UNIVERSITÀ DEGLI STUDI DI ROMA TOR VERGATA

MACROAREA DI SCIENZE MATEMATICHE, FISICHE E NATURALI

CORSO DI STUDIO IN
FISICA - Curriculum Erasmus Mundus MASS

TESI DI LAUREA IN
Cosmology

TITOLO

Detecting Primordial B-modes in the Cosmic Microwave Background: Tackling Component Separation Challenges

Supervisors:

Prof. Marina Migliaccio
Dr. Alessandro Carones

Candidate:
matricola: *0325655*
Aliza Mustafa

Collaborator:

Dr. Simon Prunet

Anno Accademico 2023/2024

Erasmus Mundus Master
in Astrophysics and Space Science

Master Thesis

Detecting Primordial B-modes in the Cosmic Microwave Background: Tackling Component Separation Challenges

Supervisors

Marina Migliaccio, Alessandro Carones
Physics Department
University of Rome, Tor Vergata

Collaborator

Simon Prunet
Université Côte d'Azur, Nice

Author

Aliza Mustafa

Academic Year 2023/2024

This Master's thesis is submitted in partial fulfilment of the requirements for the degree Master's in Astrophysics and Space Science as part of a multiple degree awarded in the framework of the Erasmus Mundus Joint Master in Astrophysics and Space Science −− MASS jointly delivered by a Consortium of four Universities: Tor Vergata University of Rome, University of Belgrade, University of Bremen, and Université Cote d'Azur, regulated by the MASS Consortium Agreement and funded by the EU under the call ERASMUS-EDU-2021-PEX-EMJM-MOB.

Contents

Acknowledgments

I would like to extend my deepest gratitude to my supervisors, Professor Marina and Alessandro, for their sincere guidance and encouragement throughout this project. Their expertise and thoughtful feedback played a crucial role in shaping the direction of my research, and I feel extremely fortunate to have had the chance to learn from them. I am profoundly grateful for the time they dedicated to the various insightful discussion sessions, which greatly enhanced the quality of this work. The knowledge and skills I have gained under their supervision will remain invaluable to me moving forward. I would also like to acknowledge Simon Prunet for his role as a collaborator for this project.

I would like to thank the MASS coordinating team for their assistance with logistics and administrative matters. Their efforts greatly contributed to the smooth running of the entire process, and I truly appreciate their help. I also want to extend my heartfelt thanks to my friends, who have been a constant source of moral and emotional support throughout this journey. Their understanding, encouragement, and companionship gave me the strength to persevere during challenging times, and I cannot thank them enough for being there for me.

Finally, I express my deepest appreciation to my mother, whose unwavering prayers and boundless support have been a cornerstone of my strength. Her affection and encouragement have consistently inspired me, and I am immensely grateful for all the sacrifices she has made for the sake of my education and ambitions.

Abstract

Next-generation experiments are targeting an accurate measurement of the Cosmic Microwave Background (CMB) polarization. Detecting a specific pattern in CMB polarization, the primordial B-modes, would confirm cosmic inflation and constrain existing models. In this framework, Simons Observatory (SO) represents the most advanced current experiment for observations of B-modes through three Small Aperture Telescopes (SATs). They are designed to observe the CMB polarization and then measure the primordial tensor-to-scalar ratio (r), with a target uncertainty of less than 0.003. This thesis applies for the first time the blind Needlet-ILC (NILC) and Multi-Clustering (MC-NILC) component separation techniques on realistic sky simulations generated from the most recent experimental configuration of SO, to forecast the sensitivity on r, specifically focusing on a 10% sky fraction. We also explored the combination of SO simulated data sets with higher frequency channels accessible from space, for example with Planck and LiteBIRD. Galactic foreground contamination is added considering realistic models with increasing complexity based on state-of-the-art data and simulations, which include spatial variability of the spectral indices of the components. The instrumental noise is incorporated by considering both a simple Gaussian, white and isotropic model and a more realistic description which includes the effects of a non-uniform scanning strategy and a $1/f$ correlated component. The results show that NILC and MC-NILC methods prove to be effective in mitigating contamination by polarized Galactic emission. NILC provides an upper bound on residual foreground contamination at 68% CL of the order of $r \sim 4 \cdot 10^{-3}$ and gives significant bias on r for a more complex foreground model. Whereas, MC-NILC provides an upper bound even for the complex foreground model at 68% CL of the order of $r \sim 2 \cdot 10^{-3}$, which is compatible with the science target of SO. As proved by the lower obtained upper bounds, the MC-NILC methodology results to be more effective in reducing foreground contamination, thanks to an optimization of the spatial domains where component separation is separately implemented.

Motivation

The Cosmic Microwave Background (CMB) is a fundamental observable in current cosmology, representing the residual light of the Big Bang that fills the universe. The accidental discovery of the CMB in 1965 by Arno Penzias and Robert Wilson marked a pivotal moment in the study of cosmology. Their findings provided compelling evidence in support of the Big Bang theory, which suggests that the universe originated from a hot, dense state. For this groundbreaking discovery, Penzias and Wilson were awarded the Nobel Prize in Physics in 1978. Their work not only contributed to a deeper understanding of the origins of the universe but also solidified the Big Bang theory as a central pillar of modern cosmology.

Current cosmology investigations and future experiments are increasingly focusing on CMB polarization. This feature of CMB could disclose a lot more about the early universe, notably about cosmic inflation. Inflation theory claims that the universe expanded rapidly after the Big Bang, this expansion brought on cosmological scales quantum vacuum fluctuations, which served as the density seeds for all the structures we observe in the universe today. Additionally, this expansion is expected to have produced gravitational waves. These waves, in turn, would imprint a distinct signature on the CMB polarization patterns, known as B-modes.

CMB polarization is categorized into two types: E-modes and B modes. E-mode polarization is associated with density fluctuations and acoustic oscillations in the primordial plasma. E-modes are mainly caused by Thomson scattering of CMB photons off unbound electrons before recombination. A curl-like pattern in the polarization of CMB photons represents B-mode polarization. Unlike E-modes, B-modes can be caused by several mechanisms, including gravitational lensing of CMB photons by large-scale structures and gravitational waves from cosmic inflation. Detecting B-mode polarization in the CMB is a key goal in cosmology since it could indicate the presence of primordial gravitational waves from the early cosmos. The amplitude of the B-mode signal is linked to the energy scale of inflation.

Future CMB experiments aim to improve the sensitivity to detect primordial B-modes, which would be a critical step towards proving the inflationary scenario. However, this signal is significantly weaker by a factor of 10^2 to 10^3 than the polarized emission of our Galaxy. Various techniques are being used to separate primordial and foreground emissions, one of which is the Internal Linear Combination (ILC), which stands out for its ability to eliminate foregrounds without considering a specific model for their emissions, which is crucial given our limited understanding of polarized galactic components. To further improve this technique, approaches in Needlet space are introduced in which the separation is done in specific portions of the sky selected to have similar foreground properties, which represents a significant stride in addressing more complex B-mode

foreground contamination challenges.

In this thesis, we investigate the robustness of the above-mentioned component separation techniques and optimise these techniques for application to future CMB experiments like Simons Observatory and LiteBIRD.

Chapter 1

Introduction

1.1 The Standard Model of Cosmology

The Lambda Cold Dark Matter (ΛCDM) model also known as the "Concordance Cosmological Model" is widely accepted as the standard model of cosmology. It provides a comprehensive framework to describe the cosmic evolution and large-scale structure of the universe. According to the cosmological principle, the universe is considered to be isotropic (same in all directions) and homogeneous (uniformly distributed on large scales). Such a model effectively explains a broad range of cosmological observations, including the expansion of the universe, the properties of the cosmic microwave background (CMB) radiation, and the distribution of galaxies.

The fundamental components of ΛCDM model, each of them characterized by an equation of state $p = \rho\omega$ with p the pressure, ρ the density and ω the state parameter, are:

- **Dark Energy:** Approximately 70% of the total energy density of the universe is attributed to dark energy ($\omega < -1/3$), whose phenomenological behaviour, at the moment, is properly described by the cosmological constant Λ having $\omega = -1$. This is responsible for the accelerated expansion of the universe observed in distant Type Ia supernovae [1, 2].

