ell^{\rm th}(r=0)$ in dashed blue, $C_\ell^{\rm obs}$ in solid red, $C_\ell^{\rm lensing}$ in dotted black, and the best-fit in solid black. The shaded area represents the standard deviation of $C_\ell^{\rm obs}$ across 300 realisations. *Inset*: normalised likelihood on the tensor-to-scalar ratio; the peak value is indicated by the dashed line.

Angular power spectra are computed with the `anafast` routine as implemented in the `healpy` package. To avoid strong contamination from the Galactic plane, the same Galactic mask we used in the previous Part (left panel of figure 6.5) is applied to the maps, retaining a sky fraction of $f_{\rm sky} \simeq 60\%$. As we described in section 3.4, the presence of a mask induces correlations among multipoles (mode-coupling [122]). This effect is not taken into account and not corrected for when estimating power spectra with `anafast`, as we are doing here. However, as discussed in [127], this approximation has a negligible impact on the power spectra estimation of foreground and noise residuals over large sky fractions, as considered in this work.

The likelihood function is evaluated over an interval of r values in the range $r \in [-1.5 \times 10^{-4}, 3 \times 10^{-2}]$, with a constant spacing of $\Delta r = 2 \times 10^{-6}$. Values of $r < 0$ are included in the analysis in order to be sensitive to negative biases on the tensor-to-scalar ratio, as our generated CMB simulations assumes $r = 0$. We take the peak value of the likelihood as our best fit tensor-to-scalar ratio $r_{\rm fgs}$, where the **fgs** label indicates that the bias on the parameter is entirely due to the presence of foreground residuals, which contribute to the power observed on the large scales and mimic the behaviour of a primordial GWs signal. An example of the likelihood inputs and outputs is provided in figure 9.9, where we report the power spectra of the recovered CMB map, noise residuals, and the input CMB lensing, together with the best-fit tensor B-mode spectra and the corresponding posterior of $r_{\rm fgs}$.

The inferred value of the tensor-to-scalar ratio gathers contributions from all the angular scales included in the analysis. In general, the largest scales (low ℓ)

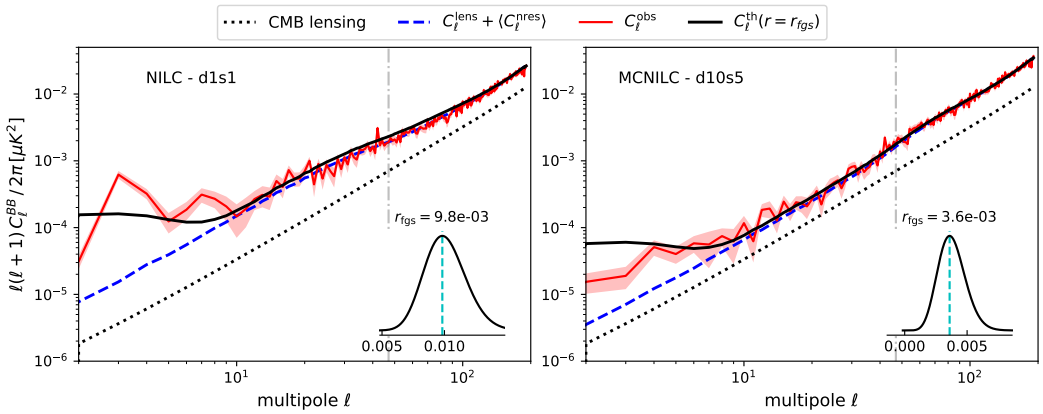


Figure 9.10. Same as figure 9.9, but considering only the largest scales in the likelihood evaluation. The maximum multipole included ($\ell = 48$) is indicated by the vertical dash-dotted line.

contain most of the information about foregrounds and inflationary physics, while the smaller scales (high ℓ) are dominated by noise and CMB lensing, having therefore a lower impact on the r measurement. To highlight the different contributions, we also evaluate the likelihood function over three multipoles range: $\Delta\ell_0 \in [2, 48]$, $\Delta\ell_1 \in [15, 175]$, and $\Delta\ell_2 \in [40, 191]$. These intervals follow the scales sampled by the needlet bands used for component separation, including all multipoles where each of the adopted needlet bands is greater than 30%. We will mostly focus on the first interval, as we are mainly interested on the impact of foreground residuals on the largest angular scales. Figure 9.10 shows the same power spectra featured in figure 9.9, but now considering only the largest scales ($\Delta\ell_0$) in the likelihood analysis, with the last included multipole indicated by the vertical line. We can visually notice a slight improvement of the fitted curve (solid black), together with a wider likelihood reflecting the increase in the uncertainty on the larger recovered value of r.

Following the procedure which would be adopted in the data analysis of a real experiment, we can propose a first robustness test by evaluating the goodness-of-fit, to understand if we are (already) able to assess whether the bias on r is induced by the presence of foreground residuals. We note that eq. (9.5) is not a Gaussian likelihood, and thus it is not possible to define a χ^2 statistic in the usual way. However, we can still evaluate the fit quality by relying on simulations, following the process below, described for a single simulation:

1. we use the inferred tensor-to-scalar ratio $r_{\rm fgs}$ to generate 300 Gaussian random realisations of CMB B-mode maps, with their power spectrum being $C_\ell^{BB} = r_{\rm fgs} \cdot C_\ell^{r=1}$. We co-add these maps with simulations of CMB lens-

Table 9.2. Percentage of incompatible simulations from the goodness-of-fit analysis, considering all scales at the same time or only the largest scales. Values are reported for each scenario considered.

all multipoles ($\Delta\ell$=2–191)

	d0s0	d1s1	d10s5
NILC	4	98	74
MC-NILC	-	28	22

large-scales only ($\Delta\ell_0$=2–48)

	d0s0	d1s1	d10s5
NILC	4	98	71
MC-NILC	-	52	43

ing and noise residuals, provided by the component separation. Doing so means taking the CMB solution and replacing the foreground residuals with realisations of tensor perturbations;

2. for each of these combined CMB lensing, noise, and tensor perturbation maps, we compute their power spectrum and we fit for r with the same likelihood function of eq. (9.5). We take each minimum of the $-\ln\mathcal{L}$ (similar to the usual χ^2) to build a distribution that we can use to compare with the value obtained by the corresponding best fit $r_{\rm fgs}$;

3. we compute the probability-to-exceed (PTE), that is the amount of the distribution which falls above the value associated to the best fit $r_{\rm fgs}$. If the PTE is less than 0.05 (95%), we deem the fit as not good for the considered simulation.

We repeat this analysis for all the 300 input simulations. In general, a reasonable goodness-of-fit is associated with a PTE distribution close to the uniform one, with the percentage of incompatible simulations being around $\sim 5\%$. The results of this analysis are reported in table 9.2, while figure 9.11 shows the $-\ln\mathcal{L}$ distribution for one simulation on the left, and the PTE distribution for the 300 simulations on the right, along with the percentage of simulations failing this χ^2-like test. For this particular realisation, the NILC - `d1s1` case fails the test (PTE < 0.05) while for MC-NILC - `d10s5` the test is passed (PTE > 0.05).

As it can be seen from the table, using the NILC component separation leads to a poor fit quality, with PTE distributions far from being uniform and the bias identified in $> 70\%$ ($> 90\%$) of the simulations for the `d10s5` (`d1s1`) model. This is suggesting that the shapes of the power spectra at the largest scale do not match enough for the fit to be acceptable: in a real experiment, we would likely get a PTE < 0.05, already ruling out the cosmological origin of the inferred tensor-to-scalar ratio. On the other hand, in the MC-NILC case the PTE distribution is closer to a uniform one,

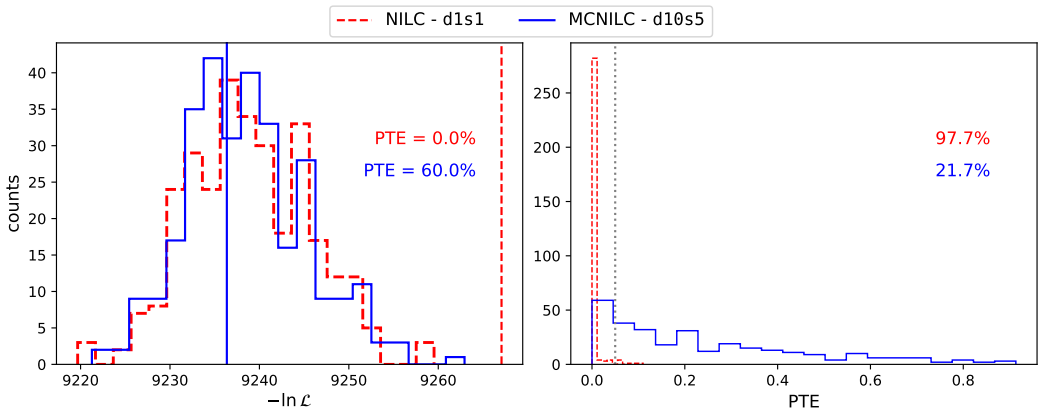


Figure 9.11. Results of the goodness-of-fit analysis for the NILC - `d1s1` (red dashed lines) and MC-NILC - `d10s5` (blue solid lines) cases. *Left panel:* one realisation of the $-\ln \mathcal{L}$ distribution of the 300 tensor simulations, compared against the value given by the best-fit on the CMB solution (vertical lines); the corresponding PTE is also indicated. *Right panel:* PTE distribution for the 300 simulations, with the percentage of incompatible simulations shown on the right.

and the percentage of simulations with a bad fit is $< 30\%$. In a real experiment, the biased r measurement would likely be accepted if we base ourselves on a simple goodness-of-fit evaluation. This calls for other and more advanced robustness tests able to reject a potential false detection.

It is worth noticing that in this analysis we are restricting ourselves to specific foreground models, component separation methods, and instrumental configurations, which may not represent the most general case. To simulate a broader scenario, we also consider a situation where the likelihood step leads to an optimal fit, where no power spectra mismatch is present. In such a case, only higher-order statistics can reveal discrepancies between residuals and primordial cosmological signal. We will further detail this point in section 9.4.2.

9.4 Reference Gaussian simulations

The goal of the robustness test is to compare the statistical properties of cleaned CMB maps (`CMBsol`) with reference Gaussian simulations, going beyond the usual two-point statistics. The foreground residuals left by component separation are expected to show non-Gaussian properties and anisotropic structures which should not be present in a field dominated by primordial CMB B-modes. Therefore, we build two sets of Gaussian simulations dubbed `CMBtens` and `FGgauss`, which are co-added with the input CMB lensing maps and the noise residuals coming from the application of component separation to the other realisations. Thus, the `CMBsol`

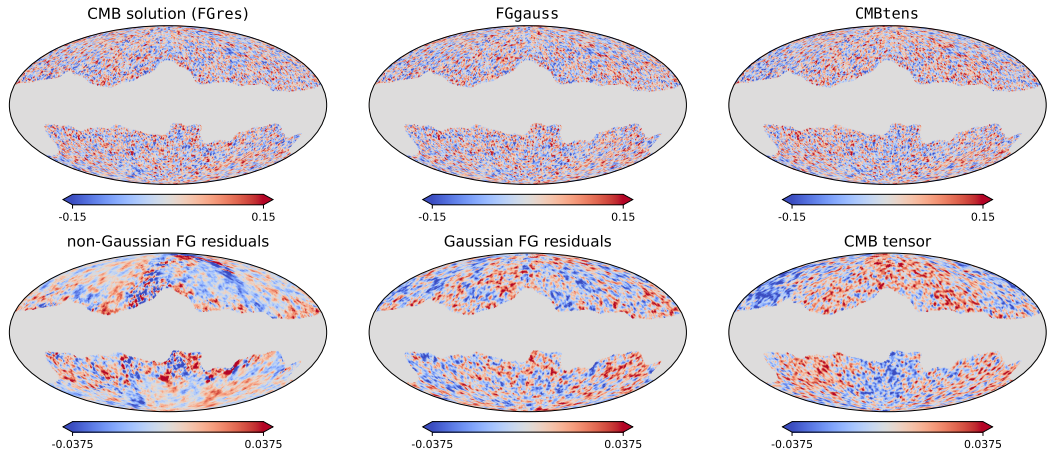


Figure 9.12. An example of the maps considered for the robustness tests, for the MC-NILC - `d10s5` case. *Upper row, from left to right:* CMB solution (`CMBsol`), Gaussian simulations `FGgauss` and `CMBtens`. *Lower row, from left to right:* as upper row, but without CMB lensing and noise, with non-Gaussian and Gaussian foreground residuals, and tensor perturbations.

and Gaussian maps have different realisations of CMB signal and noise in them.

If the detected tensor-to-scalar ratio is of Galactic origin, then the `CMBsol` map (that includes anisotropic contamination) and the Gaussian simulations should be statistically incompatible when observed through a higher-order statistics, failing the robustness test. Vice versa, if no deviation is observed between two maps, it means that the statistics is not able to detect the foreground residuals non-Gaussianity. In the latter case, the test is passed and no flag is raised. An example of these two types of maps and their co-addition with CMB lensing and noise residuals is reported in figures 9.12 and 9.13, including all scales or only the largest, respectively, and after the application of the Galactic mask. We notice how the three maps to be compared (top row) look very compatible at least by eye, since they are dominated by noise. Foreground residuals display features of Galactic origin, with evident anisotropic structures over the sky, but these structures are not so obvious when foreground residuals are mixed with CMB and noise. This is why we need sophisticated statistical tools to identify non-Gaussianity. We now describe these Gaussian simulations in more details.

9.4.1 `CMBtens`

For each of the 300 component-separated `CMBsol`, we estimated a tensor-to-scalar ratio $r_{\rm fgs}$, eventually biased by foreground residuals with respect to the input value ($r_{\rm in} = 0$ or 5×10^{-3}). We use the inferred value to generate 5000 Gaussian realisations of B-mode maps sourced only by primordial GWs, from the power

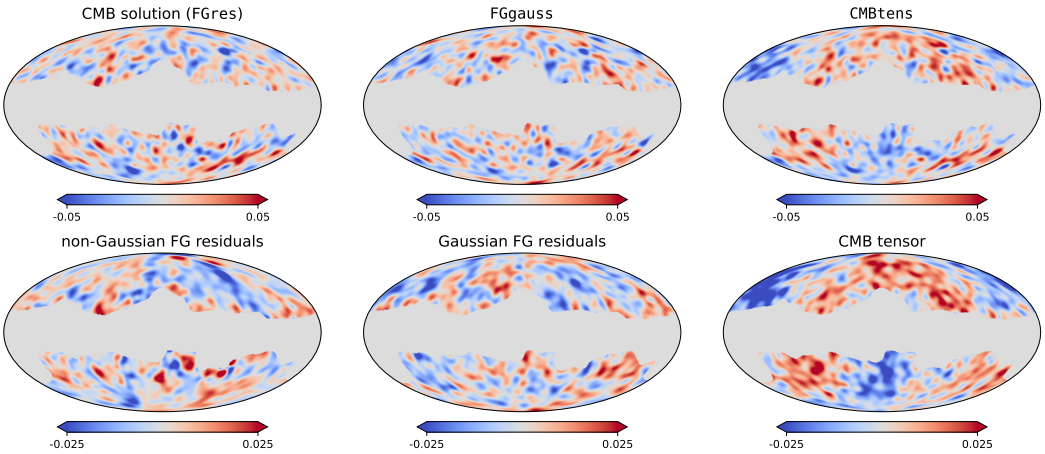


Figure 9.13. Same maps as figure 9.12, but filtered with the first needlet band b_0 to retain only the largest scales, at multipoles $\ell \lesssim 60$.

spectrum $C_\ell^{BB} = r_{\text{fgs}} \cdot C_\ell^{r=1}$, using the `synfast` routine from `healpy`. Maps are smoothed to the same resolution of `CMBsol`. We then co-add these simulations with different, independent realisations of CMB lensing and noise residuals maps, obtained from the application of component separation on the other input simulations. We refer to these simulated Gaussian maps as `CMBtens`. We thus end up with 5000 generated maps for each of the 300 input simulations.

9.4.2 `FGgauss`

Based on the discussion of section 9.3, the power spectrum of the `CMBtens` simulations may or may not be compatible with the one of `CMBsol` maps. Their CMB lensing and noise contributions are statistically the same, while the tensor contribution is fitted on the foreground spectrum contamination. A first robustness test is represented by a standard goodness-of-fit analysis, that can be able to identify a bias on the r estimate if the fitted curve has a much different spectral shape than the recovered CMB signal. Vice versa, if the power spectra are similar enough, the χ^2-analysis is passed and no flag is raised, as it happens in many of our sky realisations.

In order to simulate other situations where the fit is optimal and the χ^2 is not informative, we build the `FGgauss` simulations by first taking the angular power spectrum of foregrounds residuals maps (right column of figure 9.7) outside of the usual Galactic mask, and then using this spectrum as the $a_{\ell m}$ variance in order to generate 5000 Gaussian realisation with the `synfast` routine from `healpy`. This means that we are treating the spectrum of a single realisation as the correct underlying theoretical spectrum, with the average spectrum of the Gaussian simulations

converging to the non-Gaussian one. Finally, we add CMB lensing and noise residuals contributions as we did for the `CMBtens` set. In practice, we are replacing the non-Gaussian foreground residuals present in `CMBsol` with a Gaussian counterpart. In this way, the robustness test will compare maps having a statistically identical power spectrum, with the only difference being the non-Gaussianity. Again, we end up with 5000 generated maps for each of the 300 input simulations.

We recall that for a real experiment we do not have access to the component separation foreground residuals and, consequently, to the possibility of generating this kind of simulations[7]. Nonetheless, they provide a critical benchmark for the characterisation of the robustness performances, in an idealised scenario where the contribution from non-Gaussianity is isolated. This is very useful to understand if the selected higher-order statistics can be reliable in the validation of a potential tensor-to-scalar ratio detection. But even in a real experiment, it is always possible to generate the `CMBtens` simulations previously described.

This approach also helps in alleviating limitations induced by some assumptions we made in our analysis, such as *i)* the adoption of a simplified sky mask, which is sub-optimal for mitigating foregrounds contamination, and *ii)* the assumption of an inverse-Wishart form for the likelihood, which may be sub-optimal in the case of an analysis on a partial sky [86]. We also note that the assumptions may contribute to the bias on r, as a mask or likelihood optimisation would reduce the bias on the parameter.

9.5 Minkowski functionals

Minkowski functionals (MFs) [144] are mathematical tools describing the morphology of fields in any dimension [8]. They have been introduced into Cosmology as descriptors for the morphological properties of large-scale structure [142], and then have successively been widely adopted in the context of CMB data analysis [187, 188, 217, 161, 171], used as descriptors of non-Gaussianity and anisotropy.

Given a scalar random field f observed on the unit sphere $\mathbb{S}^2$ (e.g., CMB temperature anisotropies, or E-/B-modes), the excursion set at a threshold u is defined as

$$A_u\left(f, \mathbb{S}^2\right) = \{x \in \mathbb{S}^2 : f(x) > u\}\,. \tag{9.7}$$

We will omit the arguments of A_u when f and its domain are clear from the context. As explained in [188], the morphological properties of these excursion sets can be

[7] Even if we would still be able to take the full component-separated map `CMBsol` and generate Gaussian versions of it.

summarised by the three MFs defined as follows:

$$\begin{aligned} V_0(A_u) &= \int_{A_u} \mathrm{d}x, \\ V_1(A_u) &= \frac{1}{4}\int_{\partial A_u} \mathrm{d}r, \\ V_2(A_u) &= \frac{1}{2\pi}\int_{\partial A_u} \kappa(r)\mathrm{d}r\,, \end{aligned} \tag{9.8}$$

where $\mathrm{d}r$ denotes a line element along the boundary ∂A_u of the excursion set and $\kappa(r)$ denotes the geodesic curvature of the said boundary. Thus, V_0 represents the excursion area, V_1 is one-fourth of the boundary length (perimeter), and V_2 (genus) is associated with the number of connected regions minus the number of holes, which on the plane corresponds to the Euler–Poincaré characteristic χ. On a spherical surface, the Euler-Poincaré characteristic no longer corresponds exactly to V_2, but includes an additional term that can be computed by means of the generalized Gauss-Bonnet theorem, yielding:

$$\chi(A_u) = V_2(A_u) + \frac{1}{2\pi}V_0(A_u)\,. \tag{9.9}$$

Generally, these MFs notation is replaced with the equivalent notion of Lipschitz-Killing curvatures, expressed as

$$\begin{aligned} \mathcal{L}_2(A_u) &= V_0(A_u)\,, \\ \mathcal{L}_1(A_u) &= 2V_1(A_u)\,, \\ \mathcal{L}_0(A_u) &= V_2(A_u) + \frac{1}{2\pi}V_0(A_u)\,. \end{aligned} \tag{9.10}$$

For an isotropic Gaussian field (normalised to have unit variance), the expected value of the Lipschitz-Killing curvatures on the sphere is given by the Gaussian kinematic formula (GKF):

$$\mathbb{E}\left[\mathcal{L}_j(A_u)\right] = \sum_{k=0}^{2-i} \frac{\omega_{k+i}}{\omega_k\omega_i}\binom{k+i}{k} \cdot \rho_k(u) \cdot \mathcal{L}_{k+i}(\mathbb{S}^2) \cdot \mu^{k/2}\,, \tag{9.11}$$

where we defined the "flag" coefficients [8] as

$$\omega_i = \frac{\pi^{i/2}}{\left[\Gamma(\frac{i}{2}+1)\right]}\,, \tag{9.12}$$

with $\Gamma(z)$ being the usual Gamma function. In general, ω_i represents the volume of the i-dimensional unit ball ($\omega_1 = 2$, $\omega_2 = \pi$, $\omega_3 = 4\pi/3$, ...). μ is the derivative of the

covariance function at the origin for the field f, computed as

$$\mu = \sum_{\ell} \frac{\ell(\ell+1)}{2} \frac{2\ell+1}{4\pi} C_\ell \,, \tag{9.13}$$

where C_ℓ, as always, denotes the angular power spectrum of the field f. The ρ_k quantity in eq. (9.11) represents a set of "density" functions defined as

$$\rho_k(u) = \frac{1}{(2\pi)^{k/2}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) H_{k-1}(u) \,, \tag{9.14}$$

with $H_k(u)$ denoting the sequence of Hermite polynomials:

$$\begin{aligned} H_{-1}(u) &= \sqrt{2\pi} \exp\left(\frac{u^2}{2}\right) \left[1 - \Phi(u)\right] \,, \\ H_0(u) &= 1 \,, \\ H_1(u) &= u \,, \\ H_k(u) &= (-1)^k \exp\left(\frac{u^2}{2}\right) \frac{\mathrm{d}^k}{\mathrm{d}u^k} \exp\left(-\frac{u^2}{2}\right) \,. \end{aligned} \tag{9.15}$$

Combination of eqs. (9.10) and (9.11) provide the MFs expected values, which we report normalised by area. The first one is:

$$\begin{aligned} \frac{\mathbb{E}(V_0)}{4\pi} &= \frac{\mathbb{E}\left[\mathcal{L}_2(A_u)\right]}{4\pi} \\ &= \frac{1}{4\pi} \cdot \rho_0(u) \cdot \mathcal{L}_2(\mathbb{S}^2) \\ &= 1 - \Phi(u) \,. \end{aligned} \tag{9.16}$$

where Φ is the cumulative function of the normal distribution. We note that the above MF depends only on the threshold u and not on the field itself, as long as it is normalised to have unit variance. For the second MF we have

$$\begin{aligned} \frac{\mathbb{E}(V_1)}{4\pi} &= \frac{\mathbb{E}\left[\mathcal{L}_1(A_u)\right]}{2 \cdot 4\pi} \\ &= \frac{1}{8\pi} \cdot 2\frac{\omega_2}{\omega_1^2} \cdot \rho_1(u) \cdot \mathcal{L}_2(\mathbb{S}^2) \cdot \mu^{1/2} \\ &= \frac{1}{8\pi} \frac{2\pi}{4} \frac{1}{(2\pi)^{1/2}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \cdot 4\pi \cdot \mu^{1/2} \\ &= \frac{1}{8} \exp\left(-\frac{u^2}{2}\right) \mu^{1/2}, \end{aligned} \tag{9.17}$$

while for the third we need to first evaluate the expected value of $\mathcal{L}_0$ as

$$\begin{aligned}\mathbb{E}\left[\mathcal{L}_0(A_u)\right] &= \mathbb{E}\left[\chi(A_u)\right] = \mathbb{E}\left[V_2(A_u) + \frac{1}{2\pi}V_0(A_u)\right] \\ &= \rho_0(u)\cdot\mathcal{L}_0(\mathbb{S}^2) + \rho_2(u)\cdot\mathcal{L}_2(\mathbb{S}^2)\cdot\mu \\ &= 2\left[1-\Phi(u)\right] + \frac{2}{\sqrt{2\pi}}\mu\cdot\exp\left(-\frac{u^2}{2}\right)\cdot u\,,\end{aligned} \tag{9.18}$$

and then

$$\begin{aligned}\frac{\mathbb{E}(V_2)}{4\pi} &= \frac{1}{4\pi}\{\mathbb{E}\left[\mathcal{L}_0(A_u)\right] - \frac{1}{2\pi}V_0(A_u)\} \\ &= \frac{1}{\sqrt{(2\pi)^3}}\mu\cdot\exp\left(-\frac{u^2}{2}\right)\cdot u\,.\end{aligned} \tag{9.19}$$

The above set of equations tell us that the MFs expected values, in the case of a Gaussian random field, depend on the combination of four different independent ingredients:

1. a set of universal coefficients ω_i (called "flag" coefficients);
2. the Lipschitz-Killing curvatures $\mathcal{L}_i$ evaluated on the original manifold (in our case, the unitary sphere);
3. a power of the derivative of the covariance function μ of the field at the origin;
4. a set of "density" functions ρ_k, dependent only on the threshold at which the MFs are evaluated, and on the statistics of the field f.

A nice feature of MFs is that the presence of a mask only affects the manifold where the map is defined, having therefore an impact only on the MFs normalisation, at leading order.

A key point to notice is that, from eq. (9.11), the MFs expected value depends both on the level of Gaussianity of the field (in general, on the statistics of the field through ρ_k), and on its spectral features through the C_ℓ term in the derivative of the covariance function, in eq. (9.13). This implies that two fields will have different MFs values either due to different statistical properties or different shape of the spectra.

In this work, MFs of scalar maps are evaluated using `Pynkowski`[8], a `python` package developed to compute MFs and other higher order statistics of fields defined on the sphere and other manifolds, in `HEALPix` convention. The package has been presented and validated in [46, 42], both for scalar ($T,\ E,\ B,\ P^2$) and polarisation (Q, U) fields.

[8] https://javicarron.github.io/pynkowski/pynkowski.html

On the computational side, the first MF for a scalar field f is calculated in `Pynkowski` as

$$v_0(u) = \frac{1}{4\pi}\int_{\mathbb{S}^2} \Theta\left[f(x) - u\right] \mathrm{d}x \simeq \frac{1}{4\pi}\sum_p \Theta\left[f(p) - u\right] \,, \tag{9.20}$$

where in the second equality the integral is replaced by a discrete sum over all the valid pixels p, and $\Theta(z)$ is the Heaviside function being 1 when $z \geq 0$ and 0 otherwise. Therefore, v_0 is just the fraction of pixels with values above the threshold u.

The second MF is computed as

$$\begin{aligned} v_1(u) &= \frac{1}{4}\frac{1}{4\pi}\int_{\partial A_u} \mathrm{d}r \\ &= \frac{1}{4}\frac{1}{4\pi}\int_{\mathbb{S}^2} \delta\left[f(x) - u\right] \cdot |\nabla f| \, \mathrm{d}x \,, \end{aligned} \tag{9.21}$$

where the second equality stems from the change of coordinates from the line element $\mathrm{d}r$ of the excursion set boundary to a surface element $\mathrm{d}x$ on the sphere. ∇f is the gradient of the field and δ is the Dirac delta. The integral is approximated as a sum in pixel space.

The third MF is derived from the equation

$$\begin{aligned} v_2(u) &= \frac{1}{4\pi}\frac{1}{2\pi}\int_{\partial A_u} \kappa(r)\mathrm{d}r \\ &= \frac{1}{4\pi}\frac{1}{2\pi}\int_{\mathbb{S}^2} \delta\left[f(x) - u\right] \cdot |\nabla f| \cdot \kappa \, \mathrm{d}x \,, \end{aligned} \tag{9.22}$$

where the change of coordinates and integral replaced by sum are the same as the previous one. κ is the geodesic curvature of the excursion set boundary, given by the implicit curve expression [39]

$$\kappa = \frac{2 f_{;\phi} f_{;\theta} f_{;\phi\theta} - f_{;\theta}^2 f_{;\phi\phi} - f_{;\phi}^2 f_{;\theta\theta}}{\left(f_{;\phi}^2 + f_{;\theta}^2\right)^{3/2}} \,, \tag{9.23}$$

where the semicolon denotes the covariant derivative. All the quantities involved in this expression can be computed for each pixel. The spatial derivatives of the maps (both partial and covariant) are computed with the `alm2map_der` function of the `healpy` package [223] in harmonic space. Derivative operations require the maps to be smooth enough at small scales to avoid edge effects; this is usually achieved through the application of a beam smoothing function, suppressing the signal at higher multipoles.

Following what has been done in previous works [171, 42], in order to charac-

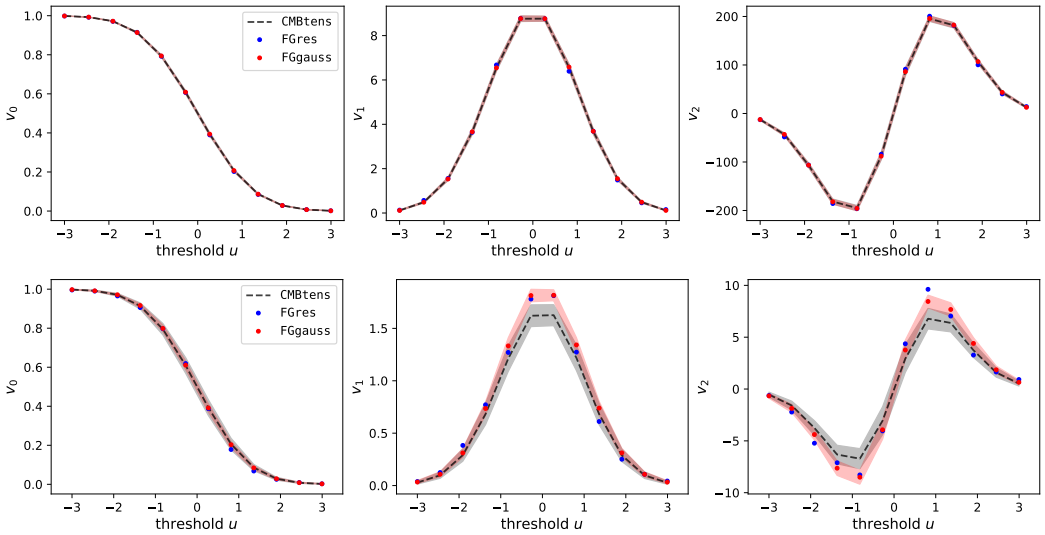


Figure 9.14. Minkowski functionals, for a single simulation, in the MC-NILC - `d10s5` case. MFs of the `CMBsol` map are shown with blue points, while the average MFs of 300 `CMBtens` (`FGgauss`) simulations are indicated by the black dashed line (red points). The shaded areas represent the standard deviation across the simulations. *Upper row*: all angular scales included. *Lower row*: largest scales only, $\ell \lesssim 60$.

terise the MFs we consider two approaches for the scale-dependent analysis of the B-mode sky maps: *i)* simultaneous exploitation of all the scales sampled in the maps (i.e., using all the multipoles available in the maps, from $\ell_{\rm min}$ to $\ell_{\rm max}$); *ii)* scales separation by using needlet bands, to isolate the contribution from different ranges of multipoles. The localisation properties of needlets (in real and harmonic space) allow a much more precise, scale-by-scale, interpretation of any possible anomaly, since non-Gaussian or statistically anisotropic signal may be relevant only in a specific range of multipoles (e.g., Galactic foregrounds residuals on large scales). The needlets approach should then provide a complete view about the nature of possible non-Gaussian features detected with the MFs.

Figure 9.14 illustrates an example of the three MFs evaluated on the three set of maps considered for the robustness test, both including all the angular scales (upper panels) or retaining only the largest scales (lower panels). In all MFs computations, the Galactic mask is applied.

9.6 Scattering transform statistics

Scattering transform (ST) statistics are a class of summary statistics for the study of non-Gaussian processes [132, 33]. They are inspired by the neural networks functioning, but without the need of any training steps. They exploit iterative wavelet

convolutions and non-linear operations as the modulus to characterise interactions between oriented spatial scales. Recently introduced in astrophysics [10, 11], ST statistics have already demonstrated their ability to efficiently characterise a variety of non-Gaussian processes, as for instance the interstellar medium [182, 118], the large-scale structures (LSSs) of the Universe [48, 210, 209], and the epoch of reionisation [104]. ST coefficients can also be used to build very efficient generative models in a maximum entropy framework, both for physical fields [34, 49, 148, 105] and CMB systematics datasets [38]. In turn, this can be exploited to build efficient statistical component separation and inverse problem algorithms, which have in part been applied to real astrophysical data [182, 58, 206, 157].

In this section we give a brief introduction of the concepts needed to discuss why are they useful for our robustness tests and to understand the results, referring to other works [132, 33, 10, 58, 49] for complete presentations and detailed discussions about the topic. ST are summary statistics for high-dimensional non-Gaussian signals. They are built from a cascade of convolutions with fixed filters (a wavelet transform), a modulus operation, and spatial averages or covariances. The filters can be analytically defined, and there is no need to learn them from data as a typical neural network would do. In this work, Morlet wavelets are adopted as filtering functions. A key advantage of ST over other higher-order statistics is their controlled variance. Indeed, estimators based directly on moments of the field, such as the bi-spectrum, have a large variance dominated by outliers [215]; instead, using the modulus as the non-linearity yields statistics with a controlled variance. This is crucially important in the context of our robustness tests for primordial B-modes, since the compared maps are noise-dominated and the non-Gaussian signal is faint, and must be told apart from sampling variance.

The wavelet transforms are computed in pixel space on the sphere by convolution with a family of dyadic Morlet wavelets $\psi_{(j,\theta)}$, obtained by adopting a mother wavelet and then dilating it by a factor 2^j and rotating it by a factor $\theta \cdot \pi/4$. In practice, we label $\lambda = (j, \theta)$ the oriented scales, and ψ_λ the related wavelets. We probe $J = 6$ dyadic scales ($j_{\text{min}} = 0$ to $j_{\text{max}} = 5$) and $\Theta = 4$ orientations between zero and π. Starting from $N_{\text{side}} = 64$ at $j = 0$, each convolution is carried out at resolution $N_{\text{side}} = 64 \cdot 2^{-j}$, so that a larger j corresponds to a larger angular scale on the sky. Coefficients are computed with the `FOSCAT`[9] package [58], which provides fast, GPU-accelerated wavelet convolutions in pixel space on the sphere.

The ST statistics provides a family of coefficients conventionally denoted S_1, S_2, S_3 and S_4, together with their cross-field counterparts $S_1^\times$, $S_2^\times$, $S_3^\times$ and $S_4^\times$. For a

[9] https://github.com/jmdelouis/FOSCAT.

single field I they read [146, 49, 148]

$$\begin{aligned} S_1(\lambda) &= \langle |I \star \psi_\lambda| \rangle \,, \\ S_2(\lambda) &= \langle |I \star \psi_\lambda|^2 \rangle \,, \\ S_3(\lambda_1, \lambda_2) &= \mathrm{Cov}\left[I \star \psi^{\lambda_1},\, |I \star \psi^{\lambda_2}| \star \psi^{\lambda_1} \right] , \\ S_4(\lambda_1, \lambda_2, \lambda_3) &= \mathrm{Cov}\left[|I \star \psi^{\lambda_3}| \star \psi^{\lambda_1},\, |I \star \psi^{\lambda_2}| \star \psi^{\lambda_1} \right] , \end{aligned} \tag{9.24}$$

and for two fields I_1 and I_2 they generalise to

$$\begin{aligned} S_1^\times(\lambda) &= \langle |(I_1 \star \psi_\lambda)(I_2 \star \psi_\lambda)^*| \rangle \,, \\ S_2^\times(\lambda) &= \langle (I_1 \star \psi_\lambda)(I_2 \star \psi_\lambda)^* \rangle \,, \\ S_3^\times(\lambda_1, \lambda_2) &= \mathrm{Cov}\left[I_1 \star \psi^{\lambda_1},\, |I_2 \star \psi^{\lambda_2}| \star \psi^{\lambda_1} \right] , \\ S_4^\times(\lambda_1, \lambda_2, \lambda_3) &= \mathrm{Cov}\left[|I_1 \star \psi^{\lambda_3}| \star \psi^{\lambda_1},\, |I_2 \star \psi^{\lambda_2}| \star \psi^{\lambda_1} \right] . \end{aligned} \tag{9.25}$$

where $\star$ indicates the convolution with the wavelet filter, $|\cdot|$ is the modulus, and $\langle \cdot \rangle$ is the spatial average. In the last two coefficients above, the covariance between two complex fields X and Y is defined as $\mathrm{Cov}[X, Y] = \langle X Y^* \rangle - \langle X \rangle \langle Y^* \rangle$. Each dimension of the S_1 and S_2 statistics is $J \cdot \Theta$, while S_3 and S_4 have respectively dimensions $J^2\Theta^2$ and $J^3\Theta^3$. The S_1 and S_2 coefficients characterise a single oriented scale λ, probing the structure contained within an individual wavelet band. In practice, the S_2 coefficient characterise the power- and cross-spectrum of the signal on the band of scales probed by ψ_λ, while the additional information brought by S_1 characterise the sparsity of the signal on these bands. But they do not encode the coupling between distinct oriented scales. Such couplings are instead captured by the higher-order coefficients S_3 and S_4, with different oriented scales λ_1, λ_2 and λ_3 entering the computation. In principle, these higher-order coefficients should contain a lot of additional information about non-Gaussianity, since they try to quantify correlation between different oriented scales, a feature typical of foreground-induced non-Gaussian structures as for instance blobs, filaments, or vortices.

The employment of both the auto- and cross-coefficients actually allows to investigate different properties of the maps:

- **non-Gaussianity.** The auto-coefficients probe the presence of non-Gaussian features associated to foreground residuals. In the robustness test, we are mainly looking for deviations between the coefficients of the `CMBsol` map and the coefficients of `CMBtens` and `FGgauss` fields. We expect these deviations to be more prominent at large scales, where foreground residuals are dominant.

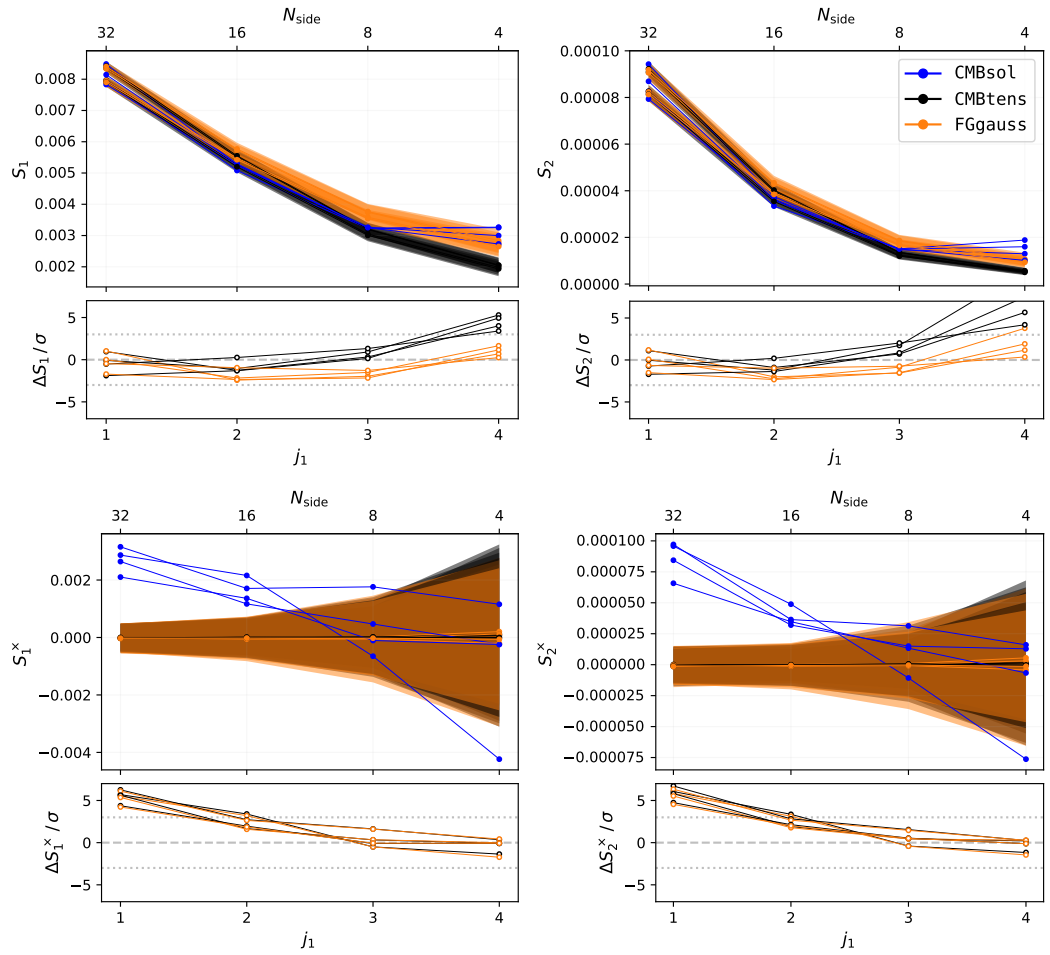


Figure 9.15. Scattering coefficients for the `d10s5` model with $r_{\rm in} = 0$, one simulation. Same-coloured lines represent different orientations. Coefficients of the component-separated CMB map are in blue, while black (orange) stands for the average coefficients of the `CMBtens` (`FGgauss`) Gaussian simulations. The secondary x-axis shows the $N_{\rm side}$ of the maps for scale j. *Upper figure:* auto-coefficients, using NILC. *Lower figure:* cross-coefficients, correlating with 40 GHz data, using MC-NILC. *Top panels:* coefficients, with the shaded areas indicating the standard deviation of the Gaussian simulations σ. *Bottom panels:* normalised difference between coefficients of `CMBsol` and of Gaussian maps, with the dotted lines indicating the $\pm 3\sigma$ region.

- **correlated structures.** The cross-coefficients probe instead the presence of correlated structures between two fields. We are mainly interested in the correlation between the foreground residuals in `CMBsol` (taken as I_1), and another selected map which possibly traces foreground emission (taken as I_2).

For the latter, in this work we adopt as the second field two of the input simulated frequency channels described in section 9.1: the low (40 GHz) and high (402 GHz) maps, containing CMB, noise, and foregrounds. We choose these maps because

they are the most foreground-dominated frequencies available in our experimental setup, with synchrotron and dust dominating the emission at 40 GHz and 402 GHz, respectively, as we listed in the first Part of the Thesis (section 4.2). If the foreground residuals in `CMBsol` are responsible for the bias on the tensor-to-scalar ratio, we expect to find some correlation with structures present in foregrounds-dominated maps, which are instead absent in fully Gaussian simulations. Consequently, we expect to register incompatible ST cross-coefficients.

Ideally, one would like to carry out the above correlation with foregrounds-only sky maps, where there is no other signal than dust or synchrotron. In general, this is not possible for a real experiment, even if component separation techniques able to remove a part of the CMB and noise contributions are available [169]. An alternative is to use external tracers of foreground emission as the correlation target, for instance exploiting *WMAP* [23] or *Planck* [172] low/high frequency maps. However, these maps are expected to have a larger noise contribution with respect to the input multi-frequency sky maps internal to the considered experiment. Anyway, here we choose to not adopt external data to compute cross-coefficients, as our foregrounds simulations generated with the `PySM` package heavily rely on the information contained in *Planck* and *WMAP* observations, with the risk of invalidating the analysis.

But we would like to stress that it is always possible to use the input data collected by the same experiment as a correlating field, even if they contain some "contamination" due to CMB and noise. Specifically, by construction of the component separation algorithm, the CMB and the noise residuals contained in `CMBsol` will be correlated with the original CMB and noise present in the data, since the separation outputs are just a linear combination of the input sky maps. The noise correlation will be more evident for the lowest and highest frequencies (40 and 402 GHz), being them the channels with the larger noise level (see table 9.1 or figure 9.4). These spurious features are instead absent when we correlate with Gaussian simulations, since they contain different realisations of CMB and noise. Therefore, in the cross-coefficients case we will register some failures of the robustness test (false positives) not due to foregrounds contamination, but induced by CMB and noise correlations. The impact of these false positives is quantified in the validation phase of the test, reported in chapter 10.

Figures 9.15 and 9.16 show an example of the computed S_1, S_2, and S_4 coefficients respectively, computed on the three maps considered for the robustness test (`CMBsol`, `CMBtens`, `FGgauss`), where the Galactic plane has been masked. Here the adopted foreground model is `d10s5`, and the input tensor-to-scalar ratio is $r_{\text{in}} = 0$, of which recovery is biased at the 10^{-3} level. The former figure shows both the auto-

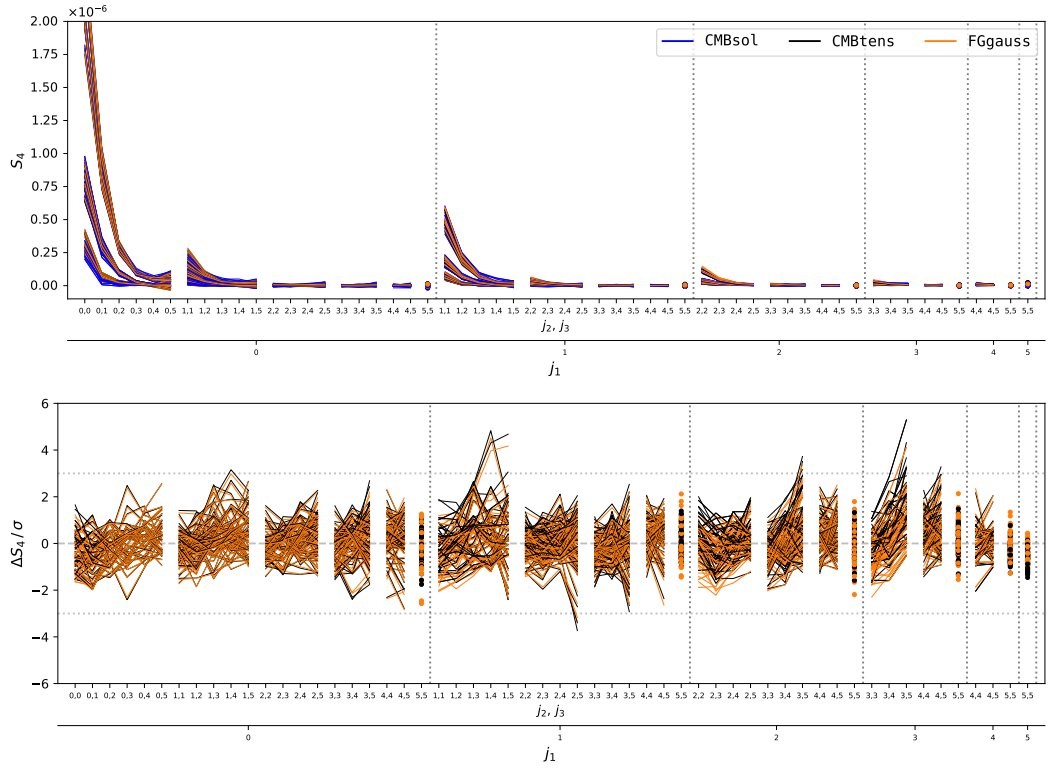


Figure 9.16. S_4 auto-coefficient for the MC-NILC - `d10s5` - $r_{\text{in}} = 0$ scenario. Colour code is the same as figure 9.15. Lower plot shows the difference between `CMBsol` and Gaussian simulations, in standard deviation units, with the dotted lines indicating the $\pm 3\sigma$ region.

and cross-coefficients, with the latter obtained by correlating with 40 GHz input maps. In all the shown auto-coefficients, the noise contribution is evident at small scales (low j), while the slight rise observed at larger scales, which is not present in the Gaussian simulations, is induced by the foregrounds non-Gaussianity.

9.7 Compatibility analysis

The goal of the robustness test is to establish the compatibility between summary statistics computed on different maps. If the statistics are compatible, foregrounds non-Gaussianity is not recognised, the test is passed, and no bias on the tensor-to-scalar ratio is identified. If the statistics are not compatible, then the test is failed, and the bias on r is identified as induced by foreground residuals. Schematically,

we compare:

$$\begin{aligned} \text{MFs}\,(\texttt{CMBsol}) \quad &\text{vs} \quad \text{MFs}\,(\texttt{CMBtens}) \\ S\,(\texttt{CMBsol}) \quad &\text{vs} \quad S\,(\texttt{CMBtens}) \\ S^{\times}\,(\texttt{CMBsol} \times 40\ \text{GHz}) \quad &\text{vs} \quad S^{\times}\,(\texttt{CMBtens} \times 40\ \text{GHz}) \\ S^{\times}\,(\texttt{CMBsol} \times 402\ \text{GHz}) \quad &\text{vs} \quad S^{\times}\,(\texttt{CMBtens} \times 402\ \text{GHz})\,, \end{aligned} \tag{9.26}$$

where S stands for S_1, S_2, S_3 or S_4, and the same is done also for the `FGgauss` simulations. Depending on the considered statistics, we carry out two slightly different compatibility analysis, one for MFs and the other for ST coefficients, which we now describe. For the remainder of this section, we will use "Gaussian simulations" to refer either to `CMBtens` or `FGgauss` maps, as we carry out the compatibility analysis in the same way for the two sets. We recall that we have 300 realisations of `CMBsol` maps, and for each one of them we generated other 5000 simulations of Gaussian (`CMBtens` or `FGgauss`) maps.