- **Cold Dark Matter:** Around 25% of the universe's energy density consists of cold dark matter with $\omega = 0$. CDM is non-relativistic at the time of decoupling and interacts primarily through gravitational forces, playing a crucial role in the formation of the large-scale structure of the universe [3, 4].

- **Baryonic Matter:** Ordinary matter with $\omega = 0$, which includes protons, neutrons, and electrons, constitutes about 5% of the total energy density and forms stars, planets, and all visible structures in the universe.

- **Radiation:** This includes photons and neutrinos with $\omega = 1/3$. Photons are primarily observed as the CMB, while neutrinos contribute a small fraction to the overall energy density. Radiation is significant in the early universe but has a negligible effect on the dynamics of the current universe.

- **Curvature:** The ΛCDM model generally assumes a flat universe ($k = 0$) based on measurements of the CMB [5].

1.1.1 The Friedmann Equations

The evolution of the universe in the ΛCDM model is governed by the Friedmann equations, which are derived from the equations of General Relativity (GR). These equations describe how the expansion of the universe is influenced by its content, including dark energy (represented by the cosmological constant Λ), dark matter, and ordinary matter.

In cosmology, the expansion of the universe is characterized by the scale factor $a(t)$, a dimensionless quantity that describes how distances between two points in the universe change over time. A scale factor of $a(t) = 1$ is often normalized to the present time, with $a(t) < 1$ corresponding to earlier times (when the universe was smaller) and $a(t) > 1$ corresponding to future times (when the universe will be larger).

Expressed in terms of the scale factor and the density parameters, the first Friedmann equation in the ΛCDM model is given by:

$$\left(\frac{H}{H_0}\right)^2 = \Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_k a^{-2} + \Omega_\Lambda, \tag{1.1}$$

where:

- $H(t) = \dot{\mathbf{a}} = \frac{d\mathbf{a}}{dt}$: Hubble parameter at time t,
- H_0: Hubble constant (current value of the Hubble parameter),
- $\Omega_m(t) = \frac{8\pi G\rho_m}{3H^2}$: density parameter for matter,
- $\Omega_r(t) = \frac{8\pi G\rho_r}{3H^2}$: density parameter for radiation,
- $\Omega_k(t) = \frac{-k}{(aH)^2}$: density parameter for curvature,
- $\Omega_\Lambda(t) = \frac{8\pi G\rho_\Lambda}{3H^2}$: density parameter for dark energy,

The second Friedmann equation, which describes the acceleration of the expansion, is:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \tag{1.2}$$

where ρ and p are defined as the sum of the contributions from all components: matter (ρ_m, p_m), radiation (ρ_r, p_r), and dark energy (ρ_Λ, p_Λ).

1.1.2 Implications of the ΛCDM Model

The ΛCDM model has several important implications for our knowledge of the universe:

- **Cosmic Expansion:** The model explains the observed expansion of the universe. The scale factor $a(t)$ describes how distances in the universe change over time, governed by the Friedmann equations mentioned above.

- **Structure Formation:** CDM provides the gravitational framework for the development of galaxies and clusters. Small initial density perturbations generated by quantum fluctuations grow over time, resulting in the large-scale structures observed today [6].
- **Cosmic Microwave Background:** The CMB provides a snapshot of the early universe, allowing precise measurements of cosmological parameters. The ΛCDM model accurately predicts the anisotropies in the CMB, which are small fluctuations of the order of 10^{-5} in the CMB temperature field (see section 1.13). They are imprinted by primordial density fluctuations [7].
- **Baryon Acoustic Oscillations (BAO):** The ΛCDM model explains the periodic fluctuations in the density field of the visible baryonic matter of the universe, providing a "standard ruler" for measuring cosmic distances [8].

1.1.3 Problems and Challenges of the ΛCDM Model

Despite its successes, the ΛCDM model faces several unresolved issues and challenges:

- **Nature of Dark Energy:** The cosmological constant Λ is an *ad hoc* addition to the model, and its physical nature is still unknown. Alternatives such as quintessence propose a dynamic field, but these theories lack conclusive observational evidence [9].
- **Nature of Dark Matter:** While CDM is a crucial component, its exact nature is still unknown. Numerous candidates, such as Weakly Interacting Massive Particles (WIMPs) and axions, have been proposed, but direct detection is still missing [10].
- **Hubble Tension:** There is a disagreement between the Hubble constant (H_0) values obtained from the early universe (CMB measurements) the and late universe (local distance measurements), indicating potential new physics beyond the ΛCDM model [11, 5].
- **Small-Scale Structure Problems:** The ΛCDM model predicts more small-scale structures, such as satellite galaxies than are observed. This discrepancy is the "missing satellites problem" [12].
- **Curvature Problem:** The curvature problem arises from the fact that the universe seems to be almost flat, which would require fine-tuning of the initial conditions in the early universe. This degree of fine-tuning would be difficult to account for within the standard ΛCDM paradigm. Cosmic inflation addresses the flatness problem by proposing a rapid exponential expansion of the early universe which stretched the universe to near-flatness, making the observable universe appear geometrically flat [5].
- **Horizon Problem:** The horizon problem refers to how distant parts of the universe, which were not in causal contact, have nearly the same temperatures and properties. The standard ΛCDM model does not naturally explain

this uniformity without adding an extra mechanism, such as cosmic inflation, which suggests that the universe expanded rapidly at the early stages, letting distant parts come in causal contact.

The ΛCDM model is phenomenologically very successful in explaining many key observations about the universe but lacks a fundamental theory motivating it. This leaves some crucial questions about the nature of dark matter and dark energy, together with the physics of cosmic inflation, largely unexplained.

1.2 Inflation

In the early 1980s, Alan Guth [13], Andrei Linde [14], and others proposed the inflationary theory. It predicts an era of extremely rapid expansion in the early universe, which occurred shortly after the Big Bang during which the scale factor grew exponentially by a factor given in terms of $a(t_{end})/a(t_{beg}) = e^N$ where N is the *e-folds*. Here, t_{beg} refers to the time at the beginning of inflation and t_{end} refers to the time at the end of inflation. This theory was introduced to address several fundamental problems with the standard ΛCDM model, including the monopole problem, the horizon problem, and the flatness problem (also known as the curvature problem).

- The monopole problem pertains to the prediction of a high density of magnetic monopoles, which are hypothetical particles expected to be present, in the early universe. However, these monopoles are not observed in the universe today, indicating a discrepancy with the standard model. Inflation reduces the density of existing monopoles by spreading them over an exponentially expanding volume, effectively making their density undetectable.
- The horizon problem can be solved by introducing an inflationary period $\left(10^{-36} \lesssim t \lesssim 10^{-32}\right)$, in which the universe expanded exponentially, causing regions that were once in causal contact to be driven far apart thus allowing for the homogenization of the temperature and density.
- Inflation causes the universe to expand rapidly, effectively eliminating any initial curvature and resulting in the appearance of a flat universe regardless of its initial state.

Cosmic inflation is typically predicted to be driven by a single scalar field denoted as $\phi(t)$, known as the inflaton, with a potential $V(\phi)$ having a negative slope, which is sufficiently flat to account for the accelerated expansion [15]. The evolution of the inflaton field $\phi(t)$ is governed by the Klein-Gordon equation in an expanding universe [13, 14]:

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0, \tag{1.3}$$

where $H = \frac{\dot{a}}{a}$ is the Hubble parameter during inflation. The energy density and pressure of the inflaton field which drive the exponential expansion, are given by:

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi); \quad p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi) \tag{1.4}$$

The slow-roll approximation is often used to describe inflation, where the inflaton field slowly rolls down its potential, leading to an almost constant Hubble parameter obtained from the first Friedmann eq. 1.1, when the energy density content of the universe is dominated by the inflation :

$$H^2 \approx \frac{8\pi G}{3} V(\phi).$$

This approximation ensures that the expansion is nearly exponential. The second derivative of the scale factor, $\ddot{a}$ can be written as

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 = H^2 \left(1 - \epsilon_H\right), \tag{1.5}$$

where ϵ_H is the first slow-roll parameter, defined as:

$$\epsilon_H = -\frac{\dot{H}}{H^2} \ll 1. \tag{1.6}$$

The above condition ensures that if the potential is sufficiently flat, $\dot{H}$ is small and $\ddot{a}$ is positive and ϕ slowly rolls down the potential. Inflation ends when the potential becomes too steep, the field then oscillates around the minimum of its potential and decays, creating a radiation-dominated universe and later matter. This process is known as reheating. Quantum fluctuations in the scalar field generate density perturbations. This will be discussed in the next section.