9.7.1 For Minkowski functionals

We follow an approach similar to the one adopted by the *Planck* team in [171], where a thorough investigation of the statistical properties of CMB anisotropies was conducted on real *Planck* data. In the same spirit, they presented a comparison of the MFs of a single component-separated map against those of a set of Gaussian simulations. Here we apply a similar analysis, but repeated for each of the 300 simulated component-separated maps.

MFs are evaluated for 12 thresholds u between -3 and 3 in σ units, on maps normalised by their variance. To quantify the level of deviation between MFs of different maps, χ^2_s values for each simulation are computed assuming a Gaussian likelihood for the MFs at every threshold:

$$\chi^2_s = \left[v_s - \bar{v}_s^{\text{sim}}\right]^{\mathrm{T}} \Sigma_s^{-1} \left[v_s - \bar{v}_s^{\text{sim}}\right], \qquad s = 1, \ldots, 300\,. \tag{9.27}$$

where v_s represents one of the three MFs ($v = v_0, v_1, v_2$) of the s-th `CMBsol` realisation, and $\bar{v}_s^{\text{sim}} \equiv \langle v_s^{\text{sim}} \rangle$ is the mean of the MFs computed on 300 of the 5000 Gaussian simulations generated for that particular `CMBsol` simulation.

Correlations between different thresholds [66] are taken into account using the appropriate covariance matrix Σ_s, different for each `CMBsol` realisation. An unbiased estimator of the inverse of the covariance is given by

$$\Sigma_s^{-1} = \frac{N - d - 2}{N - 1} \mathbf{C}_s^{-1}\,, \tag{9.28}$$

where $N = 300$ is the number of `CMBsol` simulations, $d = 12$ is the number of thresholds (degrees of freedom), and $\mathbf{C}_s$ is the sample covariance computed from the 300 Gaussian simulations:

$$\mathbf{C} = C_{ij} \equiv \langle (v_i^{\text{sim}} - \bar{v}_i^{\text{sim}})(v_j^{\text{sim}} - \bar{v}_j^{\text{sim}}) \rangle \,, \tag{9.29}$$

with i, j referring to MFs computed at different thresholds. The pre-factor in eq. (9.28) is known as the Hartlap factor [98] and it is needed for the estimator to be unbiased. We adopt only 300 of the 5000 Gaussian realisations because they are enough for the covariance to be unbiased and converge to the sample covariance, since we have a low number of statistics (12 thresholds). The Hartlap factor is close to 0.96.

For each `CMBsol` realisation we compute its probability-to-exceed (PTE, or equivalently, the p-value), which is the amount of the χ^2_{Gauss} distribution of the MFs computed on Gaussian simulation exceeding the observed χ^2_s:

$$\text{PTE}_s = \chi^2_{\text{Gauss}} > \chi^2_s. \tag{9.30}$$

We say that the MFs of the s-th simulation are not compatible when $\text{PTE}_s < 0.05$, for at least one of the MFs. Incompatibility means that statistically significant deviations are observed between the MFs computed on a map contaminated by non-Gaussian residuals and the mean MFs of Gaussian realisations with similar power. Many incompatibilities over a large number of simulations would indicate that MFs are a reliable robustness test; vice versa, a small amount of observed deviations would suggest that MFs are not able, most of the times, to distinguish between Gaussian and non-Gaussian contributions. We note that we could also combine the three MFs into a single vector and run a combined compatibility analysis; here we prefer to keep them separate to check the impact of the single quantities in identifying non-Gaussianity.

We define the test efficiency η as the percentage of registered incompatible simulations:

$$\eta = 100 \times \frac{\text{number of failing sims}}{\text{number of sims } N} = 100 \times \frac{\#\left[\text{PTE} < 0.05\right]}{N} \,, \tag{9.31}$$

where $\#[\cdot]$ stands for the number of simulations satisfying the condition in brackets. We will report this number η to quantify the ability of MFs as a robustness test, for each combination of component separation and foregrounds model, and using the two sets of Gaussian simulations as reference.

9.7.2 For ST coefficients

In the case of scattering coefficients, the compatibility analysis is slightly different from the previous one. The main reason is the larger dimension of the covariance matrix Σ coming from the ST statistics of the Gaussian simulations; evaluating the χ^2 requires inverting this covariance. Estimating and inverting a $D \times D$ covariance from a finite number N of realisations is reliable only when $N \gg D$: as D approaches N, the estimate becomes ill-conditioned, and its inverse will be strongly biased [98]. The raw ST statistics, especially when S_3 and S_4 are included, have a dimension D that is too large even when 5000 Gaussian simulations are considered. For instance, for the S_4 coefficient we have three oriented scales sampling four different directions each, leading to 3584 numbers to compare. For MFs, they were 12 for each one of them. Moreover, this kind of compatibility analysis aggregates all the available coefficients into a single scalar quantity. This means that coefficients that carry a little amount of interesting information for our purpose will only inflate the variance of the estimator, diluting the ability of the robustness test in identifying the faint non-Gaussian features.

We therefore reduce the dimensionality of the statistics with two goals: first, to obtain stable, well-conditioned estimates of the covariance Σ; and second, to retain the coefficients most likely to carry information about the non-Gaussian features of the `CMBsol` map. Each ST coefficient depends on one (for S_1 and S_2) or several (for S_3 and S_4) scales j, as well as on angles θ. This dependence on scale and angle is typically smooth, and we exploit this regularity to compress the representation [10]. We Fourier transform the coefficients over the angles and keep only the zeroth mode and the first harmonic. For S_3 and S_4, we additionally retain only those coefficients whose scales lie close together: a coefficient depending on two scales j_1 and j_2 is discarded whenever $|j_1 - j_2| > d_j$, with the threshold set to $d_j = 4$. As an example, the number of values entering the compatibility analysis goes from 3584 to 512 for the S_4 coefficient.

After this dimensionality reduction we quantify the compatibility of `CMBsol` with a given Gaussian reference set in the same way as we did for MFs in the previous section, using eq. (9.27):

$$T_s = \left[S_s - \bar{S}_s^{\mathsf{sim}}\right]^{\mathrm{T}} \Sigma_s^{-1} \left[S_s - \bar{S}_s^{\mathsf{sim}}\right] , \qquad s = 1, \ldots, 300 \, . \tag{9.32}$$

The S_s stands for the ST coefficients S_1, S_2, S_3, and S_4, while the mean is computed over 5000 simulations. The first two coefficients S_1 and S_2 can be combined into a single array of values $S_{1,2} = S_1 \cup S_2$. To convert the T_s into a PTE (or p-value), we need to know its distribution under the Gaussian null hypothesis. For Gaussian

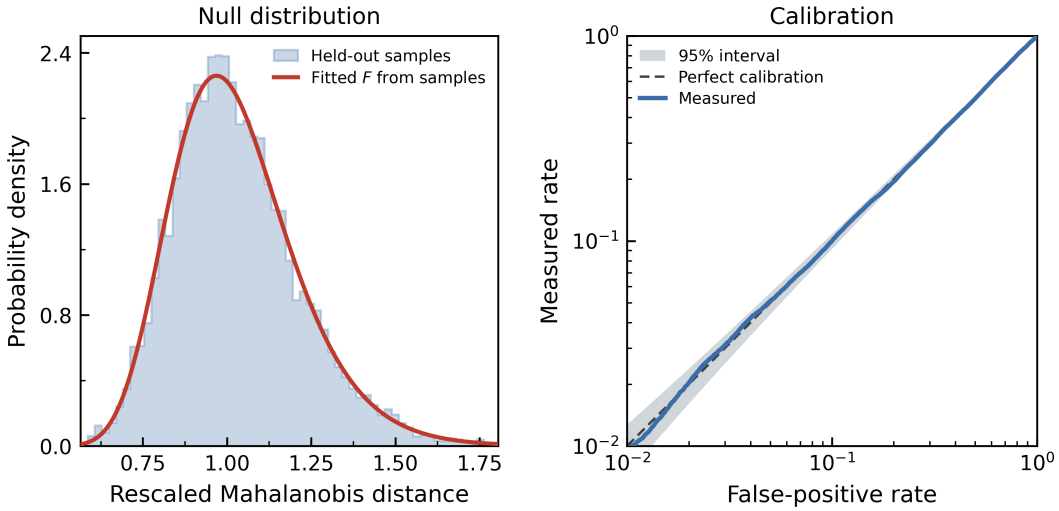


Figure 9.17. Check of the calibration of the compatibility analysis for the ST-based test, on held-out reference simulations not used in the fit. The considered foreground model is `d0s0`, with $r_{\rm in} = 0$, using the S_1 and S_2 coefficients combined together. *Left:* rescaled Mahalanobis distances against the fitted F distribution. *Right:* measured rejection rate versus theoretical false-positive rate (blue), with the perfect-calibration diagonal (dashed) and its 95% interval (grey).

features this would be a χ^2 distribution, as we implicitly assumed for the MFs. However, the ST statistics of the reference maps are themselves non-Gaussian and heavier-tailed than a multivariate normal, even though the maps are Gaussian random fields. This effect thickens the tail of T_s; adopting a χ^2 reference would underestimate this feature and produce an overconfident, mis-calibrated test. To mitigate this, we calibrate the null distribution of T empirically using an F distribution fitted to the features of the reference Gaussian maps. We can then say whether a `CMBsol` map is incompatible with the reference Gaussian simulations when the p-value is lower than 0.05. Exactly as we did for MFs, we count how many times we register incompatibilities, and report the percentage as the efficiency of the robustness test η.

In order to verify the correct calibration of the compatibility analysis, we carry out the same process on held-out reference realisations that were not used to estimate the mean and covariance, or the fitted F parameters. The left panel of figure 9.17 compares the distribution of their rescaled Mahalanobis distances with the fitted F density: the two agree over the full range, including the upper tail that sets the p-values. The right panel shows the measured rejection rate as a function of the theoretical false-positive rate: it follows the diagonal within the 95% interval expected from the finite number of held-out realisations, down to a false-positive rate of 10^{-2}. The test is therefore correctly calibrated at the significance levels we use.

Chapter 10

Results of the tests characterisation

The aim of this chapter is to present a comprehensive view on the characterisation of our proposed robustness tests for the detection of primordial B-modes from CMB observations, reporting the results obtained in the experimental scenarios we described in the previous chapter. These include foreground models of different complexity (`d0s0`, `d1s1`, `d10s5`), different component separation methods (NILC and MC-NILC), and different input cosmological signals ($r_{\text{in}} = 0$ or $r_{\text{in}} = 5 \times 10^{-3}$). We will independently discuss the outcome of using the two different statistics, namely Minkowski functionals and scattering transforms, and successively compare them. The main descriptor of the results is the efficiency of the test η, defined in eq. (9.31) as the percentage of simulations in which a bias on the tensor-to-scalar ratio r has been identified. A value of η close to 100% may suggest that the summary statistics adopted for the robustness test can confidently recognise a bias on r, while a lower value indicates that the test is not sensitive enough to be reliable in the identification of a false detection induced by foregrounds. In the compatibility analysis from which η is derived (section 9.7), we compare the statistics of the CMB map recovered by component separation `CMBsol` with two sets of Gaussian simulations, `CMBtens` and `FGgauss`. We report results for both comparisons. Once again, this chapter closely follows what is contained in two of the papers I worked on during my PhD, [178] and [179], with a somewhat extended discussion and more detailed explanations.

10.1 Minkowski functionals test

Back in figure 9.14, we already showed an example of how the MFs look like both in pixel and needlet space; we computed the statistics of the component-separated

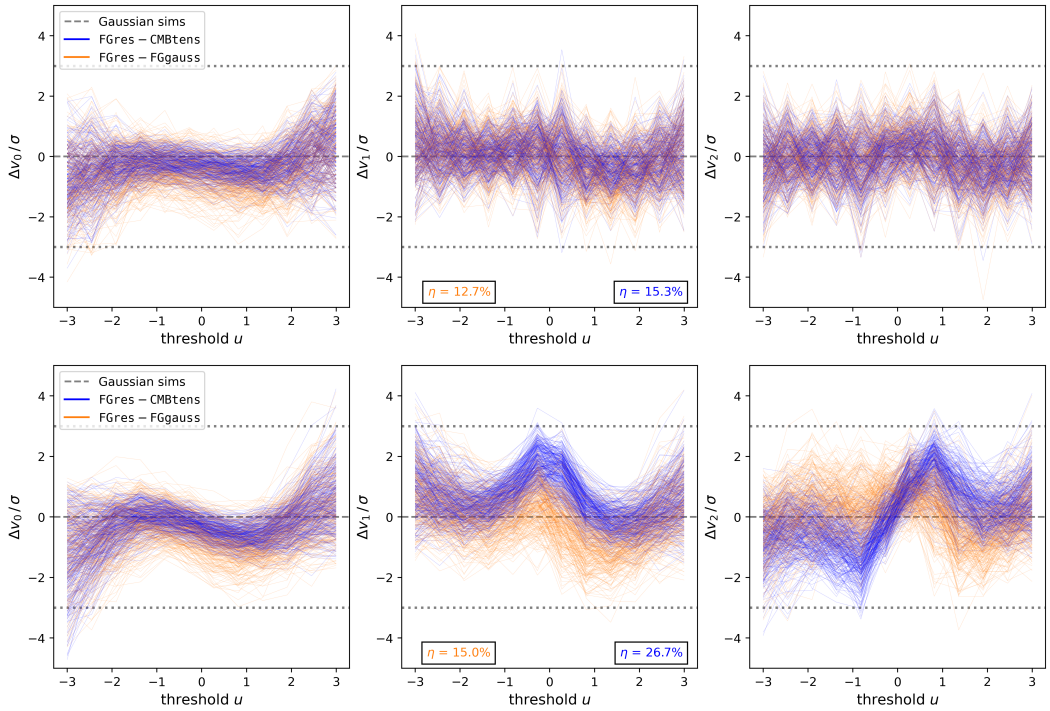


Figure 10.1. MFs weighted difference for the MC-NILC - `d10s5` scenario. Each line is one of the 300 input realisations. Blue (orange) lines indicate `CMBtens` (`FGgauss`) as the reference Gaussian simulations. Plots are shown for all multipoles (*upper row*) and for large scales only (*lower row*) analysis. Dotted lines delimit the $\pm 3\sigma$ region. Efficiencies are reported in the lower corners of the middle panels.

CMB map `CMBsol` (in blue) and the average MFs of 300 Gaussian simulations `CMBtens` (in black) and `FGgauss` (in red). Additionally, here in figure 10.1 we report the normalised difference $\Delta v/\sigma = \left(v - \bar{v}^{\rm sim}\right)/\sigma$ between the MFs of `CMBsol` and the averaged MFs of the Gaussian simulations, for all the 300 input sky realisations, weighted by the standard deviation computed across simulations. In both figures we are showcasing the use of the most advanced component separation method MC-NILC, applied on the most complex foreground model `d10s5`, where no input primordial B-modes are present ($r_{\rm in} = 0$). Tables 10.1 and 10.2 offer a complete overview on the characterisation of the MFs-based robustness test, for the different cases investigated in this manuscript. The former table report results for the pixel space analysis, while in the latter only the largest scales are considered.

Concerning the real-space analysis, we find a high test efficiency when comparing MFs of `CMBsol` against `CMBtens`, while using NILC on the `d1s1` and `d10s5` models, with $\eta = 95\%$ and 88%, respectively. This means that over 300 simulations, MFs are able to identify a bias on r in 95% (or 88%) of the cases, detecting deviations from Gaussianity induced by foregrounds contamination. However, these discrepancies are mainly driven by the large differences at the power spectrum

Table 10.1. Efficiency η (in %) of the MFs-based robustness test on polarisation B-mode maps, for the different scenarios considered in the study. Values are reported for a pixel-space analysis (i.e., including all angular scales), both for no delensing ($A_{\rm lens} = 1.0$) or 50% delensing ($A_{\rm lens} = 0.5$) applied.

no delensing ($A_{\rm lens} = 1$)				
	reference sims	d0s0	d1s1	d10s5
NILC	CMBtens	13	95	88
	FGgauss	11	39	27
MC-NILC	CMBtens		13	15
	FGgauss		10	13
50% delensing ($A_{\rm lens} = 0.50$)				
NILC	CMBtens	10	100	95
	FGgauss	11	43	28
MC-NILC	CMBtens		14	20
	FGgauss		15	24

level between the two maps. As discussed in section 9.5, the MFs expected values depend both on the level of Gaussianity in the map, and on the shape of the power spectrum included in the μ term of eq. (9.11). MFs computed on maps with mismatching power spectrum features will deviate between each other, with discrepancies sourced by both contributions. As already pointed out in section 9.4.2, the large-scale shapes of the power spectra do not match, at least for the two scenarios we are discussing in this paragraph. For comparison, the percentages of simulations providing a bad fit are 98% for d1s1 and 74% for d10s5. Hence, one could already rule out the cosmological origin of the signal by evaluating the goodness-of-fit, and the hypothesis would then be confirmed by the MFs test. It is worth noticing that in some cases, such as NILC - d10s5, we find that the MFs robustness test is more powerful than the simpler χ^2 analysis.

The same conclusion can also be drawn from the lower value of the efficiency when FGgauss simulations are considered. We find $\eta = 39\%$ and 27% for d1s1 and d10s5, respectively, after the application of the NILC algorithm. This amount of identified biases is entirely due to the different statistics between the two datasets, since the corresponding power spectra are equivalent. This interpretation of the results is also confirmed by the efficiencies obtained in the case of the MC-NILC method, where the large-scale features of the power spectra are more similar: the number of detections is relatively low, and close to what we find in the d0s0 validation scenario, where the tensor-to-scalar ratio is not biased and for which we do not expect a significantly high efficiency.

In the needlet-based analysis, we compute MFs on the same datasets as before,

Table 10.2. Same as table 10.1 but in needlet space, retaining only the largest scales in the analysis ($\ell \lesssim 60$).

no delensing ($A_{\text{lens}} = 1$)				
	reference sims	d0s0	d1s1	d10s5
NILC	CMBtens	11	93	52
	FGgauss	12	73	23
MC-NILC	CMBtens		22	27
	FGgauss		16	15
50% delensing ($A_{\text{lens}} = 0.5$)				
NILC	CMBtens	12	96	57
	FGgauss	11	75	20
MC-NILC	CMBtens		24	42
	FGgauss		18	18

but now we first filter the maps with a needlet band that smooths out the higher multipoles. This means that we are focusing our analysis on the largest angular scales in the sky ($\ell \lesssim 60$), since we expect the most relevant sources of bias on r to be in this multipole range. In particular, foreground residuals tend to dominate the sky emission at large scales, over the CMB and noise residuals components. Consequently, non-Gaussian features typical of foregrounds should be enhanced and easier to identify by MFs, while the Gaussian contribution from noise is filtered out. In general, we observe a slight increase in the efficiency for some of the cases (e.g., for MC-NILC), while in others the efficiency remains the same or is even reduced. This is due to the fact that in the needlet analysis, the r_{fgs} adopted for generating the Gaussian simulations is fitted using only the lower multipoles included in the band, providing a better power spectrum fit and thus reducing spectral shape deviations. Still, the efficiency remains considerably low and close to the "no detection" level typical of the d0s0 benchmark.

We can explain why MFs can not clearly detect non-Gaussianity by looking at the compared maps, in figures 9.12 and 9.13. The Gaussian component (made by CMB lensing and noise residuals) is dominating the maps structures, hiding the lower contribution coming from non-Gaussian foreground residuals. The level of foreground residuals is too low with respect to the noise residuals in order for it to be identified by MFs. This is easily visualised in figure 10.1: in the majority of the simulations, the MFs weighted difference is enclosed in the $\pm 3\sigma$ region, with σ being the standard deviation of the Gaussian simulations. Even if we can recognise a pattern in the difference suggesting an underlying deviation, the variance of the estimator is just too high to reliably identify it. We recall that in a real experiment, we would only have one of the 300 lines drawn in the figure, and we would probably

deem the MFs as compatible, passing the test. For a different example, in the NILC - `d1s1` case foreground residuals are relatively larger than noise residuals with respect to other scenarios, and we indeed register a higher number of detections.

The above tentative explanation can be straightforwardly tested by repeating the analysis after the removal of CMB and noise contributions. In practice, we can take figure 9.12 and look at discrepancies between the non-Gaussian (lower left map) and Gaussian (lower middle map) foreground residuals. In this case, the efficiency of the MFs test is always 100%, meaning that MFs are perfectly able to distinguish the non-Gaussianity. Once added, CMB lensing and noise residuals become the driving contribution in the MFs and tend to overwhelm the foregrounds, hindering the power of the robustness test.

To better understand the relative contribution of CMB lensing and noise residuals, we repeat the same analysis while applying two independent delensing and denoising factors on the CMB and noise residuals components, respectively, after the component separation (see section 9.2 for a couple of words on delensing). In this way, we are removing part of the Gaussian contribution at the map level, in order to check which is the dominant component and its impact on the MFs ability to detect non-Gaussianity. Tables 10.1 and 10.2 report results for a realistically attainable delensing of 50% ($A_{\rm lens} = 0.5$), showing that in general we obtain an increase of the efficiency as expected: we reduced one of the Gaussian contribution, and the anisotropy induced by foreground residuals is more evident to MFs. However, the detection number in the most realistic scenario (MC-NILC - `d10s5`) is still considerably low, being $< 30\%$. For this particular case only, we also rescale the noise residuals maps by a denoising factor of 75% ($A_{\rm noise} = 0.25$), observing an efficiency raise to 93% and 92% comparing with `CMBtens` and `FGgauss` maps, respectively, in pixel space. In the needlet-based analysis, the efficiencies reach $\sim 99\%$ in both cases. This confirms the role of noise residuals as the dominant component over CMB lensing, driving the Gaussian contribution. It also provides the noise level needed for the MFs to be considered as reliable tools for this robustness test.

It is also worth to discuss the `d10s5` scenario, which at first sight provides a counter-intuitive result: in general, we expect that an increase in foregrounds complexity should correspond to an increase in the number of identified non-Gaussian features, with respect to the simpler `d1s1` case, since foreground residuals should be more difficult to clean and thus more significant. Instead, we obtain a lower amount of detections for `d10s5`, both in pixel- and needlet-based analysis. We explain this fact by considering that in general, the application of a minimum-variance method to a more complex foreground model leads to the increase of both foreground and

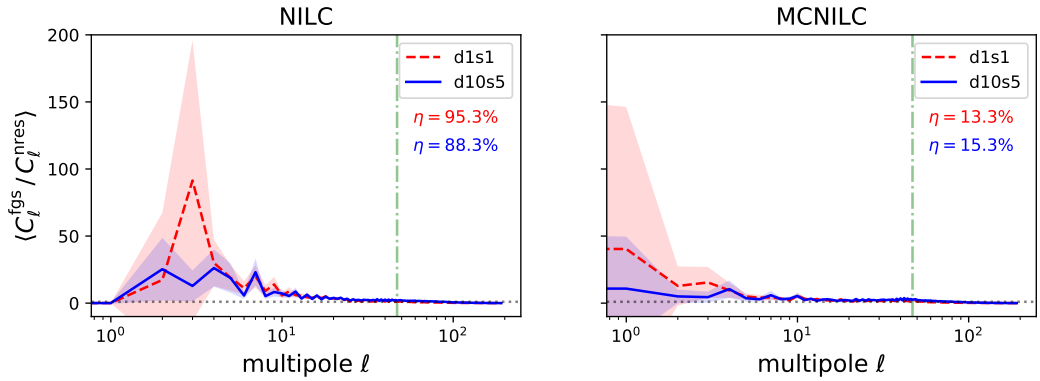


Figure 10.2. Average ratio between spectra of foreground and noise residuals. Dashed red lines indicate the `d1s1` model, while blue solid lines are for `d10s5`. Shaded areas represent the standard deviation across 300 simulations. MFs test efficiencies are also reported for comparison. The vertical line indicates the smallest scale included in the first needlet band b_0.

noise residuals. It means that the CMB solution will contain more contamination due to foreground residuals, but it will also carry a stronger Gaussian component driven by noise residuals, which is responsible for reducing the MFs ability to identify foregrounds (as discussed in the previous paragraph). To further explore this hypothesis, for each simulation we evaluate the relative amount of residuals by computing the ratio between the power spectra of foreground ($C_\ell^{\rm fgs}$) and noise ($C_\ell^{\rm nres}$) residuals. Figure 10.2 reports the mean of this quantity across simulations: for the NILC algorithm (left panel) the average ratio in the `d10s5` case is slightly lower than `d1s1`, meaning that the separation between foregrounds and noise is less clear for MFs to detect, leading to a lower η. For the MC-NILC case, the residual ratio is similar for the two models, leading to roughly the same efficiency. These considerations suggest that the MFs efficiency is not heavily dependent on the complexity of the foreground model, but rather on the relative levels of foregrounds and noise residuals, which can both increase or decrease based on the performance of the component separation.