The theory of cosmic inflation is strongly supported by several key pieces of observational evidence that align well with the predictions of inflationary models. Some of the most compelling evidence includes:

- **Support from CMB Observations:** The Planck satellite has provided detailed measurements of the CMB power spectrum, which are in excellent agreement with the predictions of inflationary models.
- **Large-Scale Structure:** The distribution of galaxies and the large-scale structure of the universe are consistent with the initial density perturbations predicted by inflation.

1.2.1 Origin of Cosmic Pertubations

Until now we have discussed an evolution of the *Background universe* which is isotropic and homogenous to a high degree based on the cosmological principle. However, an unperturbed and initially homogenous universe will not give rise to the planets, stars galaxies, and clusters that we observe today. This issue of the standard ΛCDM model is resolved by inflation which not only addresses the horizon and flatness problems but also provides a natural mechanism for generating the initial density perturbations that lead to the formation of the large-scale

structure of the universe [16, 17, 18]. These perturbations are seen as anisotropies in the CMB temperature and the distribution of galaxies in the universe.

During the inflationary period, quantum fluctuations in the inflaton field were stretched to large scales due to the rapid expansion of the universe and then propagated to matter density and metric perturbations, finally leading to the large-scale structures we observe today.

The evolution of quantum fluctuations of the inflaton field $\delta\phi$ in an expanding universe can be described by the Klein-Gordon equation [19]:

$$\ddot{\delta\phi} + 3H\dot{\delta\phi} - \frac{\nabla^2\delta\phi}{a^2} + V''\delta\phi = 0\,. \tag{1.7}$$

where $\dot{\phi} = d\phi/dt$ and $V' = dV/d\phi$.

Scalar Perturbations

The scalar perturbations, which are responsible for the density fluctuations, can be characterized by means of the gauge-invariant curvature perturbation $\mathcal{R}$ which represents the gravitational potential on comoving hypersurfaces where $\delta\phi = 0$. For Gaussian perturbations, the power spectrum fully describes their statistics. The power spectrum of curvature fluctuations, $\mathcal{P}_\mathcal{R}(k)$, is predicted to follow a power-law given by:

$$\mathcal{P}_\mathcal{R}(k) = A_s \left(\frac{k}{k_0}\right)^{n_s - 1}. \tag{1.8}$$

Here, k is the wavenumber defined as $k = \frac{2\pi}{\lambda}$, A_s is the amplitude of the scalar perturbations, n_s is the spectral index, and k_0 is a pivot scale [1].

Tensor Perturbations

Inflation also predicts the generation of tensor perturbations i.e. primordial gravitational waves. The power spectrum of the tensor perturbations, $\mathcal{P}_t(k)$, is given by:

$$\mathcal{P}_t(k) = A_t \left(\frac{k}{k_0}\right)^{n_t}, \tag{1.9}$$

where A_t is the amplitude of the tensor perturbations and n_t is the tensor spectral index.

Future experiments, like the Simons Observatory [20] and the LiteBIRD satellite [21], aim to measure with even greater precision the CMB polarization and primordial gravitational waves, providing further tests of inflationary theories.

Tensor-to-Scalar Ratio

The ratio between the amplitudes of the tensor and scalar power spectra defines the so-called *tensor-to-scalar ratio* r, which quantifies the relative contribution of

[1] The pivot scale, denoted by the wavenumber k_0, is a specific reference scale at which the amplitude of the power spectrum A_s is defined, it is typically set to be 0.05 Mpc^{-1} for CMB measurements.

scalar and tensor perturbations. It reads as [22]:

$$r = \frac{A_t}{A_s} = 16\epsilon\,. \tag{1.10}$$

Since ϵ is expected to be small, we have not yet detected gravitational waves. This ratio is strongly correlated to the model, different inflationary models predict distinct values of r, thus its measurement may allow identifying the correct scenario.

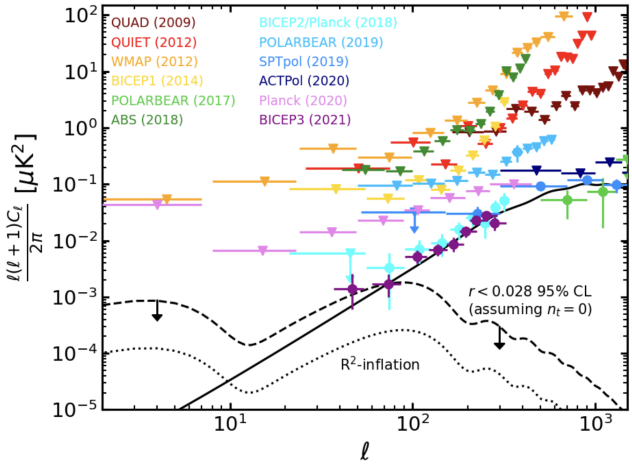


FIGURE 1.1: Adapted from [23], shows CMB B-modes measurements by several experiments. The dashed and dotted lines correspond to the theoretical predictions of the primordial signal for r, while the solid line represents the lensing signal.

The current most stringent upper bound on r from the Planck satellite observations, combined with data from BICEP2/Keck, shown in Figure 1.1, is [23, 24]:

$$r < 0.028 \text{ (95\% CL)} \tag{1.11}$$

which means that the amplitude of tensors cannot be greater than 0.028 the scalars. This is why we need next-generation CMB experiments. Future CMB experiments, such as the Simons Observatory, aim to achieve a sensitivity of $\sigma(r) < 0.003$ [25], while LiteBIRD aims to reach a sensitivity level of 10^{-3} [26].

1.3 Cosmic Microwave Background

The discovery of the CMB radiation by Arno Penzias and Robert Wilson [27] in 1965 provided strong evidence for the Big Bang theory origin of the universe [28]. The CMB spectrum can be described to very high precision by a black body with temperature $T = 2.725\,K$. However, it exhibits small deviations from isotropy at the order of $\Delta T/T \lesssim 10^{-5}$. These are also known as anisotropies and can be mapped through sensitive detectors. Space-based experiments such as COBE [29], WMAP [7] and Planck [30] have been used to measure these temperature inhomogeneities. The measurement of these anisotropies has been extended

to smaller angular scales by ground-based experiments such as the Atacama Cosmology Telescope (ACT) [31] and the South Pole Telescope (SPT) [32] through much improved angular resolution.

The CMB spectrum as measured by COBE is shown in Figure 1.2 and, as anticipated above, is perfectly described by a blackbody emission:

$$I_v(\hat{n}) = \frac{2hv^3}{c^2}\frac{1}{\exp\left(\frac{hv}{kT}\right)-1} = 2\frac{(kT)^3}{(hc)^2}\left[\frac{x^3}{e^x-1}\right], \tag{1.12}$$

where $x = \frac{hv}{kT}$, $\hat{n}$ is a direction in the sky, and T is the associated black-body temperature.

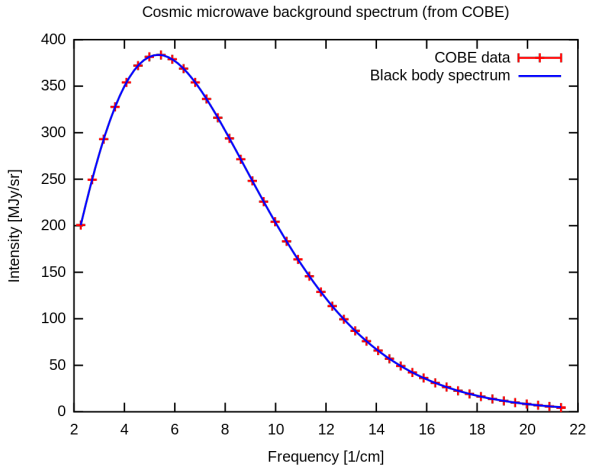


FIGURE 1.2: CMB spectrum measured by the FIRAS instrument onboard the COBE satellite. The theoretical model of the black-body spectrum with temperature 2.725 K is represented by the blue curve, while the red bars represent the negligible uncertainty in the measured intensity values.

1.3.1 Temperature Anisotropies

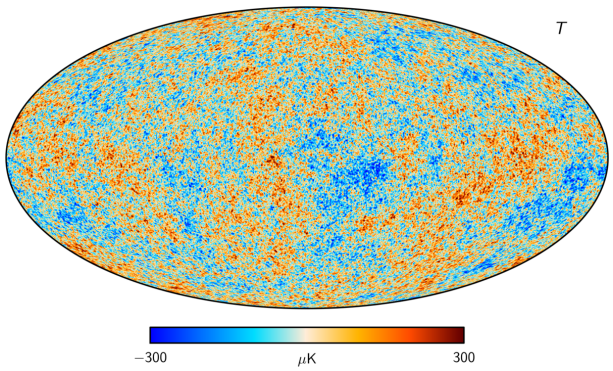


FIGURE 1.3: Full sky CMB temperature anisotropies mapped by the Planck satellite

CMB temperature anisotropies δT can be expanded into a spherical harmonic basis:

$$\delta T(\varphi, \vartheta)/T_o = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\varphi, \vartheta), \tag{1.13}$$

where, $a_{\ell m}$ are the coefficients of the expansion, (φ, ϑ) is the direction on the sky and T_o is the averaged temperature given as:

$$T_0 = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} T(\varphi, \vartheta) \sin\vartheta d\vartheta d\varphi. \tag{1.14}$$

The spherical harmonics are defined as:

$$Y_{\ell m}(\varphi, \vartheta) = \sqrt{\frac{(2\ell+1)}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\vartheta) e^{im\varphi}, \tag{1.15}$$

where the P_ℓ^m are the associated Legendre polynomials,

$$P_\ell^m(x) = \frac{(-1)^m}{2^\ell \ell!} \left(1 - x^2\right)^{m/2} \frac{d^m}{dx^m} P_\ell(x) \tag{1.16}$$

and the $P_\ell(x)$ are the Legendre polynomials,

$$P_\ell(x) = \frac{1}{2^\ell \ell!} \frac{d^\ell}{dx^\ell} \left(\left(x^2 - 1\right)^\ell\right). \tag{1.17}$$

The dipole term in the CMB anisotropies was measured around 1970 and was found to be $\lesssim 10^{-3}$ [33]. When analysing CMB anisotropies, the dipole term is usually subtracted and the following terms are considered:

$$\delta^T(\hat{n}) = \frac{T(\hat{n})}{T_0} - 1 - \sum_{m=-1}^{1} a_1^m Y_{1m}(\vartheta, \varphi) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} a_\ell^m Y_{\ell m}(\vartheta, \varphi) \tag{1.18}$$

where $\hat{n}$ also identified by the angles (ϑ, φ) is the direction along the line of sight. These fluctuations are observed to be $\left|\delta^T(\hat{n})\right| \lesssim 10^{-5}$. Such fluctuations are statistically isotropic and Gaussian according to the inflationary scenario, and can be completely described by the two-point autocorrelation function between $\hat{n}$ and $\hat{n}'$ on the sky

$$C^T(\vartheta) = \left\langle \delta^T(\hat{n}) \delta^T\left(\hat{n}'\right) \right\rangle_{\hat{n}\cdot\hat{n}'=\cos\vartheta}, \tag{1.19}$$

where ϑ is the angular separation between $\hat{n}$ and $\hat{n}'$. Using eq. (1.13), the eq. (1.19) can be written in terms of spherical harmonics as:

$$C^T(\vartheta) = \left\langle \delta^T(\hat{n}) \delta^T\left(\hat{n}'\right) \right\rangle = \sum_{\ell,\ell'} \sum_{m,m'} \langle a_{\ell m} a^*_{\ell' m'} \rangle Y_{\ell m}\left(\hat{n}\right) \cdot Y^*_{\ell' m'}\left(\hat{n}'\right). \tag{1.20}$$

The above expansion in terms of Legendre polynomials reads:

$$C^T(\vartheta) = \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} C_\ell^T P_\ell(\cos\vartheta). \tag{1.21}$$

The above equation can be solved for the coefficients C_ℓ^T using the orthogonality of the Legendre polynomials with the scalar product:

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx \tag{1.22}$$

which gives us:

$$C_\ell^T = 2\pi \int_0^\pi C^T(\vartheta) P_\ell(\cos\vartheta)\sin\vartheta d\vartheta. \tag{1.23}$$

The dimensionless coefficients C_ℓ^T represent the *Angular Power Spectrum* of the CMB temperature anisotropies shown in Figure. 1.4. The peaks in the CMB power spectrum are the imprints left by the acoustic oscillations in the primordial plasma which refer to sound waves that propagated through the hot, dense mixture of photons, electrons, and baryons in the early universe. These oscillations occurred before the universe cooled enough for atoms to form, during the period known as the photon-baryon fluid era. Largest scales anisotropies (low multipoles) are the direct imprint of primordial perturbations unaffected by acoustic oscillations. The

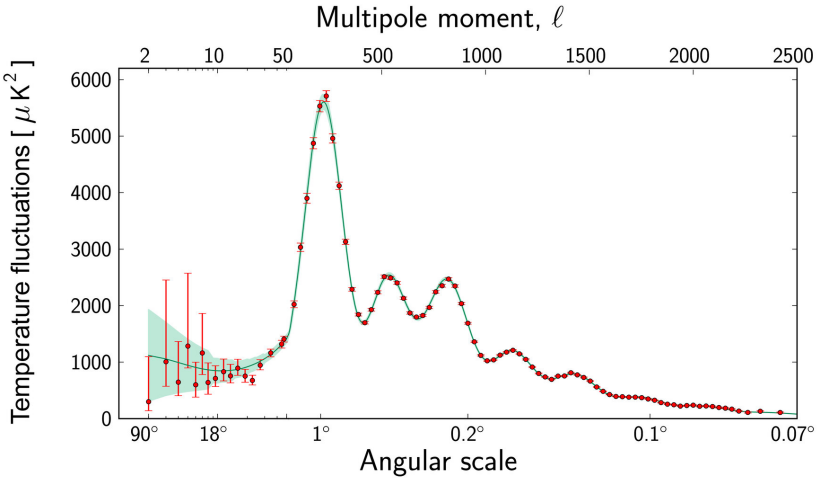


FIGURE 1.4: Angular power spectrum of the CMB temperature anisotropies represented as $D_\ell \equiv \ell(\ell+1)C_\ell/(2\pi)$ in μK^2 units. The green curve is the ΛCDM theoretical model, while the red points are the measurements taken by the Planck mission. The error bars for the largest angular scales reflect both the uncertainties in measurements and the impact of cosmic variance. The angular scale θ is connected to the multipole moment as $\ell = 180°/\theta$.

C_ℓ power spectrum also represents the variance of the harmonic coefficients $a_{\ell m}$,

which, due to statistical isotropy, is independent of m:

$$C_\ell \equiv \langle a_{\ell m} a^*_{\ell m} \rangle \,. \tag{1.24}$$

In the above equation, <> represents the ensemble average typically taken over independent realizations but since we only have one observable universe the only option is to estimate the observed angular power spectrum coefficients by averaging, for every multipole ℓ, over the allowed orientations m of the observed harmonic coefficients $a^{obs}_{\ell m}$. Therefore, the observed angular power spectrum is estimated as:

$$C^{obs}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{l} \left| a^{obs}_{\ell m} \right|^2 . \tag{1.25}$$

However, this estimate is subject to statistical uncertainty, called cosmic variance, originating from the fact that we can only observe one CMB sky realization. Cosmic variance sets a fundamental limit on the precision with which we can measure the true power spectrum C_ℓ, as, even with perfect measurements, the finite number of m modes available for each ℓ introduces an unavoidable uncertainty [34]:

$$\frac{\Delta C_\ell}{C_\ell} \approx \sqrt{\frac{2}{2\ell+1}}. \tag{1.26}$$

1.3.2 CMB Polarization

CMB is linearly polarized due to the Thomson scattering of the photons off free electrons in the primordial plasma. This linear polarization can be expressed in terms of Stokes parameters Q and U. The components of an electric field vector for a monochromatic wave with frequency ω propagating along the $\hat{z}$ direction are given by:

$$E_x = a_x \cos\left(\omega t - \xi_x\right); \quad E_y = a_y \cos\left(\omega t - \xi_y\right) \tag{1.27}$$

.