The last point we make in this section is about some approximation we did in our simulations. As mentioned back in section 9.1, we ignored the eventual artifacts introduced in B-mode maps by the harmonic transformation from the originally simulated Q/U sky maps. We decided to blindly propagate this contribution through the pipeline. In any case, its missing mitigation (before component separation) should increase the level of non-Gaussianity in the maps with respect to the reference Gaussian simulations, meaning that the number of identified false detections (i.e., the efficiency of the test η) is expected to be overestimated. However, our results suggest that even in this pessimistic scenario, MFs are not able to reliably flag a

false detection of the tensor-to-scalar ratio; therefore, removing this contamination would only reduce the efficiency of the test, not changing the overall outcome.

To summarise the results described in this section, we concluded that in our experimental setup, MFs are not able to reliably identify a bias on the measurement of the tensor-to-scalar ratio, induced by the residual contamination of Galactic foregrounds. The level of noise residuals is masking the anisotropic non-Gaussian contribution due to foreground residuals, preventing MFs from correctly identifying the contamination. The false cosmological signal passes the MFs test in the majority of the simulations, while not providing additional information on the astrophysical nature of the contamination. Thus, MFs can not be reliably used as higher-order statistical tool for a robustness test in the event of a future tensor-to-scalar ratio detection, at least if we consider a *LiteBIRD*-like experimental configuration. However, they may still be useful in scenarios where foreground residuals are relatively more dominant than noise residuals.

10.2 Scattering transform test

In the previous chapter we presented figures 9.15 and 9.16, where we showed an example of the typical ST coefficients evaluated on the three different sets of maps considered in this work. As for MFs, we report the outcomes of the compatibility analysis in terms of the efficiency η for the combined S_1 and S_2 in table 10.3, and for S_4 in table 10.4. The "auto" columns list results coming from the adoption of auto-coefficients, while the $\times$ sign indicates cross-correlation with foregrounds-dominated frequency maps. We also consider a delensed CMB contribution, reduced by a factor $A_{\text{lens}} = 0.5$. We notice that here, differently from the MFs analysis, we dropped the `d1s1` foreground model, as it is not expected to give additional information with respect to the other models. We are first going to report results obtained using the simpler S_1 and S_2 coefficients, and then the higher-order S_4. The S_3 coefficient is not reported in this work, as we have verified its performances and concluded that it is not helpful to characterise deviation from Gaussianity, at least in the context of CMB robustness tests.

In the MC-NILC - `d10s5` - $r_{\text{in}} = 0$ scenario we obtain a bias on the tensor-to-scalar ratio of $\delta(r) \sim 3 \times 10^{-3}$, being the same order of magnitude for all the 300 realisations. Using the S_1 and S_2 auto-coefficients, the ST test identifies the bias in 52% and 29% of the simulations when `CMBsol` is compared with the `CMBtens` and `FGgauss` maps, respectively. The efficiency improves to 66% and 45% when some of the CMB lensing Gaussian contribution is reduced, sharpening the foreground residuals features. Similarly to what we saw for MFs, we register a lower efficiency

Table 10.3. Efficiency η (in %) of the ST-based robustness test on polarisation B-mode maps, for the different scenarios considered in the study. Results are shown for the S_1 and S_2 coefficients combined together, between scales $j_{\text{min}} = 1$ and $j_{\text{max}} = 4$.

		MC-NILC			
	reference sims	auto $A_{\text{lens}} = 1$	auto $A_{\text{lens}} = 0.5$	$\times$ 402 GHz	$\times$ 40 GHz
d0s0	CMBtens	5	7	6	8
$r_{\text{in}} = 0$	FGgauss	5	6	6	8
no fgs	CMBtens	6	-	6	16
$r_{\text{in}} > 0$	FGgauss	6	-	5	15
d10s5	CMBtens	52	66	45	93
$r_{\text{in}} = 0$	FGgauss	29	45	50	92
		NILC			
d0s0	CMBtens	4	6	5	13
$r_{\text{in}} = 0$	FGgauss	4	6	5	13
d10s5	CMBtens	100	100	96	98
$r_{\text{in}} = 0$	FGgauss	88	93	81	90

for FGgauss simulations, being them more similar to CMBtens at the power spectrum level. Thus, in this scenario auto-coefficients are not reliable enough to consistently identify the bias on r, as the test only raises a flag in roughly half of the simulations. The same is true if we look at the cross-coefficients computed by correlating with the dust-dominated data at 402 GHz, with the efficiency being around 50%. However, the test becomes much more powerful when we cross-correlate with synchrotron-dominated data at 40 GHz, being $\eta \sim 92\%$ both when comparing with CMBtens and FGgauss. This means that in case of a biased measurement of r of the order of $\delta(r) \sim 10^{-3}$, the ST-based robustness test can identify it in almost the totality of the simulations, raising a flag on the presence of synchrotron-correlated foreground residuals.

Compared with MC-NILC, the NILC technique leaves more foregrounds contamination on the largest scales, leading to a higher bias at the $\delta(r) \gtrsim 5 \times 10^{-3}$ level. In this case, both the auto- and cross-coefficients are able to identify the bias in a good number of simulations, with the efficiency reaching $\sim 90\%$ and even 100% in the FGgauss and CMBtens comparisons, respectively. We also notice that ST-based test performs way better than the simpler robustness test provided by the goodness-of-fit analysis described in section 9.3, which was able to recognise a bias in $\sim 70\%$ of the simulations, in the same scenario. It thus guarantees a more reliable statistical analysis for the identification of foreground residuals. We recall

that even if the NILC implementation is less optimised than MC-NILC, it will still be one of the primary component separation in current and future experiments like *SO* [6, 218] and *LiteBIRD* [127].

Since ST seems to be more promising in the context of robustness tests, we carry out a more detailed validation phase with respect to the MFs one. A thorough validation of the robustness test is a crucial step for its characterisation and for trusting the above results. Here we consider three different validation scenarios:

1. We adopt the simpler foregrounds model `d0s0`. Component separation can easily mitigate this contamination, and the input tensor-to-scalar ratio $r_{\text{in}} = 0$ is recovered with small or no bias, of the order of $\delta(r) \lesssim 10^{-4}$. In this case the efficiency of the test is around $\sim 5 - 10\%$ for both MC-NILC and NILC, for all scenarios, with the small number of incompatibilities due to chance or noise correlations in the samples. Thus, if r is not biased, the robustness test is passed and no flag is raised. We assume the $5 - 10\%$ range as the benchmark value for the "no bias" case. The left plot of figure 10.3 showcases an example of these compatible S_1 coefficients.

2. We consider a scenario where, in addition to the B-modes induced by gravitational lensing, a realistic primordial B-mode signal is also included in the input simulations, as described in section 9.1. The injected tensor-to-scalar ratio is $r_{\text{in}} = 5 \times 10^{-3}$. In order to recover an unbiased estimate of r_{in}, we manually remove the foreground residuals contamination from the component separated CMB maps `CMBsol`. We recall that, since we are working with simulated data, we have access to all the isolated contribution to the `CMBsol` map. In this scenario, the efficiency is around $\sim 5 - 10\%$, similar to the previous validation case. This certifies that the incompatibilities measured in the `d1s5` scenario are entirely due to the presence of foreground residuals, and not induced by other components. The right plot of figure 10.3 reports an example of the compatible S_1 coefficients.

   The slightly higher number of false positives registered in the cross-40 GHz coefficients (15%) quantifies the impact of the correlation between the input data and the recovered CMB and noise residuals, which we discussed in section 9.6. Since the component separation algorithm linearly combines the input multi-frequency data with some weights, the CMB and noise residuals in `CMBsol` are necessarily correlated with the CMB and noise of the input maps. This contribution is instead absent in the correlation with Gaussian simulations, as they are built with different noise realisations which are uncorrelated with the input data. Thus, compared to the previous $r_{\text{in}} = 0$ case,

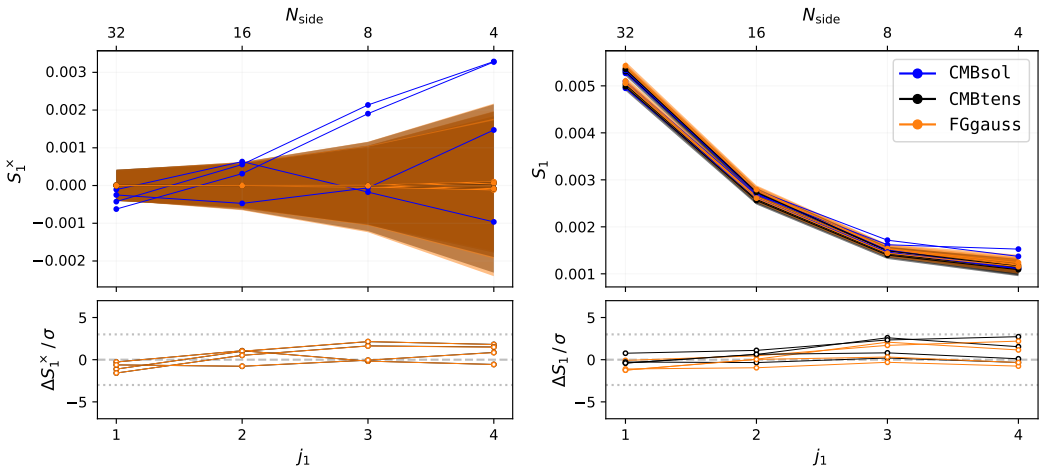


Figure 10.3. ST S_1 coefficient for two validation cases, with the same colour code of figure 9.15. *Left:* cross-coefficients correlating with synchrotron-dominated input data, for the MC-NILC - `d0s0` - $r_{\text{in}} = 0$ case. *Right:* auto-coefficients, for the NILC - `d0s0` - $r_{\text{in}} = 5 \times 10^{-3}$ scenario. In both plots, the tensor-to-scalar ratio estimate is not biased, coefficients are compatible, and the robustness test raises no strong warnings.

we now have more false positives because there is an additional primordial signal in the CMB, strengthening the correlation. This effect can be visualised in figure 10.4, where we use the cross angular power spectrum to quantify the correlation between noise residuals and the noise in the input data, for different frequencies.

The net impact of this contamination is a resulting efficiency of 100% even for the `d0s0` model, where foreground residuals are completely mitigated. In our case, this extra correlation is dominant at smaller angular scales (high multipoles). To overcome this problem, we restrict the compatibility analysis to not include the smaller scales contaminated by this feature, which translates to dropping the $j = 0$ scale. We notice that in using this range of scales, we are still retaining some noise correlation in the summary statistics; we can not drop all the impacted scales as we would lose too much sensitivity for the robustness test. In any case, this just contributes in adding a few percent to the test efficiencies in the `d0s0` validation scenario (see again table 10.3), slightly raising the benchmark values above 10%, but not changing the final conclusions of the study.

3. To further confirm that the test is really flagging the presence of foregrounds, we repeat the analysis in the MC-NILC - `d10s5` - $r_{\text{in}} = 0$ scenario, computing cross-coefficients with the `PySM` pure-synchrotron maps at 40 GHz, without CMB and noise. In this way, we only keep the foregrounds correlation, while

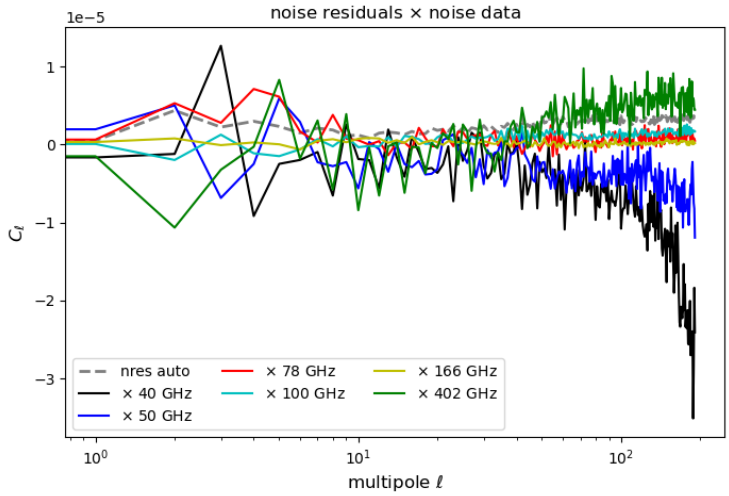


Figure 10.4. Cross angular power spectrum between noise residuals and noise maps, for different input frequency channels (in different colours), one simulation. The dashed line indicates the noise residuals auto spectrum. The considered scenario here is MC-NILC applied on the `d10s5` model. We notice the strong negative correlation for the 40 GHz channel.

avoiding CMB and noise correlations. We apply a rescaling factor $A_{\text{fgs}} < 1$ of different values to the foreground residuals contained in `CMBsol` (lower left map of figure 9.12). In this way, we gradually lower the amplitude of the foreground residuals and, consequently, the r bias. If the incompatibilities detected by the test are only due to foreground residuals, then the efficiency should decrease and approach the $\sim 5 - 10\%$ range. This step also provides the sensitivity of the test as a function of the bias on r: for each decreasing value of A_{fgs}, the bias becomes lower, and we measure in how many simulations it is identified. The left panel of figure 10.5 shows the PTE (p-value) given by the compatibility analysis (section 9.7) for different values of the r bias, for 50 simulations. If the PTE for a simulation is less than the 0.05 threshold (dashed horizontal line in the figure), contamination has been recognised by the test. From the plot, we see that when the bias is $\delta(r) \gtrsim 1 \times 10^{-3}$, the robustness test can confidently identify it, with PTEs being < 0.05 for almost all simulations. When the bias is lower than $\sim 5 \times 10^{-4}$ the test is not able to confidently flag the bias any more, as the PTEs increasingly exceed the 0.05 threshold. This also shows that foreground residuals are the responsible for the incompatibilities previously reported, as it is the only component we are reducing here, while noise and CMB lensing are the same.

Thus, the validation stage consolidates the reliability of the ST-based robustness test. Results show how the test can confidently warn about the presence of foreground

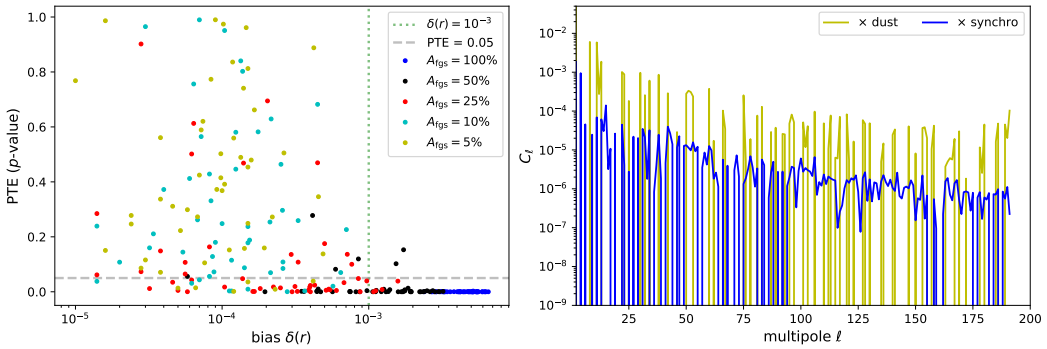


Figure 10.5. The considered scenario is MC-NILC - `d10s5` - $r_{\text{in}} = 0$. *Left:* probability-to-exceed (PTE, or p-value) given by the compatibility analysis for the cross-synchrotron test using the S_1 and S_2 coefficients, as a function of the bias on the tensor-to-scalar ratio r. Results are shown for 50 simulations and for different values of foregrounds rescaling A_{fgs}. The dotted vertical line is a reference value of $\delta(r) = 10^{-3}$, while the horizontal dashed line is the PTE threshold for incompatibility: ST coefficients are compatible above 0.05, and incompatible below, when foreground contamination is detected. *Right:* cross-angular power spectrum between `CMBsol` and foreground maps; cross-correlation with dust is indicated in yellow, while blue is for synchrotron. See text for discussion.

residuals contamination in a biased detection of the tensor-to-scalar ratio, while it raises no strong flags when the estimate is unbiased.

We now turn our attention to the higher-order ST coefficients explored in this analysis, S_3 and S_4. Since they characterise the interaction between two and three oriented scales (see section 9.6 for details), they are expected to give additional information on non-Gaussian features present in our maps, with respect to the simpler S_1 and S_2 [49, 148, 206]. However, as we mentioned at the beginning of the section, we empirically found that in our setup S_3 is not quite useful in distinguishing the faint non-Gaussian signal from a finite set of reference simulations. We thus decided to not report its complete characterisation, but we just mention that it provides an efficiency of $\sim 5\%$ in the case of a biased detection of r, meaning that it can not identify foregrounds contamination.

Table 10.4 reports the test efficiencies for the S_4 coefficient, obtained with the same compatibility analysis, after having reduced its dimensionality (see section 9.7 for details). Compared to S_1 and S_2 in the biased `d10s5` scenario, we measure a similar or slightly higher sensitivity for the auto-coefficients for the `FGgauss` simulations, while the efficiency drops when cross-correlating with foreground-dominated maps. Even if S_3 and S_4 should be (by construction) more sensitive to non-Gaussianity, the larger variance of these estimators makes it difficult to separate fields that are overall dominated by Gaussian components such as CMB lensing and noise. This is very similar to what we saw with MFs. Despite being simpler, S_1 and S_2 were found

Table 10.4. Same as table 10.3, but here results are shown for the S_4 coefficient, after dimensionality reduction.

		MC-NILC			
	reference sims	auto $A_{\rm lens}=1$	auto $A_{\rm lens}=0.5$	$\times$ 402 GHz	$\times$ 40 GHz
d0s0	CMBtens	5	6	1	2
$r_{\rm in}=0$	FGgauss	4	6	1	2
no fgs	CMBtens	5	-	4	4
$r_{\rm in}>0$	FGgauss	3	-	1	3
d10s5	CMBtens	42	52	38	38
$r_{\rm in}=0$	FGgauss	39	56	33	35
		NILC			
d0s0	CMBtens	4	4	1	3
$r_{\rm in}=0$	FGgauss	3	4	1	3
d10s5	CMBtens	100	100	78	75
$r_{\rm in}=0$	FGgauss	89	93	51	48

to be more powerful for the robustness test, in particular for the cross-correlation with foregrounds tracers. Nonetheless, S_4 provides a very good efficiency when using the NILC method, making it a reliable, additional statistics in the robustness test chain, together with S_1 and S_2. The S_4 coefficient is also able to pass the same validation phase we introduced before, with the efficiencies being around 5% or lower.

If we compare the above performances with the MFs ones listed in tables 10.1 and 10.2 of the previous section, we see how ST statistics provide an evident improvement. MFs are able to detect non-Gaussian foreground residuals in just 27% (13%) of the simulations when comparing with FGgauss maps, when NILC (MC-NILC) is applied on the d10s5 model. In the same scenario, they are clearly outperformed by ST statistics, which obtain instead 88% (29%) using auto-coefficients, and $\sim 90\%$ when cross-correlating with low frequency, synchrotron-dominated data. This improvement is mainly due to the combination of multi-scale approach, angular orientation, and low variance of the ST coefficients, which allows them to stand out in the search for non-Gaussian features.

For completeness, we also compare ST performances with other summary statistics that can be used for the same robustness test. For instance, the commonly used isotropic angular power spectrum C_ℓ has a mathematical formalism very similar to the first two ST coefficients. Both descriptor are based on the convolution of the considered fields with some kernel ψ, which is a Fourier basis in plane sine

Table 10.5. Efficiency η (in %) of the C_ℓ-based test on polarisation B-mode maps, for the MC-NILC component separation. Results are shown for angular scales between $\ell_{\text{min}} = 11$ and $\ell_{\text{max}} = 95$ (scales range similar to the one sampled by $1 < j < 4$ in table 10.3).

	reference sims	$\times$ 402 GHz	$\times$ 40 GHz
d0s0	CMBtens	2	12
$r_{\text{in}} = 0$	FGgauss	2	12
no fgs	CMBtens	4	15
$r_{\text{in}} > 0$	FGgauss	2	12
d10s5	CMBtens	10	34
$r_{\text{in}} = 0$	FGgauss	10	27

waves for the isotropic power spectrum, and a more localised wavelet basis for ST. One could object the ST statistics to be just a "fancy" version of the usual angular power spectrum, which could as well be used for the robustness test. We checked this hypothesis, reporting results in table 10.5 about the efficiency of a robustness test built using the isotropic angular power spectrum C_ℓ instead of ST coefficients. In this case we can only rely on the cross-input data power spectra, since the Gaussian simulations are built to have auto-spectra similar to CMBsol. In practice, we evaluate the cross-spectra of CMBsol and of Gaussian simulations with foregrounds-dominated data, and we check if they are compatible or not. If not, it means that the recovered CMB map is contaminated by some non-Gaussian structure which are correlating with the input Galactic foregrounds. Compared to ST coefficients in the MC-NILC - d10s5 scenario, the isotropic power spectrum performs way worse, with the percentage of identified biased simulations being $\sim 30\%$ and 10%, when correlating with synchrotron- or dust-dominated frequency maps, respectively. This clearly shows the advantage of the ST approach: with respect to the simpler isotropic power spectrum, ST coefficients retain precious information about the oriented scales through the angular dependency of the wavelet filter, providing a far greater sensitivity than cross-C_ℓ.

Similarly to what we did for the ST statistics, the compatibility analysis does not include the largest scales ($\ell < 11$) which are variance-dominated, and does not include the smallest scales ($\ell > 95$) to avoid the spurious noise correlation between residuals and input data, which was discussed in section 9.6 and also in the validation phase of this section. The angular scales considered here (in ℓ) are roughly the same as the ones included in the ST compatibility analysis (in j), following the relation $N_{\text{side}} = 64 \cdot 2^{-j}$ and $\ell_{\text{max}} = 3N_{\text{side}} - 1$. Figure 10.6 illustrate the noise correlation feature by showing the weighted difference between the cross-

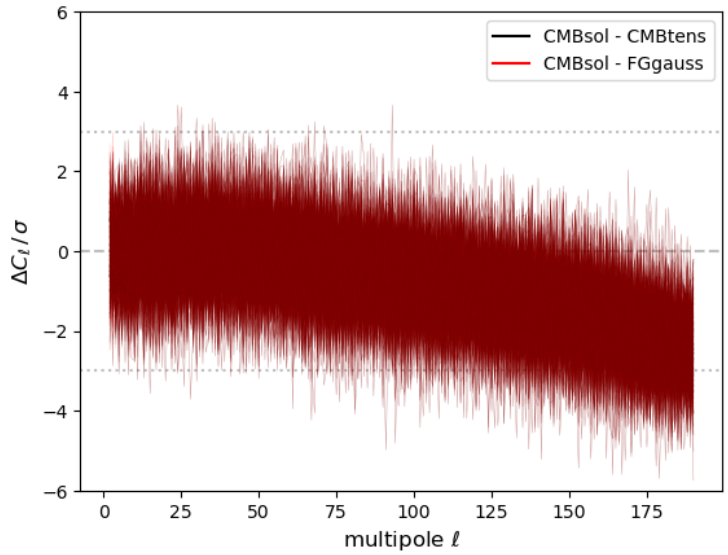


Figure 10.6. Weighted difference between the cross-40 GHz angular power spectrum of `CMBsol` maps and of Gaussian simulations, in units of the standard deviation of the simulations σ. The foregrounds model under study is `d0s0`, and the input tensor-to-scalar ratio $r_{\rm in} = 0$ is correctly recovered without bias. The observed negative trend is due to noise correlation between residuals and input data. See text for details.

spectra of `CMBsol` and of the Gaussian simulations with the 40 GHz input data in the `d0s0` case, for all the 300 realisations and in units of the standard deviations of the simulations σ. When this simple foreground model is considered, foreground residuals are negligible and the r estimate is not biased, and thus we do not expect to detect deviations. Instead, we notice a negative trend which becomes stronger at smaller scales, and which is the same trend observed in the previous figure 10.4 for the correlation with the 40 GHz channel. If not taken into account, this spurious correlation would give rise to a high value of the test efficiency even in the case of a unbiased detection of r. This is, again, the reason why we do not include the smaller scales in the compatibility analysis, both for ST coefficients and cross-power spectra.

The last, important point to clarify concerns the higher test efficiency we get when correlating the recovered CMB map with low frequency data, in comparison with high frequency ones, both for ST coefficients and cross-power spectra. This is not so obvious, as we generally expect dust emission to have a similar or larger impact than synchrotron radiation on the r recovery. An immediate way to understand this is to look at the cross angular power spectrum between `CMBsol` and the raw input foregrounds maps from `PySM` at 40 and 402 GHz, not including CMB lensing and noise. Power spectra are shown in the right panel of figure 10.5. Correlation with dust maps is compatible with zero, oscillating between positive and negative values.

Instead, we see a clear positive correlation with synchrotron maps, at intermediate scales (from $\ell \sim 85$), which are the same scales where we observe deviations between `CMBsol` and Gaussian simulations in the ST coefficients ($j = 1$ in figure 9.15). This correlation suggests the presence of synchrotron-related structures at intermediate scales in the foreground residuals, which seems to be less mitigated by the foreground-cleaning algorithms adopted in this work.

It is good to notice here that in this work we focused on blind component separation methods based on the internal linear combination approach. In general, different component separation algorithms built upon other formalisms may lead to different properties of the foreground residuals left in the recovered CMB map. This might consequently result in different efficiencies of the robustness tests. It is worth stressing that the methodology presented in this manuscript does not intrinsically depend on the adopted foregrounds-cleaning algorithm; it is completely general and can be straightforwardly applied to other map-based component separation techniques (parametric, template-fitting, etc.), keeping in mind that each method requires a characterisation suite similar to the one described here.

Part IV

Conclusions

Chapter 11

Some final words

We finally reached the end of our trip. This Thesis attempted to condense all of the efforts I carried out during the last three years, while pursuing a PhD degree, into a single manuscript. It has been a wonderful time, during which I had the possibility to investigate some of the deep mysteries of our Universe, together with the CMB group at SISSA but also with the help of many people around the planet. This chapter only serves at summarising and concluding what I presented.

In this document, I reported the various contribution to Cosmology I was able to work out during the PhD. In particular, I focused on the observational probe known as the Cosmic Microwave Background (CMB) and on the immense wealth of information it contains about the very first moment of the Universe. The major current and future experiments mapping the microwave sky target the measurement a specific pattern hidden in the polarised CMB anisotropies, called the B-mode signal and parametrised by the tensor-to-scalar ratio r. The detection of such pattern is of paramount importance for the CMB scientific community, as it is thought to be induced by the presence of primordial gravitational waves generated by inflation, a theoretical mechanism responsible for the extremely accelerated expansion our Universe experienced immediately after the Big Bang. A measurement of these B-modes in CMB maps would thus grant us access to phases and energy scales of the Universe never seen before, shedding light on the physics governing the most profound mysteries of Nature.

In Part II, I described a first series of my contribution. At the beginning of my PhD I joined the scientific collaboration of the *Simons Observatory* (*SO*), a brand new microwave experiment located in Chile, which at the time of writing is taking its first data of the sky. Among others, a major goal of the *SO* is to constrain the tensor-to-scalar ratio r. This is an extremely difficult measurement: on one side, the underlying signal is very faint, where an exquisite control of the instrumental chain is required; on the other side, Galactic foregrounds (thermal dust emission,

synchrotron radiation) represent a formidable contamination which needs to be taken into account.

In order to mitigate the latter, observations of the sky at foreground-dominated frequencies are crucial for the characterisation and subtraction of dust and synchrotron emission. Unfortunately, the *SO* will not have foregrounds-dedicated channels for the first years of operations, which will instead focus on frequencies optimal for CMB observations (90, 150 GHz), where the contamination is minimum. One of my roles in the *SO* collaboration was to address this issue by introducing external data into the analysis pipeline: it is possible to take sky observations of other microwave experiments and combine them with *SO* data in order to fill the frequency gaps and gain important information on foregrounds behaviour. In particular, we chose to use data from three external experiments: *WMAP* (space-based, full-sky, low frequencies), *Planck* (space-based, full-sky, low and high frequencies), and *S-PASS* (ground-based, partial-sky, very low frequency). All these maps provide crucial insights on foregrounds emission at low and high frequencies, with different instrumental properties. Before being combined together, data must be brought to the same experimental configuration of the target *SO* one. In chapter 6 we described how these reprocessing steps were carried out, including the downgrading of the pixel resolution, the rotation from galactic to equatorial coordinates, and the restriction to the same observed sky regions. We verified that the reprocessing phase does not alter the cosmological information contained in the original maps, by looking at the preservation of the angular power spectrum features. We also provided some characterisation of the selected external data, to investigate the presence of any unexpected issue.

Additionally to the inclusion of external data, I also contributed to the generation of different types of simulations useful for the *SO* project, described in chapter 7. In particular, I built two sets of artificial skies: for the first one, I simulated realisations of a sky emission composed by CMB and Gaussian foregrounds signals, which can be used to estimate a covariance matrix for the power spectrum-based component separation method, one of the main *SO* pipelines for the inference of the tensor-to-scalar ratio; for the second set, I produced Gaussian realisations of polarised fields starting from a power-law shaped power spectrum, which can be used to estimate the transfer function induced by instrumental and atmospheric effects typical of the *SO* experimental configuration, dealt with through the application of numerical filters at the timestreams level. All of these contributions to the *SO* project are currently and will be used in the data analysis pipeline which aim is the measurement of primordial B-modes.

Looking more at the future, in Part III, I tackled the problem of validating an

eventual, future detection of the primordial B-modes from CMB observations. Experiments like *SO* or the future *LiteBIRD* could be able to provide a measurement of the tensor-to-scalar ratio with a certain degree of significance. But how can we be sure that the detected signal is of cosmological origin and not a false detection induced by the contaminating polarised emission of Galactic foregrounds? To address this question, we proposed a methodology to build robustness tests for general map-based component separation algorithms. All these foreground-cleaning methods leave a certain level of residual contamination in the recovered CMB field. This contamination is expected to be statistically different from a pure CMB signal, as it contains non-Gaussian features typical of Galactic foregrounds emission. The goal of the robustness test is to identify such features by adopting different statistical descriptors sensitive to anisotropic structures in a given field. In particular, we presented the performances of Minkowski functionals (MFs) and scattering transform (ST) statistics in identifying foregrounds contamination when we have a biased detection of the tensor-to-scalar ratio. We found that MFs are not able to reliably recognise the foregrounds contamination, when the considered maps are dominated by noise. ST statistics demonstrated instead to be very promising, being able to flag a contaminated detection in 90% or more of our realistic simulations. They also outperformed simpler statistical tools usually adopted for CMB science, as the goodness-of-fit analysis or the isotropic angular power spectrum. This suggest the possibility of building a robustness tests chain where several statistical analysis are progressively carried out on the available data, in order to validate or reject a potential detection of primordial B-modes. These results were summarised into the B-sure project, leading to two of the published papers associated with this Thesis, [178] and [179]. The mentioned works open the way for the development of new and diverse robustness test methodologies, and also for their application to different experimental configurations, which is left for future projects.

Now I will just use this last paragraph to conclude the story. We began this manuscript by interrogating ourselves on existential questions, wondering in particular about *where do we come from* as a whole, and what are the origins of our Universe. Having now survived my PhD adventure, I like to think that the efforts I collected in this Thesis can be a (small) step forward in answering this question, and a minimal contribution in the understanding of the Universe. Many mysteries of Nature awaits to be unveiled by human beings. I really hope to have left a little something in this collective journey.

List of acronyms

ΛCDM: Λ **C**old **D**ark **M**atter

ACT: **A**tacama **C**osmology **T**elescope

BAO: **B**aryonic **A**coustic **O**scillations

BBN: **B**ig **B**ang **N**ucleosynthesis

BICEP: **B**ackground **I**maging of **C**osmic **E**xtragalactic **P**olarization

CIB : **C**osmic **I**nfrared **B**ackground

CMB: **C**osmic **M**icrowave **B**ackground

CNN: **C**onvolutional **N**eural **N**etwork

CDM: **C**old **D**ark **M**atter

CL: **C**onfidence **L**evel

CWST: **c**ross **W**avelet **S**cattering **T**ransform

DES: **D**ark **E**nergy **S**urvey

DESI: **D**ark **E**nergy **S**pectroscopic **I**nstrument

FLRW: **F**riedmann-**L**emaître-**R**obertson-**W**alker (metric)

FWHM: **F**ull **W**idth at **H**alf **M**aximum

GR: **G**eneral **R**elativity

(P)GW: (**P**rimordial) **G**ravitational **W**aves

GUT: **G**rand **U**nified **T**heories

HEALPix: **H**ierarchical **E**qual **A**rea iso**L**atitude **Pix**elisation of a 2-sphere

LAT: **L**arge **A**perture **T**elescope

LiteBIRD: **Lite** (Light) satellite for the studies of ***B***-mode polarisation and **I**nflation from cosmic background **R**adiation **D**etection

LSS: **L**arge-**S**cale **S**tructure

MC-NILC: **M**ulti **C**lustering **N**eedlet **I**nternal **L**inear **C**ombination

MFs: **M**inkowski **F**unctional**s**

MBB: **M**odified **B**lack**b**ody

NCP: **N**orth **C**elestial **P**ole

NERSC: **N**ational **E**nergy **R**esearch **S**cientific **C**omputing **C**enter

`NPIPE`: **N**ERSC **Pipe**line

NILC: **N**eedlet **I**nternal **L**inear **C**ombination

PLA: ***P****lanck* **L**egacy **A**rchive

POLARBEAR: **POLAR**ization of the **B**ackground **R**adiation

PTE: **P**robability-**t**o-**E**xceed

PySM: **P**ython **S**ky **M**odel

RJ: **R**ayleigh-**J**eans

SAT: **S**mall **A**perture **T**elescope

SED: **S**pectral **E**nergy **D**istribution

SNIa: Type **Ia** **S**uper**n**ovae

SO: **S**imons **O**bservatory

S-PASS: **S**-band **P**olarisation **A**ll **S**ky **S**urvey

SPT: **S**outh **P**ole **T**elescope

SZ: **S**unyaev-**Z**el'dovich

WMAP: **W**ilkinson **M**icrowave **A**nisotropy **P**robe

WST or **ST**: **W**avelet **S**cattering **T**ransform

List of Figures

1.1 Expansion of the Universe. The comoving distance between points x_1, x_2 on a grid remains constant as the Universe expands. The physical distance is proportional to the comoving distance times the scale factor $a(t)$. Figure taken from [62]. 16

1.2 Constraints on the energy densities of matter ($\Omega_{\rm m}$) and dark energy (Ω_Λ), and the dark energy equation of state parameter (w). *Left:* constraint contours for $\Omega_{\rm m}$ and Ω_Λ, from a recent analysis of type Ia supernovae distances, allowing for an open Universe. Filled contours are the more recent results, while the black lines show constraints from the original 1998 sample of [183]. Dashed lines indicate the parameters combination to have a flat Universe and the separation between an accelerating or decelerating one. Data clearly prefer a flat, accelerating scenario. *Right:* contours for w and $\Omega_{\rm m}$, using type Ia supernovae, CMB, and BAO measurements, and assuming a flat Universe. The combination of different probes points to the cosmological constant ($w = -1$). Both figures are from [189]. 22

1.3 Dynamics of the Universe provided by the Friedmann equations. *Left:* energy density ρ as a function of the scale factor a. Different constituents of the universe dominates at different time of the history: radiation (red dashed), non-relativistic matter (blue dotted) and a cosmological constant (black dash-dotted). *Right:* behaviour of the scale factor a as a function of the cosmic time t, in the frame of the ΛCDM model. 24

1.4 Simplified chronology of the Universe, as per the ΛCDM concordance model. The timeline in the upper part shows an artistic view of the evolution of the cosmos on large scales, passing through inflation, CMB emission, "Dark Ages", birth of the first stars and galaxies, and until present time. The inserts below show some of the microscopic processes taking place during cosmic history. Credit: ESA. 28

1.5 Illustration of a propagating gravitational wave (tensor mode), perturbating the spacetime. The k-vector is along the z axis, coming out of the page. The upper panel shows the time evolution of the wave amplitude (neglecting the damping due to the expansion), while the lower panel shows the stretching and compression of spacetime perpendicular to the wave's direction of propagation at the different points in time during a single wave cycle. Figure from [62]. 34

2.1 A scalar field slowly rolling down a potential $V(\phi)$. Since it rolls slowly, it has little kinetic energy. The potential energy is non-zero, however, so the pressure is negative. The inflationary epoch ends once the field has reached the minimum of the potential. Figure taken from [62]. 41

2.2 Comoving scales evolution during and after the inflationary period. Scales are initially into the horizon and thus in casual connection. Inflation stretches the scales and make them exit the horizon, where they are not casually connected any more, and remain frozen. Finally, scales re-enter the horizon during the hot Big Bang era. Figure from [177]. 44

2.3 Marginalized joint 68% and 95% CL regions for $n_{\rm s}$ and r at $k_{\rm p} = 0.002\,{\rm Mpc}^{-1}$ from *Planck* data alone and in combination with *BICEP/Keck* and BAO data, compared to the theoretical predictions of some inflationary models. Figure from [167]. 50

3.1 *Left:* monopole and dipole maps of the CMB intensity. Upper panel illustrate the smooth isotropic CMB temperature $T_{\rm CMB} \simeq 2.725\,{\rm K}$, almost the same in every pixel. The lower panel shows a *Planck* intensity map at a frequency of 100 GHz [173], after removal of the monopole, in ${\rm K}_{\rm CMB}$ units. The dipole anisotropy is evident in the hotter and colder regions, typical of a Doppler shift. *Right:* the monopole spectrum of the CMB, from the original data by the *FIRAS* instrument of the *COBE* experiment [140, 219, 78], available at this link. The vertical axis is in surface intensity units, while the horizontal axis is the reciprocal of the microwave wavelength, proportional to the frequency. The black curve shows the expected intensity from a single temperature blackbody spectrum at $T = 2.725\,{\rm K}$. The error bars are there in the plot, smaller than the line width (!). 55

3.2 Sky emission fluctuations in each of the nine *Planck* frequency bands, after removal of the common dipole component. The fluctuations are expressed as equivalent temperature variations (in $K_{\rm CMB}$ units), so that CMB fluctuations will appear the same in each map. Figure from [173]. 57

3.3 Caricature of the Thomson scattering, comparing the output linear polarisation generated by monopole, dipole and quadrupole configurations. The monopole and dipole generate unpolarised light, while the quadrupole leaves a net linear polarisation on the outgoing photons. Figure from [207]. 58

3.4 Polarised microwave sky as observed by the seven frequency bands of the *Planck* satellite. The first two columns show the Q and U Stokes parameters; the last column shows the polarised intensity, $P = \sqrt{Q^2 + U^2}$. Figure from [173]. 60

3.5 *Upper panel:* polarisation pattern composed only of E-modes. *Lower panel:* polarisation pattern composed only of B-modes. As indicated on the right, the E-mode pattern is tangential around hot spots (red) and radial around cold spots (blue). B-mode polarisation show instead a characteristic swirling pattern (with different orientations around hot and cold spots). Figure taken from [111]. 62

3.6 Generation of E-mode polarisation from density perturbations. The fluid either falls into the potential well while accelerating, or climbs up while decelerating (yellow arrows), creating quadrupole anisotropy in the radiation. Black arrows show the fluid motion in the electron rest frame; red lines trace the electric field direction of the radiation scattered by the electrons, resulting in positive or negative E-mode polarisation patterns. Figure from [147]. . . 63

3.7 Compilation of recent CMB angular power spectrum measurements. *Upper:* temperature, E- and B-mode polarisation. *Middle:* TE cross-correlation. *Lower:* lensing deflection power spectrum. Different colours correspond to different experiments. The dashed line indicates the best-fit ΛCDM model to the *Planck* temperature, polarisation, and lensing data. Figure from [168]. . 72

3.8 Comparison between the CMB anisotropies observed by the three satellites *COBE* (left), *WMAP* (middle), and *Planck* (right), showing the amazing improvement in sensitivity and resolution. The three panels show 10-square-degree patches of all-sky microwave maps. Source: link. Courtesy NASA/JPL-Caltech. 74

4.1 Frequency dependence of the main components of the sub-millimetre sky in temperature (*left*) and polarisation (*right*, shown in P). The vertical grey bands indicate the Planck channels, while the coloured bands show the major signal and foreground components. For temperature, the widths of the bands show the range for sky masks with 81-93% sky coverage. For polarisation, the range is 73-93%. Figure from [168]. 79

4.2 *Left:* diagram showing the mechanism of synchrotron radiation emitted by an electron moving along a helical path around an ordered magnetic field. The Lorentz force is perpendicular to the magnetic field. Picture taken from [9]. *Right:* polarised synchrotron amplitude map by *Planck*, obtained with the `Commander` algorithm, at 40' resolution, and evaluated at the reference frequency of 30 GHz. From [169]. 81

4.3 *Left:* diagram showing how dust clouds are able to polarise background starlight. Asymmetric and elongated grains tend to align with the direction of the magnetic field, emitting polarised radiation along the big axis of the grain. Picture from the *Pasiphae* website. *Right:* polarised thermal dust amplitude map by *Planck*, obtained with `Commander` algorithm, at 5' resolution, and evaluated at the reference frequency of 353 GHz. From [169]. 82

4.4 Diagram illustrating the component separation principle, combining *Planck* multi-frequency sky maps (in temperature) to obtain a foreground-cleaned CMB reconstruction. From this website. 85

4.5 Foreground-cleaned CMB Q and U maps derived from the *Planck* `NPIPE` data set with `Commander` (top row) and `SEVEM` (bottom row), plotted with an angular resolution of 40'. From [173]. 88

5.1 Anticipated coverage (lighter region) of the SATs (*left*) and LAT (*right*) in equatorial coordinates, overlaid on a map of Galactic dust emission. The survey regions of other experiments are also indicated. Figure from [6]. . . 98

5.2 *Left*: summary of *SO* forecasts for the primordial power spectrum parameters $n_{\rm s}$ and r for two example cases: vanishingly small primordial tensor modes and primordial perturbations with $r = 0.01$. The contours correspond to 68% and 95% confidence levels. The *SO* baseline (goal) forecasts are shown as filled (dashed) contours. Constraint in this parameter space from *Planck*, *BICEP2/Keck*, and BAO are shown in grey. *Right:* same as above, with predictions for $n_{\rm s}$ and r from some inflationary models. 100

5.3 Small cutout of the preliminary *SO* CMB maps in polarisation, as observed by one of the SAT, co-added from the 90 and 150 GHz bands, and including data from ~ 14 months of observations. *Left* and *right* columns show the maps produced with two different, independent map-making algorithms, namely maximum likelihood and filter and bin, giving consistent results. Figure credit: *SO* Collaboration. 102

6.1 Frequency dependence, in Rayleigh-Jeans (RJ) brightness temperature, of the synchrotron and thermal dust emission at degree scales within the SATs footprint, compared to the CMB lensing B-mode signal. The turnover of the modified blackbody law for dust lies above this frequency range. The SATs channels are shown as vertical bands. Figure from [6]. 104

6.2 Examples of the external datasets adopted to integrate *SO* SATs observations, in Stokes Q maps. Units are $\mu{\rm K}_{\rm CMB}$. *From left to right: Planck* `NPIPE` 353 GHz channel, *WMAP* 23 GHz channel, and *S-PASS* 2.3 GHz. *From top to bottom:* input maps, then downgraded to $N_{\rm side} = 512$, and rotated from galactic to equatorial coordinates. *WMAP* data have not been downgraded as their are natively at the same resolution of the SATs. 106

6.3 Nominal sky regions observed by the *SO* SATs [218], in equatorial coordinates. Unobserved pixels are in black. *Left:* binary mask, covering $f_{\rm sky} = 10\%$ of the sky. *Right:* apodised mask weighted by the forecasted pixel observing time. Mask edges are apodised using a C1-type kernel with an 10 degrees apodisation scale. 110

6.4 Validation of the external data pre-processing, comparing angular power spectra before (solid lines) and after (circles) the downgrading and rotation steps carried out in harmonic space (*left panel*). The spectral features are preserved, with differences (*right panel*) being at the 5% level at very small scales $\ell > 1200$. Data shown are from the *Planck* 100 GHz channel, in EE (red) and BB (blue). 111

6.5 Binary version of the masks adopted for the characterisation of external data, in galactic coordinates. *Left:* mask covering the Galactic plane. *Right:* official mask used for the *Planck* analysis of polarised data, covering the Galactic plane, extra-galactic point sources, and other polarised structures. 112

6.6 EE and BB auto-frequency angular power spectra of *Planck* `NPIPE` data, evaluated adopting the mask covering the Galactic plane, point sources, and other polarised structures (right panel of figure 6.5). At each frequency, cross-splits (A$\times$B) spectra are computed to avoid noise correlation bias. . . 113

6.7 Same as figure 6.6, but here spectra are evaluated on the *SO* SATs sky patch of figure 6.3. 114

6.8 EE and BB auto-frequency angular power spectra of *Planck* `NPIPE` residuals simulation, evaluated adopting the mask covering the Galactic plane (left panel of figure 6.5). At each frequency, cross-splits (A$\times$B) spectra are computed to avoid noise correlation bias. Results are reported normalised by the error on the simulations. 115

6.9 Same as figure 6.8, but here spectra are evaluated on the *SO* SATs sky patch of figure 6.3. 116

6.10 EE (*left*) and BB (*right)* auto- (black and red) and cross-frequency (cyan) angular power spectra of *WMAP* data, evaluated adopting the mask covering the Galactic plane, point sources, and other polarised structures (right panel of figure 6.5). 116

6.11 EE (*left column*) and BB (*right column*) auto-angular power spectra of *S-PASS* data. Spectra in top panels are evaluated on the mask covering the Galactic plane, while the ones in the lower panels are computed on the *SO* SATs sky patches. For black circles, point sources are included in the analysis, while for cyan circles point sources are masked out. The improvement in removing the flattening due to polarised point sources is evident. 117

7.1 Example of sky maps simulated for the purpose of covariance matrix estimation, at a frequency of 353 GHz. *From left to right:* total intensity I, and Stokes Q and U maps. *From top to bottom:* CMB, synchrotron radiation, dust emission, and all the above co-added maps. At the selected frequency, thermal dust is the dominant component. Units are $\mu\mathrm{K}_{\mathrm{CMB}}$. 121
7.2 An example of Q and U maps containing E- and B-modes only. The characteristic $+$ and $\times$ patterns in the Q_E and Q_B maps indicate the dominance of E- and B-modes, respectively. 124
7.3 An example of transfer function for a 90 GHz SATs-like channel, showing the 16 $X'Y' \rightarrow XY$ polarisations, averaged over $N = 100$ realisations of the power-law simulations shown in figure 7.2. The (very thin) shaded area represents the standard deviation divided by $\sqrt{N}$. This transfer function is estimated using more aggressive filtering than what expected in the real *SO* data analysis. Figure from [101]. 126

8.1 B-mode signal reported by the *BICEP2* collaboration in 2014, with the black rods indicating the polarisation amplitude in different spots of the sky region observed by the telescope. Figure taken from the *BICEP2* 2014 Release Image Gallery, realised using data from [4]. 131

9.1 *Top panels:* Gaussian CMB field on the left, foregrounds map on the right, both in Stokes parameter Q. Visually, they clearly have a different morphology. *Bottom panel:* EE angular power spectrum of the above maps, over the full-sky or adopting a mask covering the Galactic plane. For a Gaussian field, the statistics is the same, when corrected for the sky fraction. 134
9.2 Flowchart illustrating the robustness test methodology and showing the various steps followed for the test characterisation and validation. 135
9.3 Input angular power spectra used to generate the Gaussian CMB simulations, in C_ℓ (*left*) and D_ℓ (*right*), adopting the best-fitting *Planck* 2018 parameters [166]. TT in black, EE in red, BB in blue. Solid lines represents spectra where B-modes arise from lensing only ($r_{\mathrm{in}} = 0$), while the dashed line show the contribution due to primordial gravitational waves for $r_{\mathrm{in}} = 5 \times 10^{-3}$. . . 138
9.4 Visualisation of the noise properties of the *LiteBIRD* configuration adopted in this work. *Left*: sensitivity levels used for generating the simulations, in $\mu\mathrm{K}_{\mathrm{CMB}} \cdot \mathrm{arcmin}$ units, with frequency bands denoted by vertical lines. They are also reported in table 9.1. *Right*: corresponding white noise power spectra, obtained from eq. (9.1). 139
9.5 Simulated polarised foregrounds emission maps, in Q (top row) and U (bottom row) Stokes parameters, for the `d10s5` model, at synchrotron-dominated (40 GHz, *left column*), foregrounds-minimum (100 GHz, *middle column*), and dust-dominated (402 GHz, *right column*) frequencies. All maps are smoothed to 70.5 arcmin, and reported in $\mu\mathrm{K}_{\mathrm{CMB}}$ units. 140
9.6 Sky partitions adopted in the MC-NILC component separation for `d1s1` (*left*) and `d10s5` (*right*) foreground models. 142

9.7 An example of the component separation outputs, for one simulation. *Left column:* reconstructed CMB B-mode signal (CMB solution), containing the CMB field, and noise and foreground residuals. *Middle column:* noise residuals. *Right column:* foreground residuals. Outputs are shown for NILC (*upper row*) and MC-NILC (*lower row*) methods, applied on the `d10s5` model with $r_{\rm in} = 0$. Units are $\mu{\rm K}_{\rm CMB}$. 142