- E_x and E_y are the components of the electric field in the x and y directions, respectively,
- a_x and a_y are the amplitudes of the electric field components,
- ω is the angular frequency of the wave,
- ξ_x and ξ_y are the phase shifts of the electric field components in the x and y directions, respectively.

The intensity is then [35]:

$$I = a_x^2 + a_y^2 \tag{1.28}$$

and linear-polarization parameters Q and U are:

$$Q = a_x^2 - a_y^2, \quad U = 2 a_x a_y \cos\left(\xi_x - \xi_y\right), \tag{1.29}$$

where Q quantifies the polarization along x or y axes while U along the axes rotated by 45° with respect to the xy reference frame. The circular-polarization parameter is expected to be null for CMB:

$$V = 2a_x a_y \sin\left(\xi_x - \xi_y\right) = 0\,. \tag{1.30}$$

Once the $Q(\hat{n})$ and $U(\hat{n})$ are measured as a function of the position on the sky $\hat{n}$ on a flat region, the spin-2 polarization tensor field can be expressed as

$$P = \frac{1}{2}\begin{pmatrix} Q(\hat{n}) & U(\hat{n}) \\ U(\hat{n}) & -Q(\hat{n}) \end{pmatrix}. \tag{1.31}$$

The CMB polarization field can also be written as a complex number,

$$P_{\pm}(\hat{n}) = |P_{\pm}(\hat{n})|e^{2i\alpha} = Q(\hat{n}) \pm iU(\hat{n}) \tag{1.32}$$

where the magnitude of the polarization is $|P| = \left(Q^2 + U^2\right)^{1/2}$ and the polarization angle is $\alpha = (1/2)\tan^{-1}(U/Q)$, i.e. its orientation relative to the x axis. This polarization field can be expanded into spin-weighted spherical harmonics, $Y^{\pm 2}_{\ell m}$, as:

$$P_{\pm}(\hat{n}) = Q(\hat{n}) \pm iU(\hat{n}) = \sum_{\ell m} a_{\pm 2,\ell m}\, Y^{\pm 2}_{\ell m}(\hat{n})\,. \tag{1.33}$$

Q and U are not rotationally invariant, since $Q \to Q$ and $U \to -U$ under parity inversion. Therefore, one cannot write a two-point correlation function for the polarization Stokes parameters as in the case of temperature anisotropies because it will depend on the direction of the line connecting the two points making it difficult to extract physical information. See [36] for further details.

Thus, we introduce E-modes and B-modes, which are polarization spin-0 quantities, and tracing the geometrical properties of the underlying fluctuations as:

$$E(\hat{n}) = \sum_{\ell=2}^{\infty} a^E_{\ell,m} Y_{\ell m}(\hat{n}) \tag{1.34}$$

and similarly $B(\hat{n})$ mode,

$$B(\hat{n}) = \sum_{\ell=2}^{\infty} a^B_{\ell,m} Y_{\ell m}(\hat{n}) \tag{1.35}$$

where,

$$a^E_{\ell m} = \left(a_{+2,\ell m} + a_{-2,\ell m}\right)/2; \quad a^B_{\ell m} = \left(a_{+2,\ell m} - a_{-2,\ell m}\right)/2i\,. \tag{1.36}$$

Where $a_{\pm,\ell m}$ are those in eq. (1.33). Therefore, the following six quantities fully characterize the statistical properties of CMB fluctuations in harmonic space:

C^{TT}_ℓ for temperature;

C^{EE}_ℓ for E-modes;

C^{BB}_ℓ for B-modes;

C_ℓ^{TB}, C_ℓ^{TE} and C_ℓ^{EB} for their cross-correlations.

Their respective power spectra are defined starting from their spherical harmonic coefficients as:

$$\left\langle \left(a_{\ell m}^{X}\right)^{*} a_{\ell' m'}^{X'} \right\rangle = C_\ell^{XX'} \delta_{\ell\ell'} \delta_{mm'} \tag{1.37}$$

where $X, X' = \{T, E, B\}$. Under parity inversion $E \to E$, while $B \to -B$. Since the standard model of cosmology is parity invariant, it follows that $C_\ell^{TB} = 0$ and $C_\ell^{EB} = 0$ for CMB and we only need to determine the remaining four quantities shown in Figure. 1.5. The temperature and E-modes are supposed to be generated by scalar primordial perturbations, while primordial B-modes are by tensors.

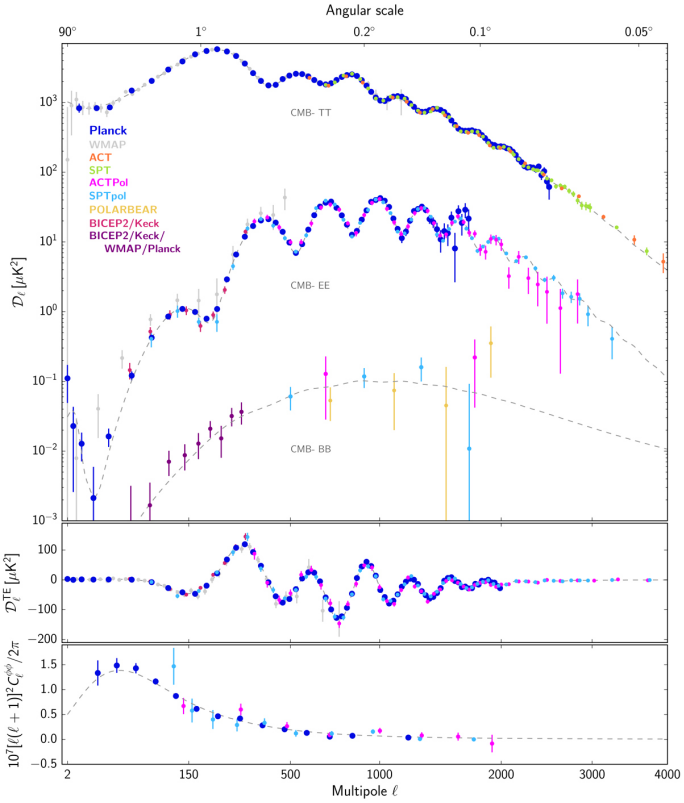


FIGURE 1.5: Comparison of theoretical predictions for the ΛCDM model with state-of-the-art observational data [37]. CMB angular power spectra for D_ℓ^{TT}, D_ℓ^{EE}, D_ℓ^{BB} (upper panel), D_ℓ^{TE} (middle panel), and the lensing $C_\ell^{\phi\phi}$.

CMB is only around 10% polarized; therefore, the temperature power spectra are several orders of magnitude greater than E-modes. Nevertheless, several experiments have detected E-modes, such as WMAP [38], BOOMERANG [39] and Planck collaboration [5]. Primordial B-modes, instead, because of their extremely faint nature, are yet to be detected and targeted by several upcoming experiments, introduced in section 1.2.1.

1.3.3 Origin of B-modes

B-modes in the CMB polarization arise from several astrophysical and cosmological phenomena. Below are the primary mechanisms responsible for generating these distinctive polarization patterns:

- The *primordial B-modes* are generated by tensor cosmological perturbations in the early universe. According to the Inflationary theory, the universe expanded exponentially immediately after the Big Bang. This process is expected to produce a stochastic background of gravitational waves at large angular scales (low multipole $\ell \lesssim 100$), which would have left a specific imprint on the polarization pattern of the CMB. The amplitude of these modes is parametrized by the tensor-to-scalar ratio and it is directly linked to the energy scale of inflation. The polarization spectra of the CMB are modified as the CMB photons interact during their journey from the last scattering surface. *Reionization*, in particular, reintroduces power in large-scale polarization by scattering CMB photons in the late universe. This creates a distinct feature called the reionisation bump, which appears as a peak in the CMB B-mode spectrum at ($\ell < 20$) values. During *Recombination* era primordial B-modes are produced due to the interaction of primordial gravitational waves with the photon-baryon plasma. These gravitational waves induce quadrupolar anisotropies in the temperature of the CMB, which, through Thomson scattering, generate a distinct pattern of polarization leading to peaks in the power spectrum at angular scales corresponding to the horizon size at the time of recombination.

- *B*-mode polarization can also originate from the *gravitational lensing* of the CMB due to the intervening large-scale structures, between the last scattering surface and us. Lensing scatters photons off their original paths, slightly distorting the original pattern of the CMB anisotropies. In polarization, the primary effect is the conversion of high power *E-mode* to lower power *B-mode* [40], making it challenging to distinguish from primordial weak *B-mode*. This phenomenon is known as 'BB lensing' [41]. BB lensing has been measured by Planck with high accuracy in harmonic spectra [42] and map domains [43], as well as by ground-based observatories POLARBEAR [44], SPTpol [45], and ACTPol [46].

- A B-mode pattern can also be generated by a phenomenon known as "Cosmic birefringence" [47]. This occurs when the orientation of the linear CMB polarization plane rotates as the photons travel through space. This rotational effect is caused by the interaction with new fields or particles such as axions, leading to the conversion of E-modes into B-modes.

- At first order, primordial B-mode polarization is only caused by the tensor perturbations of the metric. However, some non-linear effects also induce B-mode polarization, even in the absence of gravitational waves. This is addressed with second-order perturbation theory. As discovered by [48], this intrinsic scalar B-polarization is unlikely to interfere with future searches for the primordial tensor B-modes.

- *Primordial magnetic fields* can also contribute to CMB anisotropies through metric perturbations and the Lorentz force felt by baryons in the pre-recombination plasma [49]. In particular, resulting rotational velocity perturbations on the last scattering surface of the CMB can lead to B-mode polarization patterns, they contribute to polarization anisotropies of the order of $0.1\mu K - 4\mu K$ on small angular scales $500 < \ell < 2000$ [50].

1.3.4 Future CMB experiments for detection of B-modes

Primordial B-modes offer one of the best observational windows to provide direct evidence for cosmic inflation and to constrain inflationary models. However, due to the lack of detection so far, improving detector sensitivities and data analysis techniques is the main challenge towards measuring CMB polarization B-modes to constrain r. Experiments such as BICEP2, Keck Array, and Planck have made significant strides and have provided detailed polarization data. The Planck satellite data combined with BICEP2 provides the tightest constraints to date on the largest angular scale B-modes, which cannot be observed from the ground [51]. There are several CMB experiments in planning and ongoing with the scientific goal of detecting the inflationary B-mode signal, with different frequency ranges, angular scales, and observational time.

- **Simons Observatory:** It is a ground-based mission being built in Chile, which started observations in early 2024. It aims to measure temperature and polarization anisotropies. The observatory will include multiple telescopes designed to observe the CMB with high sensitivity and resolution covering a frequency range from 30 GHz to 270 GHz. The small aperture telescopes (SATs) will target the largest angular scales observable mapping a sky fraction of $f_{sky} = 30 - 40\%$, to measure the primordial tensor-to-scalar ratio r at a target level of $r \approx 0.003$. While large aperture telescopes (LATs) will provide high angular resolution data covering 40% of the sky [20].

- **CMB-S4:** CMB-S4 is a next-generation ground-based experiment designed to measure the CMB temperature and polarization at high sensitivity and high resolution, with a sky coverage of about 40%, reaching a sensitivity of $r < 0.001$, in the presence of Galactic foregrounds and gravitational lensing of the CMB [52].

- **LiteBIRD** (Light satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection): A Japanese-led mission set to launch in the early 2030s. It aims to map CMB polarization at large angular scales. LiteBIRD will observe the CMB sky in multiple frequency bands (from 40 GHz to 400 GHz) with the satellite designed to achieve a sensitivity sufficient to detect or constrain the tensor-to-scalar ratio $r \approx 0.001$ or further [26].

- **PICO** (Probe of Inflation and Cosmic Origins): A NASA mission study-concept. It is designed to observe through 21 frequency bands ranging from 20 to 800 GHz aiming to detect or constrain $r \lesssim 10^{-4}$ [53].

CMB measurements are contaminated by the emission in the microwave range (typically between 30 and 300 GHz) from both the Galactic interstellar medium (ISM) and extra-galactic sources, known as *foregrounds*, which makes it difficult to extract cosmological information from the observed frequency maps. These foregrounds are expected to emit polarized light, with a polarization fraction that is often equal to or greater than the CMB, hence making separating these components from the CMB signal essential. This is particularly important for the B-modes of CMB polarization, which, if measurable, will be sub-dominant at every scale and frequency. In the next chapter, we will explore the properties of these foreground contaminants.

Chapter 2

CMB Foregrounds

Several processes contribute to the total sky emission in the frequency range of interest for CMB observations, typically between 30 and 300 GHz. Foreground emission comes from the solar system, extragalactic sources, the galactic interstellar medium (ISM), commonly known as diffuse Galactic emission, and from distortions of the CMB itself through its interaction with structures in the nearby universe. These foregrounds dominate throughout much of the sky and across most frequencies, particularly in polarization, but they are less dominant in temperature near the CMB peak frequency from 70 to 120 GHz [54]. Galactic foreground sources include:

- **Synchrotron Emission:** Radiation emitted by electrons spiralling along the Galactic magnetic field dominates at low frequencies $\nu \lesssim 70\,\mathrm{GHz}$ and has large-scale smooth structures.
- **Thermal Dust:** Radiation emitted by the interstellar dust grains getting heated by the star-light dominating at frequencies from around 90 GHz.
- **Free-Free Emission:** Emission from the electrons scattering off ions in the ionized regions of the Milky Way. This process contributes at frequencies below 100 GHz. This emission has a flatter spectrum than synchrotron.
- **Anomalous Microwave Emission (AME):** Emission from the spinning dust grains at microwave frequencies with a peak around 20-40 GHz. The spectral distribution is different from thermal dust and synchrotron [55].

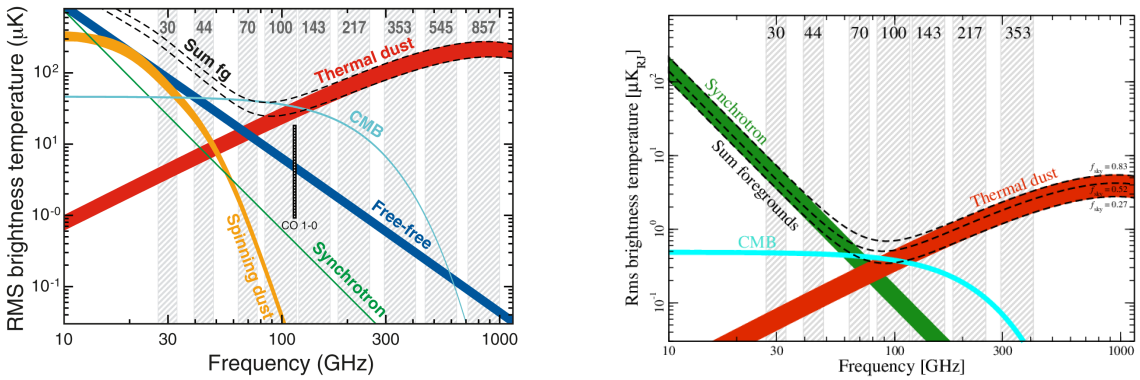
Their expected SEDs at microwave frequencies are shown in Figure 2.1a. Among these processes, thermal emission from galactic dust and synchrotron radiation are significantly linearly polarized (see section 2.1 and 2.2), while free-free emission is generally not polarized because of the isotropic and random scattering of electrons. Anomalous microwave emission is found to be very weakly polarized with current upper bounds on polarization fraction of 1% [56]. Therefore, the focus of the present thesis will be on the removal of the first two contaminants in CMB polarization. Figure 2.2 shows full-sky polarization maps observed by Planck at different frequencies contaminated by foreground emission, where Galactic contamination is visible at all frequencies with synchrotron dominating for $\nu <$ 70 GHz and thermal dust from around 90 GHz. Galactic emission also includes contributions from compact regions such as supernovae remnants and molecular clouds with specific emission properties. The intensity of these emission processes

varies across the sky, but the relative amplitudes are fairly typical. In Figure 2.1b polarized intensity rms (root mean square) amplitude of synchrotron and thermal dust emissions as a function of frequency are compared with the CMB rms for the ΛCDM model, as observed by Planck. This comparison highlights that the CMB signal is sub-dominant at all frequency channels of the Planck mission[1] with the contamination being minimal around $\nu \approx 100$ GHz. Thus, modelling synchrotron and dust emissions is important to determine the level of contamination to CMB. This also enables the magnetic field to be traced over the entire Galaxy.