9.8 *Left panel*: needlet bands Mexican configuration adopted in this work for component separation, where the first bands have been merged. *Right panel*: full-sky angular power spectra evaluated on foreground (red) and noise (blue) residuals, for NILC (dashed lines) and MC-NILC (solid lines), applied to the `d10s5` model. Black dotted line is the input CMB spectrum (lensing-only). . 144

9.9 Angular power spectra going into the likelihood evaluation for a single simulation, for the NILC - `d1s1` (*left*) and MC-NILC - `d10s5` (*right*) scenarios. *Main plot*: spectra, with $C_\ell^{\rm th}(r=0)$ in dashed blue, $C_\ell^{\rm obs}$ in solid red, $C_\ell^{\rm lensing}$ in dotted black, and the best-fit in solid black. The shaded area represents the standard deviation of $C_\ell^{\rm obs}$ across 300 realisations. *Inset*: normalised likelihood on the tensor-to-scalar ratio; the peak value is indicated by the dashed line. 145

9.10 Same as figure 9.9, but considering only the largest scales in the likelihood evaluation. The maximum multipole included ($\ell = 48$) is indicated by the vertical dash-dotted line. 146

9.11 Results of the goodness-of-fit analysis for the NILC - `d1s1` (red dashed lines) and MC-NILC - `d10s5` (blue solid lines) cases. *Left panel:* one realisation of the $-\ln\mathcal{L}$ distribution of the 300 tensor simulations, compared against the value given by the best-fit on the CMB solution (vertical lines); the corresponding PTE is also indicated. *Right panel:* PTE distribution for the 300 simulations, with the percentage of incompatible simulations shown on the right. 148

9.12 An example of the maps considered for the robustness tests, for the MC-NILC - `d10s5` case. *Upper row, from left to right:* CMB solution (`CMBsol`), Gaussian simulations `FGgauss` and `CMBtens`. *Lower row, from left to right:* as upper row, but without CMB lensing and noise, with non-Gaussian and Gaussian foreground residuals, and tensor perturbations. 149

9.13 Same maps as figure 9.12, but filtered with the first needlet band b_0 to retain only the largest scales, at multipoles $\ell \lesssim 60$. 150

9.14 Minkowski functionals, for a single simulation, in the MC-NILC - `d10s5` case. MFs of the `CMBsol` map are shown with blue points, while the average MFs of 300 `CMBtens` (`FGgauss`) simulations are indicated by the black dashed line (red points). The shaded areas represent the standard deviation across the simulations. *Upper row*: all angular scales included. *Lower row*: largest scales only, $\ell \lesssim 60$. 156

9.15 Scattering coefficients for the d10s5 model with $r_{\text{in}} = 0$, one simulation. Same-coloured lines represent different orientations. Coefficients of the component-separated CMB map are in blue, while black (orange) stands for the average coefficients of the CMBtens (FGgauss) Gaussian simulations. The secondary x-axis shows the N_{side} of the maps for scale j. *Upper figure:* auto-coefficients, using NILC. *Lower figure:* cross-coefficients, correlating with 40 GHz data, using MC-NILC. *Top panels:* coefficients, with the shaded areas indicating the standard deviation of the Gaussian simulations σ. *Bottom panels:* normalised difference between coefficients of CMBsol and of Gaussian maps, with the dotted lines indicating the $\pm 3\sigma$ region. 159

9.16 S_4 auto-coefficient for the MC-NILC - d10s5 - $r_{\text{in}} = 0$ scenario. Colour code is the same as figure 9.15. Lower plot shows the difference between CMBsol and Gaussian simulations, in standard deviation units, with the dotted lines indicating the $\pm 3\sigma$ region. 161

9.17 Check of the calibration of the compatibility analysis for the ST-based test, on held-out reference simulations not used in the fit. The considered foreground model is d0s0, with $r_{\text{in}} = 0$, using the S_1 and S_2 coefficients combined together. *Left:* rescaled Mahalanobis distances against the fitted F distribution. *Right:* measured rejection rate versus theoretical false-positive rate (blue), with the perfect-calibration diagonal (dashed) and its 95% interval (grey). . 165

10.1 MFs weighted difference for the MC-NILC - d10s5 scenario. Each line is one of the 300 input realisations. Blue (orange) lines indicate CMBtens (FGgauss) as the reference Gaussian simulations. Plots are shown for all multipoles (*upper row*) and for large scales only (*lower row*) analysis. Dotted lines delimit the $\pm 3\sigma$ region. Efficiencies are reported in the lower corners of the middle panels. 168

10.2 Average ratio between spectra of foreground and noise residuals. Dashed red lines indicate the d1s1 model, while blue solid lines are for d10s5. Shaded areas represent the standard deviation across 300 simulations. MFs test efficiencies are also reported for comparison. The vertical line indicates the smallest scale included in the first needlet band b_0. 172

10.3 ST S_1 coefficient for two validation cases, with the same colour code of figure 9.15. *Left:* cross-coefficients correlating with synchrotron-dominated input data, for the MC-NILC - d0s0 - $r_{\text{in}} = 0$ case. *Right:* auto-coefficients, for the NILC - d0s0 - $r_{\text{in}} = 5 \times 10^{-3}$ scenario. In both plots, the tensor-to-scalar ratio estimate is not biased, coefficients are compatible, and the robustness test raises no strong warnings. 176

10.4 Cross angular power spectrum between noise residuals and noise maps, for different input frequency channels (in different colours), one simulation. The dashed line indicates the noise residuals auto spectrum. The considered scenario here is MC-NILC applied on the d10s5 model. We notice the strong negative correlation for the 40 GHz channel. 177

10.5 The considered scenario is MC-NILC - `d10s5` - $r_{\rm in} = 0$. *Left:* probability-to-exceed (PTE, or p-value) given by the compatibility analysis for the cross-synchrotron test using the S_1 and S_2 coefficients, as a function of the bias on the tensor-to-scalar ratio r. Results are shown for 50 simulations and for different values of foregrounds rescaling $A_{\rm fgs}$. The dotted vertical line is a reference value of $\delta(r) = 10^{-3}$, while the horizontal dashed line is the PTE threshold for incompatibility: ST coefficients are compatible above 0.05, and incompatible below, when foreground contamination is detected. *Right:* cross-angular power spectrum between `CMBsol` and foreground maps; cross-correlation with dust is indicated in yellow, while blue is for synchrotron. See text for discussion. 178

10.6 Weighted difference between the cross-40 GHz angular power spectrum of `CMBsol` maps and of Gaussian simulations, in units of the standard deviation of the simulations σ. The foregrounds model under study is `d0s0`, and the input tensor-to-scalar ratio $r_{\rm in} = 0$ is correctly recovered without bias. The observed negative trend is due to noise correlation between residuals and input data. See text for details. 181

List of Tables

1.1 Best-fit values of the concordance ΛCDM model parameters combining CMB data from the *Planck* experiment with BAO measurements. We report the flat ($\Omega_\kappa = 0$) `base_plikHM_TTTEEE_lowl_lowE_lensing_post_BAO` set from [166]. The six parameters in the upper part are the primary parameters of the concordance cosmology, while those below are derived parameters. The last column lists the 95% confidence-level parameter limits. More details are on the *Planck* wiki. 27

5.1 Properties of the planned *SO* surveys, from the 2019 forecast paper [6]. Some of these values will obviously be different in the current operational phase. Noise columns give anticipated white noise levels for temperature, with polarisation noise $\sqrt{2}$ higher as both Q and U Stokes parameters are measured. Noise levels are quoted for a homogeneous hits map, and are reported in $\mu\mathrm{K_{CMB}}\cdot$ arcmin units. 99

6.1 Properties of the different datasets considered in this Part, including the *SO* SATs and external data (*Planck*, *WMAP*, *S-PASS*) configurations. 108

9.1 *LiteBIRD* specifics used to generate the simulations, with values related to 15 observational bands (15 frequencies, 22 effective channels). From left to right the columns are: telescope covering the band (Low-Frequency Telescope LFT, Medium-Frequency Telescope MFT, High-Frequency Telescope HFT); band identification number (ID); band centre frequency in GHz; main beam size (FWHM) in arcmin; sensitivity in $\mu\mathrm{K_{CMB}}\cdot\mathrm{arcmin}$. 137

9.2 Percentage of incompatible simulations from the goodness-of-fit analysis, considering all scales at the same time or only the largest scales. Values are reported for each scenario considered. 147

10.1 Efficiency η (in %) of the MFs-based robustness test on polarisation B-mode maps, for the different scenarios considered in the study. Values are reported for a pixel-space analysis (i.e., including all angular scales), both for no delensing ($A_{\rm lens} = 1.0$) or 50% delensing ($A_{\rm lens} = 0.5$) applied. 169

10.2 Same as table 10.1 but in needlet space, retaining only the largest scales in the analysis ($\ell \lesssim 60$). 170

10.3 Efficiency η (in %) of the ST-based robustness test on polarisation B-mode maps, for the different scenarios considered in the study. Results are shown for the S_1 and S_2 coefficients combined together, between scales $j_{\text{min}} = 1$ and $j_{\text{max}} = 4$. 174
10.4 Same as table 10.3, but here results are shown for the S_4 coefficient, after dimensionality reduction. 179
10.5 Efficiency η (in %) of the C_ℓ-based test on polarisation B-mode maps, for the MC-NILC component separation. Results are shown for angular scales between $\ell_{\text{min}} = 11$ and $\ell_{\text{max}} = 95$ (scales range similar to the one sampled by $1 < j < 4$ in table 10.3). 180

Bibliography

[1] M. Abitbol et al. “The Simons Observatory: science goals and forecasts for the enhanced Large Aperture Telescope”. In: *Journal of Cosmology and Astroparticle Physics* 2025.8, 034 (Aug. 2025), p. 034. DOI: 10.1088/1475-7516/2025/08/034. arXiv: 2503.00636 [astro-ph.IM].

[2] I. Abril-Cabezas et al. “Impact of Galactic dust non-Gaussianity on searches for B-modes from inflation”. In: *Monthly Notices of the Royal Astronomical Society* 527.3 (Jan. 2024), pp. 5751–5766. DOI: 10.1093/mnras/stad3529. arXiv: 2309.09978 [astro-ph.CO].

[3] I. Abril-Cabezas et al. “The Simons Observatory: forecasted constraints on primordial gravitational waves with the expanded array of Small Aperture Telescopes”. In: *Journal of Cosmology and Astroparticle Physics* 2026.4, 051 (Apr. 2026), p. 051. DOI: 10.1088/1475-7516/2026/04/051. arXiv: 2512.15833 [astro-ph.CO].

[4] P. A. R. Ade et al. “Detection of B-Mode Polarization at Degree Angular Scales by BICEP2”. In: *Phys. Rev. Lett.* 112 (24 June 2014), p. 241101. DOI: 10.1103/PhysRevLett.112.241101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.112.241101.

[5] P. A. R. Ade et al. “Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum with the POLARBEAR Experiment”. In: *Phys. Rev. Lett.* 113 (2 July 2014), p. 021301. DOI: 10.1103/PhysRevLett.113.021301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.021301.

[6] P. A. R. Ade et al. “The Simons Observatory: science goals and forecasts”. In: *Journal of Cosmology and Astroparticle Physics* 2019.02 (Feb. 2019), p. 056. DOI: 10.1088/1475-7516/2019/02/056. URL: https://dx.doi.org/10.1088/1475-7516/2019/02/056.

[7] P. A. R. Ade et al. “Improved Constraints on Primordial Gravitational Waves using Planck, WMAP, and BICEP/Keck Observations through the 2018

Observing Season". In: *Physical Review Letters* 127.15, 151301 (Oct. 2021), p. 151301. DOI: 10.1103/PhysRevLett.127.151301. arXiv: 2110.00483 [astro-ph.CO].

[8] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. Springer Monographs in Mathematics. Springer New York, 2009. ISBN: 9780387481166. URL: https://books.google.it/books?id=R5BGvQ3ejloC.

[9] E. Alexander. "Magnetic fields around Radio Galaxies". PhD thesis. The University of Manchester, 2022.

[10] E. Allys et al. "The RWST, a comprehensive statistical description of the non-Gaussian structures in the ISM". In: *Astronomy and Astrophysics* 629 (2019), A115. DOI: 10.1051/0004-6361/201834975. URL: https://doi.org/10.1051/0004-6361/201834975.

[11] E. Allys et al. "New interpretable statistics for large-scale structure analysis and generation". In: *Phys. Rev. D* 102 (10 Nov. 2020), p. 103506. DOI: 10.1103/PhysRevD.102.103506. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.103506.

[12] D. Alonso et al. "A unified pseudo-C_ℓ framework". In: *Monthly Notices of the Royal Astronomical Society* 484.3 (Apr. 2019), pp. 4127–4151. DOI: 10.1093/mnras/stz093. arXiv: 1809.09603 [astro-ph.CO].

[13] R. A. Alpher, H. Bethe, and G. Gamow. "The Origin of Chemical Elements". In: *Phys. Rev.* 73 (7 Apr. 1948), pp. 803–804. DOI: 10.1103/PhysRev.73.803. URL: https://link.aps.org/doi/10.1103/PhysRev.73.803.

[14] R. A. Alpher and R. Herman. "Evolution of the Universe". In: *Nature* 162.4124 (1948), pp. 774–775. DOI: 10.1038/162774b0.

[15] S. Azzoni et al. "A minimal power-spectrum-based moment expansion for CMB B-mode searches". In: *Journal of Cosmology and Astroparticle Physics* 2021.05 (May 2021), p. 047. DOI: 10.1088/1475-7516/2021/05/047. URL: https://doi.org/10.1088/1475-7516/2021/05/047.

[16] S. Azzoni et al. "A hybrid map-C_ℓ component separation method for primordial CMB B-mode searches". In: *Journal of Cosmology and Astroparticle Physics* 2023.3, 035 (Mar. 2023), p. 035. DOI: 10.1088/1475-7516/2023/03/035. arXiv: 2210.14838 [astro-ph.CO].

[17] C. Baccigalupi. "Cosmic microwave background polarisation: foreground contrast and component separation". In: *New Astronomy Reviews* 47.11-12 (Dec. 2003), pp. 1127–1134. DOI: 10.1016/j.newar.2003.09.038. arXiv: astro-ph/0306181 [astro-ph].

[18] N. Bartolo et al. “Non-Gaussianity from inflation: theory and observations”. In: *Physics Reports* 402.3 (2004), pp. 103–266. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2004.08.022. URL: https://www.sciencedirect.com/science/article/pii/S0370157304003151.

[19] S. Basak and J. Delabrouille. “A needlet internal linear combination analysis of WMAP 7-year data: estimation of CMB temperature map and power spectrum”. In: *Monthly Notices of the Royal Astronomical Society* 419.2 (Dec. 2011), pp. 1163–1175. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.19770.x. URL: https://doi.org/10.1111/j.1365-2966.2011.19770.x.

[20] R. A. Battye et al. “Statistical properties of polarized radio sources at high frequency and their impact on cosmic microwave background polarization measurements”. In: *Monthly Notices of the Royal Astronomical Society* 413.1 (May 2011), pp. 132–148. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.18115.x. eprint: https://academic.oup.com/mnras/article-pdf/413/1/132/18581492/mnras0413-0132.pdf. URL: https://doi.org/10.1111/j.1365-2966.2010.18115.x.

[21] E. J. Baxter et al. “A Measurement of Gravitational Lensing of the Cosmic Microwave Background by Galaxy Clusters Using Data from the South Pole Telescope”. In: *The Astrophysical Journal* 806.2, 247 (June 2015), p. 247. DOI: 10.1088/0004-637X/806/2/247. arXiv: 1412.7521 [astro-ph.CO].

[22] C. L. Bennett et al. “First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission”. In: *The Astrophysical Journal Supplement Series* 148.1 (Sept. 2003), p. 97. DOI: 10.1086/377252. URL: https://dx.doi.org/10.1086/377252.

[23] C. L. Bennett et al. “Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results”. In: *ApJ Supplements* 208.2, 20 (Oct. 2013), p. 20. DOI: 10.1088/0067-0049/208/2/20. arXiv: 1212.5225 [astro-ph.CO].

[24] BICEP/Keck Collaboration et al. “Measurements of Degree-Scale B-mode Polarization with the BICEP/Keck Experiments at South Pole”. In: *arXiv e-prints*, arXiv:1807.02199 (July 2018), arXiv:1807.02199. DOI: 10.48550/arXiv.1807.02199. arXiv: 1807.02199 [astro-ph.CO].

[25] BICEP2 Collaboration et al. “Bicep2 III: Instrumental Systematics”. In: *The Astrophysical Journal* 814.2, 110 (Dec. 2015), p. 110. DOI: 10.1088/0004-637X/814/2/110. arXiv: 1502.00608 [astro-ph.IM].

[26] BICEP2 Collaboration et al. “Improved Constraints on Cosmology and Foregrounds from BICEP2 and Keck Array Cosmic Microwave Background Data with Inclusion of 95 GHz Band”. In: *Physical Review Letters* 116.3, 031302 (Jan. 2016), p. 031302. DOI: 10.1103/PhysRevLett.116.031302. arXiv: 1510.09217 [astro-ph.CO].

[27] BICEP2/Keck Collaboration et al. “Joint Analysis of BICEP2/Keck Array and Planck Data”. In: *Physical Review Letters* 114.10, 101301 (Mar. 2015), p. 101301. DOI: 10.1103/PhysRevLett.114.101301. arXiv: 1502.00612 [astro-ph.CO].

[28] M. Birkinshaw. “The Sunyaev-Zel’dovich effect”. In: *Phys. Rep.* 310.2-3 (Mar. 1999), pp. 97–195. DOI: 10.1016/S0370-1573(98)00080-5. arXiv: astro-ph/9808050 [astro-ph].

[29] L. E. Bleem et al. “Galaxy Clusters Discovered via the Sunyaev-Zel’dovich Effect in the 2500-Square-Degree SPT-SZ Survey”. In: *ApJ Supplements* 216.2, 27 (Feb. 2015), p. 27. DOI: 10.1088/0067-0049/216/2/27. arXiv: 1409.0850 [astro-ph.CO].

[30] J. R. Bond and G. Efstathiou. “Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter”. In: *ApJ Letters* 285 (Oct. 1984), pp. L45–L48. DOI: 10.1086/184362.

[31] L. Boubekeur and D. H. Lyth. “Hilltop inflation”. In: *Journal of Cosmology and Astroparticle Physics* 2005.07 (July 2005), p. 010. DOI: 10.1088/1475-7516/2005/07/010. URL: https://doi.org/10.1088/1475-7516/2005/07/010.

[32] R. Brout, F. Englert, and E. Gunzig. “The creation of the universe as a quantum phenomenon”. In: *Annals of Physics* 115.1 (1978), pp. 78–106. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(78)90176-8. URL: https://www.sciencedirect.com/science/article/pii/0003491678901768.

[33] J. Bruna and S. Mallat. “Invariant Scattering Convolution Networks”. In: *arXiv e-prints*, arXiv:1203.1513 (Mar. 2012), arXiv:1203.1513. DOI: 10.48550/arXiv.1203.1513. arXiv: 1203.1513 [cs.CV].

[34] J. Bruna and S. Mallat. “Multiscale Sparse Microcanonical Models”. In: *arXiv e-prints*, arXiv:1801.02013 (Jan. 2018), arXiv:1801.02013. DOI: 10.48550/arXiv.1801.02013. arXiv: 1801.02013 [math-ph].

[35] E. F. Bunn. “Calculation of Cosmic Background Radiation Anisotropies and Implications”. In: *The Cosmic Microwave Background*. Ed. by C. H. Lineweaver et al. Vol. 502. NATO Advanced Study Institute (ASI) Series

C. Jan. 1997, p. 135. DOI: 10.48550/arXiv.astro-ph/9607088. arXiv: astro-ph/9607088 [astro-ph].

[36] E. F. Bunn et al. "E/B decomposition of finite pixelized CMB maps". In: *Phys. Rev. D* 67 (2 Jan. 2003), p. 023501. DOI: 10.1103/PhysRevD.67.023501. URL: https://link.aps.org/doi/10.1103/PhysRevD.67.023501.

[37] P. Campeti et al. "LiteBIRD science goals and forecasts. A case study of the origin of primordial gravitational waves using large-scale CMB polarization". In: *Journal of Cosmology and Astroparticle Physics* 2024.06 (June 2024), p. 008. DOI: 10.1088/1475-7516/2024/06/008. URL: https://doi.org/10.1088/1475-7516/2024/06/008.

[38] P. Campeti et al. "From few to many maps: A fast map-level emulator for extreme augmentation of cosmic microwave background systematics datasets". In: *Astronomy and Astrophysics* 700, A136 (Aug. 2025), A136. DOI: 10.1051/0004-6361/202554540. arXiv: 2503.11643 [astro-ph.CO].

[39] M. P. do Carmo. *Differential geometry of curves and surfaces*. Prentice Hall, 1976.

[40] A. Carones et al. "Analysis of Needlet Internal Linear Combination performance on B-mode data from sub-orbital experiments". In: *A&A* 677 (2023), A147. DOI: 10.1051/0004-6361/202244824. URL: https://doi.org/10.1051/0004-6361/202244824.

[41] A. Carones et al. "Multiclustering needlet ILC for CMB B-mode component separation". In: *Monthly Notices of the Royal Astronomical Society* 525.2 (Oct. 2023), pp. 3117–3135. DOI: 10.1093/mnras/stad2423. arXiv: 2212.04456 [astro-ph.CO].

[42] A. Carones et al. "Minkowski functionals of CMB polarization intensity with PYNKOWSKI: theory and application to Planck and future data". In: *Monthly Notices of the Royal Astronomical Society* 527.1 (Jan. 2024), pp. 756–773. DOI: 10.1093/mnras/stad3002. arXiv: 2211.07562 [astro-ph.CO].

[43] A. Carones et al. "BROOM: a python package for model-independent analysis of microwave astronomical data". In: *arXiv e-prints*, arXiv:2604.14088 (Apr. 2026), arXiv:2604.14088. DOI: 10.48550/arXiv.2604.14088. arXiv: 2604.14088 [astro-ph.CO].

[44] E. Carretti et al. "S-band Polarization All-Sky Survey (S-PASS): survey description and maps". In: *Monthly Notices of the Royal Astronomical Society* 489.2 (Oct. 2019), pp. 2330–2354. DOI: 10.1093/mnras/stz806. arXiv: 1903.09420 [astro-ph.GA].

[45] J. Carrón Duque et al. “Point source detection and false discovery rate control on CMB maps”. In: *Astronomy and Computing* 28, 100310 (July 2019), p. 100310. DOI: 10.1016/j.ascom.2019.100310. arXiv: 1902.06636 [astro-ph.CO].

[46] J. Carrón Duque et al. “Minkowski Functionals in SO(3) for the spin-2 CMB polarisation field”. In: *Journal of Cosmology and Astroparticle Physics* 2024.1, 039 (Jan. 2024), p. 039. DOI: 10.1088/1475-7516/2024/01/039. arXiv: 2301.13191 [astro-ph.CO].

[47] R. Cepeda-Arroita et al. “Spectral properties of Anomalous Microwave Emission in 144 Galactic clouds”. In: *Astronomy and Astrophysics* 708, A14 (Mar. 2026), A14. DOI: 10.1051/0004-6361/202557578. arXiv: 2510.05067 [astro-ph.GA].

[48] S. Cheng et al. “A new approach to observational cosmology using the scattering transform”. In: *Monthly Notices of the Royal Astronomical Society* 499.4 (Nov. 2020), pp. 5902–5914. ISSN: 0035-8711. DOI: 10.1093/mnras/staa3165. URL: https://doi.org/10.1093/mnras/staa3165.

[49] S. Cheng et al. “Scattering spectra models for physics”. In: *PNAS Nexus* 3.4 (Apr. 2024). ISSN: 2752-6542. DOI: 10.1093/pnasnexus/pgae103. URL: https://doi.org/10.1093/pnasnexus/pgae103.

[50] J. Chluba et al. “Spectral Distortions of the CMB as a Probe of Inflation, Recombination, Structure Formation and Particle Physics”. In: *Bulletin of the AAS* 51.3, 184 (May 2019), p. 184. DOI: 10.48550/arXiv.1903.04218. arXiv: 1903.04218 [astro-ph.CO].

[51] J. Chluba et al. “New horizons in cosmology with spectral distortions of the cosmic microwave background”. In: *Experimental Astronomy* 51.3 (June 2021), pp. 1515–1554. DOI: 10.1007/s10686-021-09729-5. arXiv: 1909.01593 [astro-ph.CO].

[52] P. Coles and J. D. Barrow. “Non-Gaussian statistics and the microwave background radiation”. In: *Monthly Notices of the Royal Astronomical Society* 228 (Sept. 1987), pp. 407–426. DOI: 10.1093/mnras/228.2.407.