Extra-galactic objects emit via several different mechanisms, each having its own spectral energy distribution and polarization properties.

- **Sunyaev-Zel'dovich Effect:** CMB photons interact with hot electrons in galaxy clusters, resulting in a distinctive spectral distortion, which is null at around 217 GHz [57].
- **Infrared Point Sources:** Infrared radiation at high frequencies above 100 GHz from distant galaxies including star-forming galaxies.
- **Radio Point Sources:** Emission from the radio wave emitting galaxies, AGN and star-forming galaxies contribute over a broad range of frequencies.

The Sunyaev-Zel'dovich effect generally lacks polarization, while infrared point sources can show polarization, especially because of magnetic field arrangements. Radio point sources frequently display polarization, mainly resulting from synchrotron radiation emitted by relativistic jets and their associated magnetic fields.



(A) (B)

FIGURE 2.1: CMB and Galactic foregrounds spectral energy distribution in temperature (A) and polarization (B). In both figures, grey bands correspond to the Planck frequency channels [58].

[1] From 2009 to 2013, the Planck satellite conducted a comprehensive survey of the entire sky in nine frequency bands ranging from 30 GHz to 857 GHz, providing the best maps to date of the CMB emission. 353 GHz is the highest-frequency polarization channel.

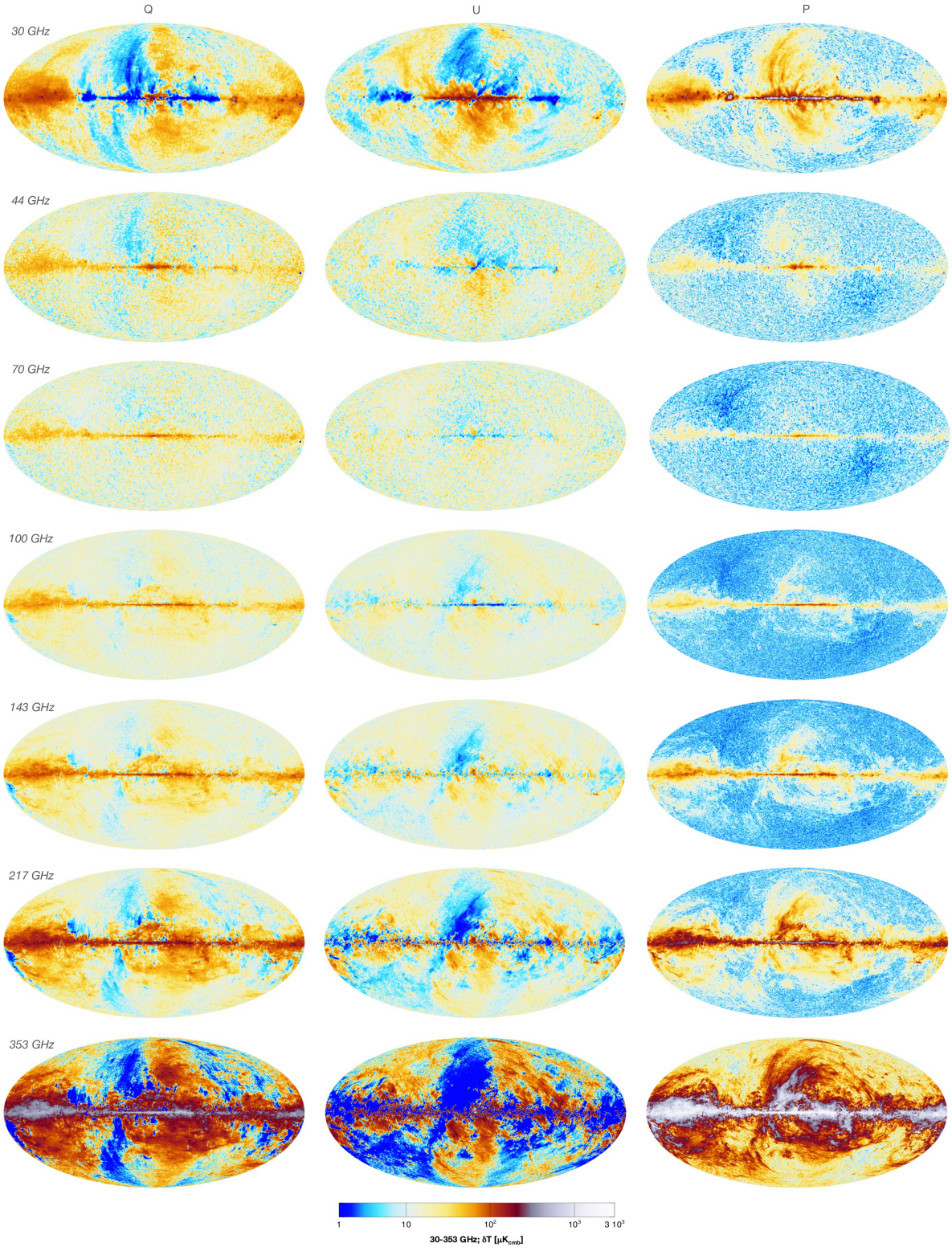


FIGURE 2.2: Full sky polarization maps are shown for seven Planck frequency channels. Polarization intensity P is defined as $P = \sqrt{Q^2 + U^2}$ shown in the last column.

2.1 Synchrotron Emission

Polarized synchrotron radiation is generated by Cosmic Ray (CR) electrons as they accelerate along the galactic magnetic field. As these electrons spiral around the magnetic field lines, they emit highly anisotropic radiation, which is polarized perpendicular to the direction of acceleration and the magnetic field's direction. This emission is naturally partially linearly polarized up to 70-75% [35].

Synchrotron emission is related to the CR energy density distribution, which follows a power law with an electron-energy distribution index p: $n(E) \propto E^{-p}$. It also depends on the strength of the magnetic field, $\mathbf{B}$, which can vary across the Galaxy, resulting in different spectral behaviour for the synchrotron emission. The synchrotron spectrum can be approximated by a power law over a wide range of frequencies as:

$$I_s(\hat{n}, \nu) = A_s(\hat{n}) \cdot \left(\frac{\nu}{\nu_s}\right)^{\beta_s}, \tag{2.1}$$

where β_s is known as the *spectral index*. For synchrotron radiation, this spectral index is typically constant over a large range of frequencies [35]. $A_s(\hat{n}) \propto B(\hat{n})^{(p+1)/2}$ is the synchrotron intensity measured at frequency ν_s in the $\hat{n}$ direction and spectral index is related to p:

$$\beta_s = -\frac{(p+3)}{2}. \tag{2.2}$$

It typically has a value of -2.7 at 22 MHz. The spectral index derived by [7] using the WMAP data shows a steepening of the spectral index around 20 GHz with $\beta_s = -3$, which is consistent with the spectral index $\beta_s = -3 \pm 0.06$ derived by [59], a re-analysis of the WMAP data using a different technique. Synchrotron polarization fraction for electron density following a power law of index p is given as [35]:

$$\Pi = \frac{3(p+1)}{3p+7}. \tag{2.3}$$

Current measurements show variations in the observed polarization fraction from about 10% near the galactic plane, to 10-40% at intermediate to high galactic latitudes [60, 61]. The degree of polarization changes slowly with small variations in p. So, the intrinsic synchrotron polarization level should be almost constant across the entire sky. However, geometric depolarization occurs because of changes in the polarization angle along the line of sight, causing partial cancellation of polarization when emissions with orthogonal polarization directions overlap. Synchrotron SED can be parameterized as a simple power law in brightness temperature [62]. However, variations in the energy distribution of electrons can cause the SED to deviate from a straight power law, resulting in a curved shape. Thus, a general model for synchrotron emission can be written to also take into account the spatial variations of the spectral index with a curvature parameter C_s:

$$I_s(\hat{n}, \nu) = A_s(\hat{n}) \cdot \left(\frac{\nu}{\nu_s}\right)^{\beta_s + C_s \log \frac{\nu}{\nu_{C_s}}}. \tag{2.4}$$

A default synchrotron emission intensity template at 408 MHz has been provided

by Haslam et al. (1982) [63], who combined sky surveys at 1.4 GHz [64], 2.3 GHz [65], and other studies [66, 67] to derive nearly full-sky spectral index maps. Figure 2.3 shows a polarized intensity emission of synchrotron radiations at 30 GHz. This map shows the strong intensity of the synchrotron in the Galactic plane, as well as the North Galactic spurs, as observed by Planck [58].