[53] R. J. Cooke, M. Pettini, and C. C. Steidel. “One Percent Determination of the Primordial Deuterium Abundance”. In: *The Astrophysical Journal* 855.2 (Mar. 2018), p. 102. DOI: 10.3847/1538-4357/aaab53. URL: https://doi.org/10.3847/1538-4357/aaab53.

[54] B. P. Crill et al. “BOOMERANG: A Balloon-borne Millimeter-Wave Telescope and Total Power Receiver for Mapping Anisotropy in the Cosmic Microwave Background”. In: *ApJ Supplements* 148.2 (Oct. 2003), pp. 527–541. DOI: 10.1086/376894. arXiv: astro-ph/0206254 [astro-ph].

[55] S. Das et al. “Detection of the Power Spectrum of Cosmic Microwave Background Lensing by the Atacama Cosmology Telescope”. In: *Physical Review Letters* 107.2, 021301 (July 2011), p. 021301. DOI: 10.1103/PhysRevLett.107.021301. arXiv: 1103.2124 [astro-ph.CO].

[56] P. de Bernardis et al. “A flat Universe from high-resolution maps of the cosmic microwave background radiation”. In: *Nature* 404.6781 (Apr. 2000), pp. 955–959. DOI: 10.1038/35010035. arXiv: astro-ph/0004404 [astro-ph].

[57] J. Delabrouille et al. “A full sky, low foreground, high resolution CMB map from WMAP”. In: *Astronomy and Astrophysics* 493.3 (Jan. 2009), pp. 835–857. DOI: 10.1051/0004-6361:200810514. arXiv: 0807.0773 [astro-ph].

[58] J. M. Delouis et al. “Non-Gaussian modelling and statistical denoising of Planck dust polarisation full-sky maps using scattering transforms”. In: *Astronomy and Astrophysics* 668 (2022), A122. DOI: 10.1051/0004-6361/202244566. URL: https://doi.org/10.1051/0004-6361/202244566.

[59] J. Dick, M. Remazeilles, and J. Delabrouille. “Impact of calibration errors on CMB component separation using FastICA and ILC”. In: *Monthly Notices of the Royal Astronomical Society* 401.3 (Jan. 2010), pp. 1602–1612. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.15798.x. eprint: https://academic.oup.com/mnras/article-pdf/401/3/1602/3811973/mnras0401-1602.pdf. URL: https://doi.org/10.1111/j.1365-2966.2009.15798.x.

[60] R. H. Dicke et al. “Cosmic Black-Body Radiation.” In: *The Astrophysical Journal* 142 (July 1965), pp. 414–419. DOI: 10.1086/148306.

[61] C. Dickinson, R. D. Davies, and R. J. Davis. “Towards a free–free template for CMB foregrounds”. In: *Monthly Notices of the Royal Astronomical Society* 341.2 (May 2003), pp. 369–384. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2003.06439.x. eprint: https://academic.oup.com/mnras/article-pdf/341/2/369/3851468/341-2-369.pdf. URL: https://doi.org/10.1046/j.1365-8711.2003.06439.x.

[62] S. Dodelson and F. Schmidt. *Modern Cosmology*. 2020. DOI: 10.1016/C2017-0-01943-2.

[63] A. G. Doroshkevich and I. D. Novikov. “Mean Density of Radiation in the Metagalaxy and Certain Problems in Relativistic Cosmology”. In: *Soviet Physics Doklady* 9 (Aug. 1964), p. 111.

[64] B. T. Draine. *Physics of the Interstellar and Intergalactic Medium*. Princeton University Press, 2011.

[65] B. T. Draine and A. Lazarian. “Electric Dipole Radiation from Spinning Dust Grains”. In: *The Astrophysical Journal* 508.1 (Nov. 1998), p. 157. DOI: 10.1086/306387. URL: https://doi.org/10.1086/306387.

[66] A. Ducout et al. “Non-Gaussianity and Minkowski functionals: forecasts for Planck”. In: *Monthly Notices of the Royal Astronomical Society* 429.3 (Mar. 2013), pp. 2104–2126. DOI: 10.1093/mnras/sts483. arXiv: 1209.1223 [astro-ph.CO].

[67] J. R. Eimer et al. “CLASS Angular Power Spectra and Map-component Analysis for 40 GHz Observations through 2022”. In: *The Astrophysical Journal* 963.2 (Mar. 2024), p. 92. DOI: 10.3847/1538-4357/ad1abf. URL: https://doi.org/10.3847/1538-4357/ad1abf.

[68] A. Einstein. “Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie”. In: *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften* (Jan. 1917), pp. 142–152.

[69] H. K. Eriksen et al. “Power Spectrum Estimation from High-Resolution Maps by Gibbs Sampling”. In: *The Astrophysical Journal Supplement Series* 155.2 (Dec. 2004), p. 227. DOI: 10.1086/425219. URL: https://doi.org/10.1086/425219.

[70] H. K. Eriksen et al. “Joint Bayesian Component Separation and CMB Power Spectrum Estimation”. In: *The Astrophysical Journal* 676.1 (Mar. 2008), p. 10. DOI: 10.1086/525277. URL: https://dx.doi.org/10.1086/525277.

[71] J. Errard and R. Stompor. “Characterizing bias on large scale CMB B -modes after Galactic foregrounds cleaning”. In: *Phys. Rev. D.* 99.4, 043529 (Feb. 2019), p. 043529. DOI: 10.1103/PhysRevD.99.043529. arXiv: 1811.00479 [astro-ph.CO].

[72] J. Errard et al. “Modeling Atmospheric Emission for CMB Ground-based Observations”. In: *The Astrophysical Journal* 809.1, 63 (Aug. 2015), p. 63. DOI: 10.1088/0004-637X/809/1/63. arXiv: 1501.07911 [astro-ph.IM].

[73] T. Essinger-Hileman et al. “CLASS: the cosmology large angular scale surveyor”. In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*. Ed. by Wayne S. Holland and Jonas Zmuidzinas. Vol. 9153. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. July 2014, 91531I, p. 91531I. DOI: 10.1117/12.2056701. arXiv: 1408.4788 [astro-ph.IM].

[74] R. Fernández-Cobos et al. “Multiresolution internal template cleaning: an application to the Wilkinson Microwave Anisotropy Probe 7-yr polarization data”. In: *Monthly Notices of the Royal Astronomical Society* 420.3 (Mar. 2012), pp. 2162–2169. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.20182.x. eprint: https://academic.oup.com/mnras/article-pdf/420/3/2162/3015247/mnras0420-2162.pdf. URL: https://doi.org/10.1111/j.1365-2966.2011.20182.x.

[75] V. Ferrari, L. Gualtieri, and P. Pani. *General Relativity and its Applications*. CRC Press, Dec. 2020. ISBN: 978-0-429-49140-5. DOI: 10.1201/9780429491405.

[76] D. P. Finkbeiner, M. Davis, and D. J. Schlegel. “Extrapolation of Galactic Dust Emission at 100 Microns to Cosmic Microwave Background Radiation Frequencies Using FIRAS”. In: *The Astrophysical Journal* 524.2 (Oct. 1999), p. 867. DOI: 10.1086/307852. URL: https://doi.org/10.1086/307852.

[77] D. J. Fixsen. “THE TEMPERATURE OF THE COSMIC MICROWAVE BACKGROUND”. In: *The Astrophysical Journal* 707.2 (Nov. 2009), p. 916. DOI: 10.1088/0004-637X/707/2/916. URL: https://doi.org/10.1088/0004-637X/707/2/916.

[78] D. J. Fixsen et al. “The Cosmic Microwave Background Spectrum from the Full COBE* FIRAS Data Set”. In: *The Astrophysical Journal* 473.2 (Dec. 1996), p. 576. DOI: 10.1086/178173. URL: https://doi.org/10.1086/178173.

[79] A. Friedmann. “Über die Krümmung des Raumes”. In: *Zeitschrift fur Physik* 10 (Dec. 1922), pp. 377–386. DOI: 10.1007/BF01332580.

[80] U. Fuskeland et al. “Spatial Variations in the Spectral Index of Polarized Synchrotron Emission in the 9 yr WMAP Sky Maps”. In: *The Astrophysical Journal* 790.2, 104 (Aug. 2014), p. 104. DOI: 10.1088/0004-637X/790/2/104. arXiv: 1404.5323 [astro-ph.CO].

[81] G. Galloni, N. Bartolo, S. Matarrese, M. Migliaccio, A. Ricciardone, and N. Vittorio. “Updated constraints on amplitude and tilt of the tensor primordial spectrum”. In: *Journal of Cosmology and Astroparticle Physics* 2023.04 (Apr. 2023), p. 062. DOI: 10.1088/1475-7516/2023/04/062. URL: https://dx.doi.org/10.1088/1475-7516/2023/04/062.

[82] N. Galitzki. “The Simons Observatory: Project Overview”. In: *arXiv e-prints*, arXiv:1810.02465 (Oct. 2018), arXiv:1810.02465. DOI: 10.48550/arXiv.1810.02465. arXiv: 1810.02465 [astro-ph.IM].

[83] V. Ganesan et al. “Primordial non-Gaussian signatures in CMB polarization”. In: *Journal of Cosmology and Astroparticle Physics* 2015.2 (Feb. 2015), pp. 028–028. DOI: 10.1088/1475-7516/2015/02/028. arXiv: 1411.5256 [astro-ph.CO].

[84] K. Ganga et al. “Interplanetary Dust as a Foreground for the LiteBIRD CMB Satellite Mission”. In: *The Open Journal of Astrophysics* 4 (Dec. 2021). DOI: 10.21105/astro.2110.13856.

[85] D. Geller and D. Marinucci. “Spin Wavelets on the Sphere”. In: *Journal of Fourier Analysis and Applications* 16.6 (Dec. 2010), pp. 840–884. DOI: 10.1007/s00041-010-9128-3. arXiv: 0811.2935 [math.CA].

[86] M. Gerbino et al. “Likelihood methods for CMB experiments”. In: *Frontiers in Physics* 8, 15 (Feb. 2020), p. 15. DOI: 10.3389/fphy.2020.00015. arXiv: 1909.09375 [astro-ph.CO].

[87] S. L. Glashow. “Partial-symmetries of weak interactions”. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692.

[88] K. M. Górski et al. “HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere”. In: *The Astrophysical Journal* 622.2 (Apr. 2005), pp. 759–771. DOI: 10.1086/427976. arXiv: astro-ph/0409513 [astro-ph].

[89] J. Grain, M. Tristram, and R. Stompor. “Polarized CMB power spectrum estimation using the pure pseudo-cross-spectrum approach”. In: *Phys. Rev. D.* 79.12, 123515 (June 2009), p. 123515. DOI: 10.1103/PhysRevD.79.123515. arXiv: 0903.2350 [astro-ph.CO].

[90] M. B. Gralla et al. “A measurement of the millimetre emission and the Sunyaev-Zel’dovich effect associated with low-frequency radio sources”. In: *Monthly Notices of the Royal Astronomical Society* 445.1 (Nov. 2014), pp. 460–478. DOI: 10.1093/mnras/stu1592. arXiv: 1310.8281 [astro-ph.CO].

[91] A. H. Guth. “Inflationary universe: A possible solution to the horizon and flatness problems”. In: *Phys. Rev. D* 23 (2 Jan. 1981), pp. 347–356. DOI: 10.1103/PhysRevD.23.347. URL: https://link.aps.org/doi/10.1103/PhysRevD.23.347.

[92] A. H. Guth and S.-Y. Pi. “Fluctuations in the New Inflationary Universe”. In: *Phys. Rev. Lett.* 49 (15 Oct. 1982), pp. 1110–1113. DOI: 10.1103/PhysRevLett.49.1110. URL: https://link.aps.org/doi/10.1103/PhysRevLett.49.1110.

[93] S. Hamimeche and A. Lewis. “Likelihood analysis of CMB temperature and polarization power spectra”. In: *Phys. Rev. D* 77 (10 May 2008), p. 103013. DOI: 10.1103/PhysRevD.77.103013. URL: https://link.aps.org/doi/10.1103/PhysRevD.77.103013.

[94] S. Hamimeche and A. Lewis. “Properties and use of CMB power spectrum likelihoods”. In: *Phys. Rev. D.* 79.8, 083012 (Apr. 2009), p. 083012. DOI: 10.1103/PhysRevD.79.083012. arXiv: 0902.0674 [astro-ph.CO].

[95] S. Hanany et al. “PICO: Probe of Inflation and Cosmic Origins”. In: *arXiv e-prints*, arXiv:1902.10541 (Feb. 2019), arXiv:1902.10541. DOI: 10.48550/arXiv.1902.10541. arXiv: 1902.10541 [astro-ph.IM].

[96] N. Hand et al. “Evidence of Galaxy Cluster Motions with the Kinematic Sunyaev-Zel’dovich Effect”. In: *Physical Review Letters* 109.4, 041101 (July 2012), p. 041101. DOI: 10.1103/PhysRevLett.109.041101. arXiv: 1203.4219 [astro-ph.CO].

[97] D. Hanson et al. “Detection of B-Mode Polarization in the Cosmic Microwave Background with Data from the South Pole Telescope”. In: *Physical Review Letters* 111.14, 141301 (Oct. 2013), p. 141301. DOI: 10.1103/PhysRevLett.111.141301. arXiv: 1307.5830 [astro-ph.CO].

[98] J. Hartlap, P. Simon, and P. Schneider. “Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix”. In: *Astronomy and Astrophysics* 464.1 (2007), pp. 399–404. DOI: 10.1051/0004-6361:20066170. URL: https://doi.org/10.1051/0004-6361:20066170.

[99] S. W. Hawking. “The development of irregularities in a single bubble inflationary universe”. In: *Physics Letters B* 115.4 (1982), pp. 295–297. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(82)90373-2. URL: https://www.sciencedirect.com/science/article/pii/0370269382903732.

[100] E. Hertig et al. “The Simons Observatory: Combining cross-spectral foreground cleaning with multitracer B-mode delensing for improved constraints on inflation”. In: *Phys. Rev. D.* 110.4, 043532 (Aug. 2024), p. 043532. DOI: 10.1103/PhysRevD.110.043532. arXiv: 2405.01621 [astro-ph.CO].

[101] C. Hervías-Caimapo et al. “The Simons Observatory: validation of reconstructed power spectra from simulated filtered maps for the small aperture telescope survey”. In: *Journal of Cosmology and Astroparticle Physics* 2025.6, 055 (June 2025), p. 055. DOI: 10.1088/1475-7516/2025/06/055. arXiv: 2502.00946 [astro-ph.CO].

[102] M. Hilton et al. “The Atacama Cosmology Telescope: The Two-season ACTPol Sunyaev-Zel’dovich Effect Selected Cluster Catalog”. In: *ApJ Supplements* 235.1, 20 (Mar. 2018), p. 20. DOI: 10.3847/1538-4365/aaa6cb. arXiv: 1709.05600 [astro-ph.CO].

[103] E. Hivon et al. “MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets”. In: *The Astrophysical Journal* 567.1 (Mar. 2002), p. 2. DOI: 10.1086/338126. URL: https://doi.org/10.1086/338126.

[104] I. Hothi et al. “Wavelet-based statistics for enhanced 21cm EoR parameter constraints”. In: *A&A* 686 (2024), A212. DOI: 10.1051/0004-6361/202348444. URL: https://doi.org/10.1051/0004-6361/202348444.

[105] I. Hothi et al. “Generative models of 21 cm Epoch of Reionisation lightcones with 3D scattering transforms”. In: *A&A* 706 (2026), A15. DOI: 10.1051/0004-6361/202555896. URL: https://doi.org/10.1051/0004-6361/202555896.

[106] E. de la Hoz et al. “LiteBIRD science goals and forecasts: constraining isotropic cosmic birefringence”. In: *JCAP* 07 (2025), p. 083. DOI: 10.1088/1475-7516/2025/07/083. arXiv: 2503.22322 [astro-ph.CO].

[107] W. Hu. “Lecture Notes on CMB Theory: From Nucleosynthesis to Recombination”. In: *arXiv e-prints*, arXiv:0802.3688 (Feb. 2008), arXiv:0802.3688. DOI: 10.48550/arXiv.0802.3688. arXiv: 0802.3688 [astro-ph].

[108] W. Hu and N. Sugiyama. “Anisotropies in the Cosmic Microwave Background: an Analytic Approach”. In: *The Astrophysical Journal* 444 (May 1995), p. 489. DOI: 10.1086/175624. arXiv: astro-ph/9407093 [astro-ph].

[109] W. Hu and M. White. “A CMB polarization primer”. In: *New Astronomy* 2.4 (1997), pp. 323–344. ISSN: 1384-1076. DOI: https://doi.org/10.1016/S1384-1076(97)00022-5. URL: https://www.sciencedirect.com/science/article/pii/S1384107697000225.

[110] E. Hubble and M. L. Humason. “The Velocity-Distance Relation among Extra-Galactic Nebulae”. In: *The Astrophysical Journal* 74 (July 1931), p. 43. DOI: 10.1086/143323.

[111] M. Kamionkowski and E. D. Kovetz. “The Quest for B Modes from Inflationary Gravitational Waves”. In: *Annual Review of Astronomy and Astrophysics* 54 (Sept. 2016), pp. 227–269. DOI: 10.1146/annurev-astro-081915-023433. arXiv: 1510.06042 [astro-ph.CO].

[112] A. Kogut et al. “Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Polarization”. In: *The Astrophysical Journal* 665.1 (Aug. 2007), pp. 355–362. DOI: 10.1086/519754. arXiv: 0704.3991 [astro-ph].

[113] J. M. Kovac et al. “Detection of polarization in the cosmic microwave background using DASI”. In: *Nature* 420.6917 (Dec. 2002), pp. 772–787. DOI: 10.1038/nature01269. arXiv: astro-ph/0209478 [astro-ph].

[114] N. Krachmalnicoff et al. “Characterization of foreground emission on degree angular scales for CMB B-mode observations. Thermal dust and synchrotron signal from Planck and WMAP data”. In: *Astronomy and Astrophysics* 588, A65 (Apr. 2016), A65. DOI: 10.1051/0004-6361/201527678. arXiv: 1511.00532 [astro-ph.CO].

[115] N. Krachmalnicoff et al. “S-PASS view of polarized Galactic synchrotron at 2.3 GHz as a contaminant to CMB observations”. In: *Astronomy and Astrophysics* 618, A166 (Oct. 2018), A166. DOI: 10.1051/0004-6361/201832768. arXiv: 1802.01145 [astro-ph.GA].

[116] H. Kragh. “The Most Philosophically Important of All the Sciences?: Karl Popper and Physical Cosmology”. In: *Perspectives on Science* 21.3 (2013), pp. 325–357. DOI: 10.1162/posc_{a}_00102.

[117] A. Kusaka et al. “Publisher’s Note: “Modulation of cosmic microwave background polarization with a warm rapidly rotating half-wave plate on the Atacama B-Mode Search instrument” [Rev. Sci. Instrum. 85, 024501 (2014)]”. In: *Review of Scientific Instruments* 85.3, 039901 (Mar. 2014), p. 039901. DOI: 10.1063/1.4867655. arXiv: 1310.3711 [astro-ph.IM].

[118] M. Lei and S. E. Clark. “Probing the Cold Neutral Medium through H I Emission Morphology with the Scattering Transform”. In: *The Astrophysical Journal* 947.2 (Apr. 2023), p. 74. DOI: 10.3847/1538-4357/acc02a. URL: https://doi.org/10.3847/1538-4357/acc02a.

[119] G. Lemaître. “Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques”. In: *Annales de la Société Scientifique de Bruxelles* 47 (Jan. 1927), pp. 49–59.

[120] G. Lemaître. “L’Hypothèse de l’Atome Primitif”. In: *Publications du Laboratoire d’Astronomie et de Geodesie de l’Universite de Louvain* 12 (Jan. 1949), pp. D1–D19.

[121] A. Lewis and A. Challinor. “Weak gravitational lensing of the CMB”. In: *Phys. Rep.* 429.1 (June 2006), pp. 1–65. DOI: 10.1016/j.physrep.2006.03.002. arXiv: astro-ph/0601594 [astro-ph].

[122] A. Lewis, A. Challinor, and N. Turok. “Analysis of CMB polarization on an incomplete sky”. In: *Phys. Rev. D* 65 (2 Dec. 2001), p. 023505. DOI: 10.1103/PhysRevD.65.023505. URL: https://link.aps.org/doi/10.1103/PhysRevD.65.023505.

[123] A. R. Liddle. *An Introduction to Modern Cosmology, Second Edition*. Wiley-VCH, Apr. 2003. ISBN: 0-470-84834-0.

[124] A. R. Liddle and D. H. Lyth. “COBE, gravitational waves, inflation and extended inflation”. In: *Physics Letters B* 291.4 (1992), pp. 391–398. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(92)91393-N. URL: https://www.sciencedirect.com/science/article/pii/037026939291393N.

[125] A. Linde. “A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems”. In: *Physics Letters B* 108.6 (1982), pp. 389–393. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(82)91219-9. URL: https://www.sciencedirect.com/science/article/pii/0370269382912199.

[126] A. Linde. “Inflationary Cosmology”. In: *Physica Scripta* 2000.T85 (Jan. 2000), p. 168. DOI: 10.1238/Physica.Topical.085a00168. URL: https://doi.org/10.1238/Physica.Topical.085a00168.

[127] LiteBIRD Collaboration et al. “Probing cosmic inflation with the LiteBIRD cosmic microwave background polarization survey”. In: *Progress of Theoretical and Experimental Physics* 2023.4, 042F01 (Apr. 2023), 042F01. DOI: 10.1093/ptep/ptac150. arXiv: 2202.02773 [astro-ph.IM].

[128] LiteBIRD Collaboration et al. “LiteBIRD Mission Overview after Mission Reformation”. In: *arXiv e-prints*, arXiv:2607.25285 (July 2026), arXiv:2607.25285. arXiv: 2607.25285 [astro-ph.IM].

[129] M. López-Caniego et al. “POLARIZATION OF THE WMAP POINT SOURCES”. In: *The Astrophysical Journal* 705.1 (Oct. 2009), p. 868. DOI: 10.1088/0004-637X/705/1/868. URL: https://doi.org/10.1088/0004-637X/705/1/868.

[130] T. Louis et al. “The Atacama Cosmology Telescope: DR6 power spectra, likelihoods and ΛCDM parameters”. In: *Journal of Cosmology and Astroparticle Physics* 2025.11, 062 (Nov. 2025), p. 062. DOI: 10.1088/1475-7516/2025/11/062. arXiv: 2503.14452 [astro-ph.CO].

[131] C. J. MacTavish et al. “Cosmological Parameters from the 2003 Flight of BOOMERANG”. In: *The Astrophysical Journal* 647.2 (Aug. 2006), p. 799. DOI: 10.1086/505558. URL: https://doi.org/10.1086/505558.

[132] S. Mallat. “Group Invariant Scattering”. In: *arXiv e-prints*, arXiv:1101.2286 (Jan. 2011), arXiv:1101.2286. DOI: 10.48550/arXiv.1101.2286. arXiv: 1101.2286 [math.FA].

[133] D. Marinucci et al. “Spherical needlets for cosmic microwave background data analysis”. In: *Monthly Notices of the Royal Astronomical Society* 383.2 (Dec. 2007), pp. 539–545. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.12550.x. eprint: https://academic.oup.com/mnras/article-pdf/383/2/539/18571602/mnras0383-0539.pdf. URL: https://doi.org/10.1111/j.1365-2966.2007.12550.x.

[134] D. J. E. Marsh. “Axion cosmology”. In: *Physics Reports* 643 (2016). Axion cosmology, pp. 1–79. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2016.06.005. URL: https://www.sciencedirect.com/science/article/pii/S0370157316301557.

[135] J. Martin, C. Ringeval, and V. Vennin. “Encyclopædia Inflationaris”. In: *Physics of the Dark Universe* 5-6 (2014). Hunt for Dark Matter, pp. 75–235. ISSN: 2212-6864. DOI: https://doi.org/10.1016/j.dark.2014.01.003. URL: https://www.sciencedirect.com/science/article/pii/S2212686414000053.

[136] E. Martínez-González et al. “Cosmic microwave background power spectrum estimation and map reconstruction with the expectation-maximization algorithm”. In: *Monthly Notices of the Royal Astronomical Society* 345.4 (Nov. 2003), pp. 1101–1109. ISSN: 0035-8711. DOI: 10.1046/j.1365-2966.2003.06885.x. eprint: https://academic.oup.com/mnras/article-pdf/345/4/1101/3696319/345-4-1101.pdf. URL: https://doi.org/10.1046/j.1365-2966.2003.06885.x.

[137] G. Martínez-Solaeche, A. Karakci, and J. Delabrouille. “A 3D model of polarized dust emission in the Milky Way”. In: *Monthly Notices of the Royal Astronomical Society* 476.1 (May 2018), pp. 1310–1330. DOI: 10.1093/mnras/sty204. arXiv: 1706.04162 [astro-ph.CO].

[138] F. A. Martire, R. B. Barreiro, and E. Martínez-González. “Characterization of the polarized synchrotron emission from Planck and WMAP data”. In: *Journal of Cosmology and Astroparticle Physics* 2022.04 (Apr. 2022), p. 003. DOI: 10.1088/1475-7516/2022/04/003. URL: https://doi.org/10.1088/1475-7516/2022/04/003.

[139] M. Massardi et al. “A polarization survey of bright extragalactic AT20G sources”. In: *Monthly Notices of the Royal Astronomical Society* 436.4 (Dec. 2013), pp. 2915–2928. ISSN: 0035-8711. DOI: 10.1093/mnras/stt1717. eprint: https://academic.oup.com/mnras/article-pdf/436/4/2915/13763025/stt1717.pdf. URL: https://doi.org/10.1093/mnras/stt1717.

[140] J. C. Mather et al. “Measurement of the Cosmic Microwave Background Spectrum by the COBE FIRAS Instrument”. In: *The Astrophysical Journal* 420 (Jan. 1994), p. 439. DOI: 10.1086/173574.

[141] A. McKellar. “Molecular Lines from the Lowest States of Diatomic Molecules Composed of Atoms Probably Present in Interstellar Space”. In: *Publications of the Dominion Astrophysical Observatory Victoria* 7 (Jan. 1941), p. 251.

[142] K. R. Mecke, T. Buchert, and H. Wagner. “Robust morphological measures for large-scale structure in the Universe”. In: *Astronomy and Astrophysics* 288 (Aug. 1994), pp. 697–704. DOI: 10.48550/arXiv.astro-ph/9312028. arXiv: astro-ph/9312028 [astro-ph].

[143] F. Menanteau et al. “The Atacama Cosmology Telescope: ACT-CL J0102-4915 “El Gordo,” a Massive Merging Cluster at Redshift 0.87”. In: *The Astrophysical Journal* 748.1, 7 (Mar. 2012), p. 7. DOI: 10.1088/0004-637X/748/1/7. arXiv: 1109.0953 [astro-ph.CO].

[144] H. Minkowski. “Volumen und Oberfläche”. In: *Mathematische Annalen* 57 (Dec. 1903), pp. 447–495. DOI: 10.1007/BF01445180. URL: https://doi.org/10.1007/BF01445180.

[145] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. San Francisco: W. H. Freeman, 1973. ISBN: 978-0-7167-0344-0.

[146] R. Morel et al. “Scale Dependencies and Self-Similar Models with Wavelet Scattering Spectra”. In: *arXiv e-prints* (Apr. 2022). eprint: 2204.10177 (physics.data-an).

[147] L. Mousset. “Exploring the primordial universe, inflation and primordial gravitational waves with QUBIC, the Q & U Bolometric Interferometer for Cosmology.” Theses. Université Paris Cité, Oct. 2021. DOI: 10.70675/9670a6e4zf25az40f6zb4edzc8488f91131a. URL: https://theses.hal.science/tel-03783687.

[148] L. Mousset et al. “Generative models of astrophysical fields with scattering transforms on the sphere”. In: *Astronomy and Astrophysics* 691 (2024), A269. DOI: 10.1051/0004-6361/202451396. URL: https://doi.org/10.1051/0004-6361/202451396.

[149] S. Naess and T. Louis. “Large-scale power loss in ground-based CMB mapmaking”. In: *The Open Journal of Astrophysics* 6, 21 (July 2023), p. 21. DOI: 10.21105/astro.2210.02243. arXiv: 2210.02243 [astro-ph.IM].

[150] S. Naess et al. “The Atacama Cosmology Telescope: DR6 maps”. In: *Journal of Cosmology and Astroparticle Physics* 2025.11, 061 (Nov. 2025), p. 061. DOI: 10.1088/1475-7516/2025/11/061. arXiv: 2503.14451 [astro-ph.CO].

[151] F. J. Narcowich, P. Petrushev, and J. D. Ward. “Localized Tight Frames on Spheres”. In: *SIAM Journal on Mathematical Analysis* 38.2 (2006), pp. 574–594. DOI: 10.1137/040614359. eprint: https://doi.org/10.1137/040614359. URL: https://doi.org/10.1137/040614359.

[152] E. A. Ohm. “Receiving System”. In: *Bell System Technical Journal* 40.4 (1961), pp. 1065–1094. DOI: https://doi.org/10.1002/j.1538-7305.1961.tb01638.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1961.tb01638.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1961.tb01638.x.

[153] D. Paoletti et al. “LiteBIRD science goals and forecasts: primordial magnetic fields”. In: *JCAP* 07 (2024), p. 086. DOI: 10.1088/1475-7516/2024/07/086. arXiv: 2403.16763 [astro-ph.CO].

[154] A. A. Penzias. *The Origin of Elements*. Nobel Lecture. 1978. URL: https://www.nobelprize.org/prizes/physics/1978/penzias/lecture/.

[155] A. A. Penzias and R. W. Wilson. “A Measurement of Excess Antenna Temperature at 4080 Mc/s.” In: *The Astrophysical Journal* 142 (July 1965), pp. 419–421. DOI: 10.1086/148307.

[156] S. Perlmutter et al. “Measurements of Ω and Λ from 42 High-Redshift Supernovae”. In: *The Astrophysical Journal* 517.2 (June 1999), p. 565. DOI: 10.1086/307221. URL: https://doi.org/10.1086/307221.

[157] S. Pierre et al. “Bayesian imaging inverse problem with scattering transform”. In: *arXiv e-prints*, arXiv:2602.05816 (Feb. 2026), arXiv:2602.05816. DOI: 10.48550/arXiv.2602.05816. arXiv: 2602.05816 [astro-ph.IM].

[158] Planck Collaboration et al. “Planck intermediate results. XXII. Frequency dependence of thermal emission from Galactic dust in intensity and polarization”. In: *Astronomy and Astrophysics* 576, A107 (Apr. 2015), A107. DOI: 10.1051/0004-6361/201424088. arXiv: 1405.0874 [astro-ph.GA].

[159] Planck Collaboration et al. “Planck 2015 results. X. Diffuse component separation: Foreground maps”. In: *Astronomy and Astrophysics* 594, A10 (Sept. 2016), A10. DOI: 10.1051/0004-6361/201525967. arXiv: 1502.01588 [astro-ph.CO].

[160] Planck Collaboration et al. “Planck 2015 results. XIX. Constraints on primordial magnetic fields”. In: *Astronomy and Astrophysics* 594, A19 (Sept. 2016), A19. DOI: 10.1051/0004-6361/201525821. arXiv: 1502.01594 [astro-ph.CO].

[161] Planck Collaboration et al. “Planck 2015 results. XVI. Isotropy and statistics of the CMB”. In: *Astronomy and Astrophysics* 594, A16 (Sept. 2016), A16. DOI: 10.1051/0004-6361/201526681. arXiv: 1506.07135 [astro-ph.CO].

[162] Planck Collaboration et al. “Planck 2015 results. XXVI. The Second Planck Catalogue of Compact Sources”. In: *Astronomy and Astrophysics* 594, A26 (Sept. 2016), A26. DOI: 10.1051/0004-6361/201526914. arXiv: 1507.02058 [astro-ph.CO].

[163] Planck Collaboration et al. “Planck 2015 results. XXVII. The second Planck catalogue of Sunyaev-Zeldovich sources”. In: *Astronomy and Astrophysics* 594, A27 (Sept. 2016), A27. DOI: 10.1051/0004-6361/201525823. arXiv: 1502.01598 [astro-ph.CO].

[164] Planck Collaboration et al. “Planck intermediate results. XXX. The angular power spectrum of polarized dust emission at intermediate and high Galactic latitudes”. In: *Astronomy and Astrophysics* 586, A133 (Feb. 2016), A133. DOI: 10.1051/0004-6361/201425034. arXiv: 1409.5738 [astro-ph.CO].

[165] Planck Collaboration et al. “Planck 2018 results - IX. Constraints on primordial non-Gaussianity”. In: *Astronomy and Astrophysics* 641 (2020), A9. DOI: 10.1051/0004-6361/201935891. URL: https://doi.org/10.1051/0004-6361/201935891.

[166] Planck Collaboration et al. “Planck 2018 results - VI. Cosmological parameters”. In: *A&A* 641 (2020), A6. DOI: 10.1051/0004-6361/201833910. URL: https://doi.org/10.1051/0004-6361/201833910.

[167] Planck Collaboration et al. “Planck 2018 results - X. Constraints on inflation”. In: *A&A* 641 (2020), A10. DOI: 10.1051/0004-6361/201833887. URL: https://doi.org/10.1051/0004-6361/201833887.

[168] Planck Collaboration et al. “Planck 2018 results. I. Overview and the cosmological legacy of Planck”. In: *Astronomy and Astrophysics* 641, A1 (Sept. 2020), A1. DOI: 10.1051/0004-6361/201833880. arXiv: 1807.06205 [astro-ph.CO].

[169] Planck Collaboration et al. “Planck 2018 results. IV. Diffuse component separation”. In: *Astronomy and Astrophysics* 641, A4 (Sept. 2020), A4. DOI: 10.1051/0004-6361/201833881. arXiv: 1807.06208 [astro-ph.CO].

[170] Planck Collaboration et al. “Planck 2018 results. V. CMB power spectra and likelihoods”. In: *Astronomy and Astrophysics* 641, A5 (Sept. 2020), A5. DOI: 10.1051/0004-6361/201936386. arXiv: 1907.12875 [astro-ph.CO].

[171] Planck Collaboration et al. “Planck 2018 results. VII. Isotropy and statistics of the CMB”. In: *Astronomy and Astrophysics* 641, A7 (Sept. 2020), A7. DOI: 10.1051/0004-6361/201935201. arXiv: 1906.02552 [astro-ph.CO].

[172] Planck Collaboration et al. “Planck 2018 results. XI. Polarized dust foregrounds”. In: *Astronomy and Astrophysics* 641, A11 (Sept. 2020), A11. DOI: 10.1051/0004-6361/201832618. arXiv: 1801.04945 [astro-ph.GA].

[173] Planck Collaboration et al. “Planck intermediate results. LVII. Joint Planck LFI and HFI data processing”. In: *Astronomy and Astrophysics* 643, A42 (Nov. 2020), A42. DOI: 10.1051/0004-6361/202038073. arXiv: 2007.04997 [astro-ph.CO].

[174] Polarbear Collaboration et al. “A Measurement of the Cosmic Microwave Background B-mode Polarization Power Spectrum at Sub-degree Scales with POLARBEAR”. In: *The Astrophysical Journal* 794.2, 171 (Oct. 2014), p. 171. DOI: 10.1088/0004-637X/794/2/171. arXiv: 1403.2369 [astro-ph.CO].

[175] D. Poletti and J. Errard. *FGBuster: Parametric component separation for Cosmic Microwave Background observations*. Astrophysics Source Code Library, record ascl:2307.021. July 2023. ascl: 2307.021.

[176] G. Puglisi et al. “Improved galactic foreground removal for B-mode detection with clustering methods”. In: *Monthly Notices of the Royal Astronomical Society* 511.2 (Apr. 2022), pp. 2052–2074. DOI: 10.1093/mnras/stac069. arXiv: 2109.11562 [astro-ph.CO].

[177] Z. Qiu and Q. Huang. “Inflation in a nutshell: From basics to latest advances”. In: *arXiv e-prints* (2026). arXiv: 2606.24474 [astro-ph.CO]. URL: https://arxiv.org/abs/2606.24474.

[178] C. Ranucci et al. “B-sure. Part I. Minkowski functionals as robustness test for tensor-to-scalar ratio detection from CMB observations”. In: *Journal of Cosmology and Astroparticle Physics* 2026.01 (Jan. 2026), p. 045. DOI: 10.1088/1475-7516/2026/01/045. URL: https://doi.org/10.1088/1475-7516/2026/01/045.

[179] C. Ranucci et al. “B-sure. Part II. Scattering transforms as robustness test for tensor-to-scalar ratio detection from CMB observations”. In: *arXiv e-prints* (Sept. 2026). arXiv: 2609.17531 [astro-ph.CO].

[180] M. Rashid, M. L. Brown, and D. B. Thomas. “CMB polarization signal demodulation with a rotating half-wave plate”. In: *Monthly Notices of the Royal Astronomical Society* 532.2 (Aug. 2024), pp. 2309–2319. DOI: 10.1093/mnras/stae1619. arXiv: 2307.01860 [astro-ph.IM].

[181] M. J. Rees and D. W. Sciama. “Large-scale Density Inhomogeneities in the Universe”. In: *Nature* 217.5128 (Feb. 1968), pp. 511–516. DOI: 10.1038/217511a0.

[182] B. Regaldo-Saint Blancard et al. “Statistical description of dust polarized emission from the diffuse interstellar medium - A RWST approach”. In: *Astronomy and Astrophysics* 642 (2020), A217. DOI: 10.1051/0004-6361/202038044. URL: https://doi.org/10.1051/0004-6361/202038044.

[183] A. G. Riess et al. “Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant”. In: *The Astronomical Journal* 116.3 (Sept. 1998), p. 1009. DOI: 10.1086/300499. URL: https://doi.org/10.1086/300499.

[184] G. B. Rybicki and A. P. Lightman. *Radiative processes in astrophysics*. Wiley-VCH, 1979.

[185] A. Salam. “WEAK AND ELECTROMAGNETIC INTERACTIONS.” In: *pp 367-77 of Elementary Particle Theory. Svartholm, Nils (ed.). New York, John Wiley and Sons, Inc., 1968.* (Oct. 1969). URL: https://www.osti.gov/biblio/4767615.

[186] L. Santos et al. “Statistical imprints of CMB B-type polarization leakage in an incomplete sky survey analysis”. In: *Journal of Cosmology and Astroparticle Physics* 2017.01 (Jan. 2017), p. 043. DOI: 10.1088/1475-7516/2017/01/043. URL: https://doi.org/10.1088/1475-7516/2017/01/043.

[187] J. Schmalzing and T. Buchert. “Beyond Genus Statistics: A Unifying Approach to the Morphology of Cosmic Structure”. In: *ApJ Letters* 482.1 (June 1997), pp. L1–L4. DOI: 10.1086/310680. arXiv: astro-ph/9702130 [astro-ph].

[188] J. Schmalzing and K. M. Gorski. “Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps”. In: *Monthly Notices of the Royal Astronomical Society* 297.2 (June 1998), pp. 355–365. DOI: 10.1046/j.1365-8711.1998.01467.x. arXiv: astro-ph/9710185 [astro-ph].

[189] D. M. Scolnic et al. “The Complete Light-curve Sample of Spectroscopically Confirmed SNe Ia from Pan-STARRS1 and Cosmological Constraints from the Combined Pantheon Sample”. In: *The Astrophysical Journal* 859.2 (May 2018), p. 101. DOI: 10.3847/1538-4357/aab9bb. URL: https://doi.org/10.3847/1538-4357/aab9bb.

[190] U. Seljak. “Measuring Polarization in the Cosmic Microwave Background”. In: *The Astrophysical Journal* 482.1 (June 1997), pp. 6–16. DOI: 10.1086/304123. arXiv: astro-ph/9608131 [astro-ph].

[191] U. Seljak and M. Zaldarriaga. “Signature of Gravity Waves in the Polarization of the Microwave Background”. In: *Phys. Rev. Lett.* 78 (11 Mar. 1997), pp. 2054–2057. DOI: 10.1103/PhysRevLett.78.2054. URL: https://link.aps.org/doi/10.1103/PhysRevLett.78.2054.

[192] B. D. Sherwin et al. “Two-season Atacama Cosmology Telescope polarimeter lensing power spectrum”. In: *Phys. Rev. D* 95 (12 June 2017), p. 123529. DOI: 10.1103/PhysRevD.95.123529. URL: https://link.aps.org/doi/10.1103/PhysRevD.95.123529.

[193] J. Silk. “Cosmic Black-Body Radiation and Galaxy Formation”. In: *The Astrophysical Journal* 151 (Feb. 1968), p. 459. DOI: 10.1086/149449.

[194] A. A. Starobinsky. “A new type of isotropic cosmological models without singularity”. In: *Physics Letters B* 91.1 (1980), pp. 99–102. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(80)90670-X. URL: https://www.sciencedirect.com/science/article/pii/037026938090670X.

[195] A. A. Starobinsky. “Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations”. In: *Physics Letters B* 117.3 (1982), pp. 175–178. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(82)90541-X. URL: https://www.sciencedirect.com/science/article/pii/037026938290541X.

[196] R. Stompor et al. “Maximum likelihood algorithm for parametric component separation in cosmic microwave background experiments”. In: *Monthly Notices of the Royal Astronomical Society* 392.1 (Dec. 2008), pp. 216–232. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2008.14023.x. URL: https://doi.org/10.1111/j.1365-2966.2008.14023.x.

[197] R. M. Sullivan, L. T. Hergt, and D. Scott. “Methods for CMB Map Analysis”. In: *Research Notes of the American Astronomical Society* 9.2, 43 (Feb. 2025), p. 43. DOI: 10.3847/2515-5172/adb610. arXiv: 2410.12951 [astro-ph.CO].

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

[199] A. Suzuki et al. “The Polarbear-2 and the Simons Array Experiments”. In: *Journal of Low Temperature Physics* 184.3-4 (Aug. 2016), pp. 805–810. DOI: 10.1007/s10909-015-1425-4. arXiv: 1512.07299 [astro-ph.IM].

[200] S. Takakura et al. “Performance of a continuously rotating half-wave plate on the POLARBEAR telescope”. In: *Journal of Cosmology and Astroparticle Physics* 2017.5, 008 (May 2017), p. 008. DOI: 10.1088/1475-7516/2017/05/008. arXiv: 1702.07111 [astro-ph.IM].

[201] M. Tegmark and G. Efstathiou. “A method for subtracting foregrounds from multifrequency CMB sky maps”. In: *Monthly Notices of the Royal Astronomical Society* 281.4 (Aug. 1996), pp. 1297–1314. ISSN: 0035-8711. DOI: 10.1093/mnras/281.4.1297. eprint: https://academic.oup.com/mnras/article-pdf/281/4/1297/3697403/281-4-1297.pdf. URL: https://doi.org/10.1093/mnras/281.4.1297.

[202] M. Tegmark, A. de Oliveira-Costa, and A. J. S. Hamilton. “High resolution foreground cleaned CMB map from WMAP”. In: *Phys. Rev. D* 68 (12 Dec. 2003), p. 123523. DOI: 10.1103/PhysRevD.68.123523. URL: https://link.aps.org/doi/10.1103/PhysRevD.68.123523.

[203] The Pan-Experiment Galactic Science Group et al. “Full-sky Models of Galactic Microwave Emission and Polarization at Sub-arcminute Scales for the Python Sky Model”. In: *arXiv e-prints*, arXiv:2502.20452 (Feb. 2025), arXiv:2502.20452. DOI: 10.48550/arXiv.2502.20452. arXiv: 2502.20452 [astro-ph.CO].

[204] B. Thorne et al. “The Python Sky Model: software for simulating the Galactic microwave sky”. In: *Monthly Notices of the Royal Astronomical Society* 469.3 (Aug. 2017), pp. 2821–2833. DOI: 10.1093/mnras/stx949. arXiv: 1608.02841 [astro-ph.CO].

[205] M. Tristram et al. “Improved limits on the tensor-to-scalar ratio using BICEP and $Planck$ data”. In: *Phys. Rev. D* 105 (8 Apr. 2022), p. 083524. DOI: 10.1103/PhysRevD.105.083524. URL: https://link.aps.org/doi/10.1103/PhysRevD.105.083524.

[206] A. Tsouros et al. “Separation of polarized dust emission in Planck observations with Scattering Transforms”. In: *arXiv e-prints*, arXiv:2602.04528 (Feb. 2026), arXiv:2602.04528. DOI: 10.48550/arXiv.2602.04528. arXiv: 2602.04528 [astro-ph.CO].

[207] L. Vacher. “Understanding the galactic polarized signal in the quest for new fundamental physics in the cosmic microwave background”. Theses. Université Paul Sabatier - Toulouse III, Sept. 2023. URL: https://theses.hal.science/tel-04406036.

[208] L. Vacher et al. “How bad could it be? Modelling the 3D complexity of the polarised dust signal using moment expansion”. In: *Astronomy and Astrophysics* 697 (2025), A212. DOI: 10.1051/0004-6361/202453066. URL: https://doi.org/10.1051/0004-6361/202453066.

[209] G. Valogiannis and C. Dvorkin. “Going beyond the galaxy power spectrum: An analysis of BOSS data with wavelet scattering transforms”. In: *Phys. Rev. D* 106 (10 Nov. 2022), p. 103509. DOI: 10.1103/PhysRevD.106.103509. URL: https://link.aps.org/doi/10.1103/PhysRevD.106.103509.

[210] G. Valogiannis and C. Dvorkin. “Towards an optimal estimation of cosmological parameters with the wavelet scattering transform”. In: *Phys. Rev. D* 105 (10 May 2022), p. 103534. DOI: 10.1103/PhysRevD.105.103534. URL: https://link.aps.org/doi/10.1103/PhysRevD.105.103534.

[211] A. van Engelen et al. “A Measurement of Gravitational Lensing of the Microwave Background Using South Pole Telescope Data”. In: *The Astrophysical Journal* 756.2, 142 (Sept. 2012), p. 142. DOI: 10.1088/0004-637X/756/2/142. arXiv: 1202.0546 [astro-ph.CO].

[212] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[213] S. Weinberg. “The cosmological constant problem”. In: *Rev. Mod. Phys.* 61 (1 Jan. 1989), pp. 1–23. DOI: 10.1103/RevModPhys.61.1. URL: https://link.aps.org/doi/10.1103/RevModPhys.61.1.

[214] S. Weinberg. *Cosmology*. OUP Oxford, 2008. ISBN: 9780191523601. URL: https://books.google.it/books?id=nqQZdg020fsC.

[215] M. Welling. “Robust higher order statistics”. In: *International Workshop on Artificial Intelligence and Statistics* (Jan. 2005), pp. 405–412. URL: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=8200InoAAAAJ&cstart=400&pagesize=100&sortby=pubdate&citation_for_view=8200InoAAAAJ:L8Ckcad2t8MC.

[216] D. T. Wilkinson and P. J. E. Peebles. “Discovery of the Cosmic Microwave Background”. In: *Physica Scripta* 2000.T85 (Jan. 2000), p. 136. DOI: 10.1238/Physica.Topical.085a00136. URL: https://doi.org/10.1238/Physica.Topical.085a00136.

[217] S. Winitzki and A. Kosowsky. “Minkowski functional description of microwave background Gaussianity”. In: *New Astronomy* 3.2 (1998), pp. 75–99. ISSN: 1384-1076. DOI: https://doi.org/10.1016/S1384-1076(97)00046-8. URL: https://www.sciencedirect.com/science/article/pii/S1384107697000468.

[218] K. Wolz et al. “The Simons Observatory: Pipeline comparison and validation for large-scale B-modes”. In: *Astronomy and Astrophysics* 686 (2024), A16. DOI: 10.1051/0004-6361/202346105. URL: https://doi.org/10.1051/0004-6361/202346105.

[219] E. L. Wright et al. “Interpretation of the COBE FIRAS CMBR Spectrum”. In: *The Astrophysical Journal* 420 (Jan. 1994), p. 450. DOI: 10.1086/173576.

[220] E. L. Wright et al. “FIVE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE* OBSERVATIONS: SOURCE CATALOG”. In: *The Astrophysical Journal Supplement Series* 180.2 (Feb. 2009), p. 283. DOI: 10.1088/0067-0049/180/2/283. URL: https://doi.org/10.1088/0067-0049/180/2/283.

[221] M. Zaldarriaga and U. Seljak. “All-sky analysis of polarization in the microwave background”. In: *Phys. Rev. D* 55 (4 Feb. 1997), pp. 1830–1840. DOI: 10.1103/PhysRevD.55.1830. URL: https://link.aps.org/doi/10.1103/PhysRevD.55.1830.

[222] M. Zaldarriaga and U. Seljak. “Gravitational lensing effect on cosmic microwave background polarization”. In: *Phys. Rev. D* 58 (2 June 1998), p. 023003. DOI: 10.1103/PhysRevD.58.023003. URL: https://link.aps.org/doi/10.1103/PhysRevD.58.023003.

[223] A. Zonca et al. “healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python”. In: *The Journal of Open Source Software* 4.35, 1298 (Mar. 2019), p. 1298. DOI: 10.21105/joss.01298.

[224] A. Zonca et al. “The Python Sky Model 3 software”. In: *Journal of Open Source Software* 6.67 (2021), p. 3783. DOI: 10.21105/joss.03783. URL: https://doi.org/10.21105/joss.03783.

[225] F. Zwicky. “Die Rotverschiebung von extragalaktischen Nebeln”. In: *Helv. Phys. Acta* 6 (1933), pp. 110–127. DOI: 10.1007/s10714-008-0707-4.