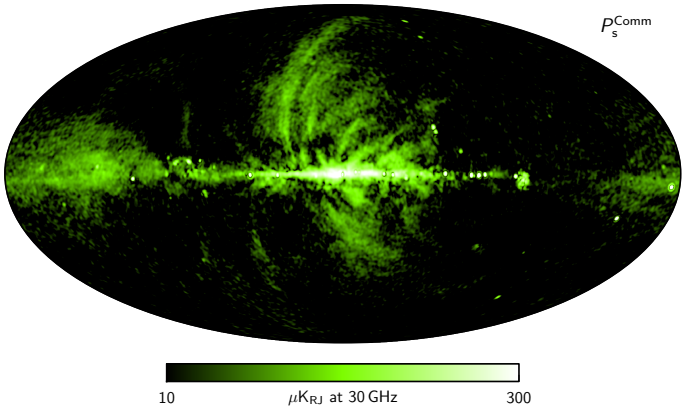


FIGURE 2.3: Synchrotron polarization intensity map observed by Planck at a frequency of 30 GHz with a resolution of 40′ FWHM.

Analysis of data from (All Sky Survey) S-PASS and WMAP shows that the spectral index of polarized synchrotron emissions, β_s, changes with Galactic latitude, becoming steeper from around 2.8 at low latitudes to 3.3 at high latitudes [68]. This change is likely due to depolarization effects at lower latitudes. Furthermore, there is a slight flattening of the spectral index at higher frequencies, though this comes with significant uncertainty. These spatial variations highlight the need for accurate modelling of synchrotron foregrounds in CMB analysis.

2.2 Thermal Dust

The ISM dust grains emit polarized thermal radiation at far-infrared and sub-millimetre wavelengths when they are heated by absorbing interstellar radiations. Dust grains are composed of graphites, silicates, and polyaromatic hydrocarbons (PAH), and they have their dipole moment which gets oriented along the direction of the local galactic magnetic field. Therefore, the dust grains' longest axis becomes aligned at right angles to the magnetic field lines, resulting in radiation being predominantly emitted along this axis. This alignment happens due to the torques exerted by the magnetic field on the dust grains, causing their longest axis to align perpendicular to the magnetic field lines. We can model the thermal emission from these grains using a modified black body (MBB) with emissivity $\epsilon_\nu = \kappa_\nu B_\nu(T_d)$, where the absorption coefficient κ_ν depends on the properties of the dust grains [69]:

$$I_{\rm d}(\hat{n},\nu) = A_{\rm d}\left(\hat{n},\nu_{\rm d}\right)\cdot\left(\frac{\nu}{\nu_{\rm d}}\right)^{\beta_{\rm d}(\hat{n})}\cdot\frac{B_\nu\left(T_{\rm d}(\hat{n})\right)}{B_{\nu_{\rm d}}\left(T_{\rm d}(\hat{n})\right)}, \tag{2.5}$$

where $B_\nu(T)$ is the black-body spectrum, $\beta_{\rm d}$ is the dust spectral index, $T_{\rm d}$ is the dust temperature, and $A_d\left(\hat{n},\nu_{\rm d}\right)$ represents the amplitude at a reference frequency

$\nu_{\rm d}$ along direction $\hat{n}$. The average temperature is $T_d \approx 19\,K$ and index $\beta_d \approx 1.6$ [43].

The polarization of thermal dust emission is a direct consequence of the alignment of dust grains with the galactic magnetic field, resulting in partial polarization of the emitted radiation. Polarization of starlight by aspherical dust grains suggests a partial alignment of elongated grains with the galactic magnetic field [70]. The alignment level varies depending on the dust grain size. Superposition of multiple magnetic field orientations and dust clouds with different SEDs along the line of sight modify the overall SED of observed dust emission.

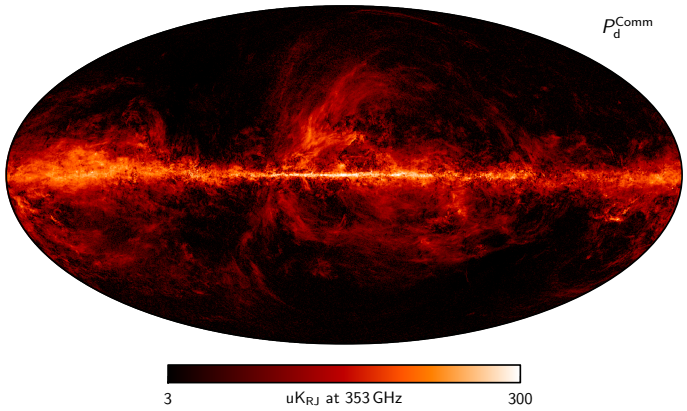


FIGURE 2.4: Dust polarization intensity map observed by Planck at a frequency of 353 GHz with a resolution of 5′.

The Planck satellite measured polarization fractions across the sky, with average values of around 10% at high latitudes, whereas at intermediate latitudes it can reach approximately $\sim 20\%$ [71]. Regions with lower intensity, typically exhibit higher polarization fractions. Conversely, areas with higher density, such as the Galactic plane, often have reduced polarization fractions as a result of line-of-sight depolarization [72]. The distribution of polarized thermal dust emission over the sky is shown in Figure 2.4, as observed by Planck.

The thermal dust emission spectral index, β_d, exhibits spatial variation, affected by factors like local dust grain characteristics and environmental conditions. Existing data has been unable to tightly constrain this variation. According to [73], the dust spectral index can vary with Galactic latitude and dust density, indicating variations in dust temperature and composition. Specifically, β_d tends to increase in regions with higher dust column density or different dust grain sizes, leading to distinct emission characteristics across the sky. These variations underscore the importance of considering spatially dependent spectral indices when modelling foreground emissions for cosmological observations. To account for this variability different dust models have been proposed in the literature, in particular multi-component models with different dust temperatures which are in agreement with current observations.

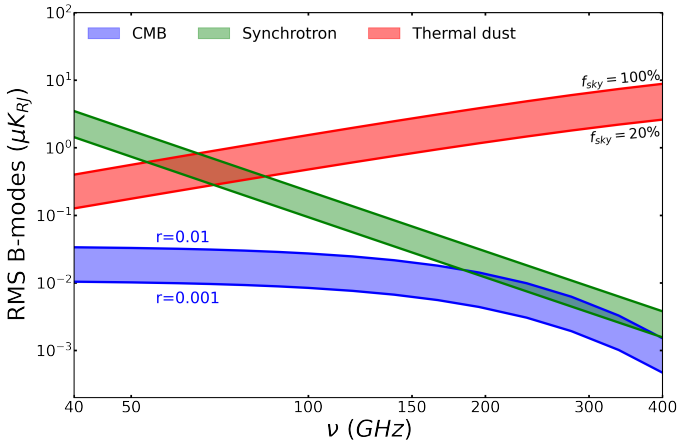


FIGURE 2.5: SED of the synchrotron and thermal dust emission compared with the expected rms amplitude of primordial B-modes. (Figure courtesy of Alessandro Carones.)

The amplitude of the primordial B-mode signal is expected to be less than 1% of the foreground emission, making its extraction quite challenging. The Galactic foregrounds consistently overshadow the signal at all frequencies. Hence, developing methods for foreground subtraction has become increasingly important. Various techniques have been utilized, such as analyzing data at different frequencies and studying the frequency dependence of astrophysical emission laws. In Chapter 4, we will outline some of the component separation methods adopted in this thesis work to isolate the primordial B-mode signal.

Chapter 3

Simulated Data Framework

This chapter describes the methodology and processes used to simulate the maps and datasets that are used for the analysis presented in the thesis. To simulate full-sky Galactic foreground emission, we have utilized a Python package called *PySM3 (Python Sky Model)* [1] [74]. Polarized extragalactic sources are excluded because they are anticipated to be significant at angular scales smaller than those primarily relevant for inflationary B-mode signals ($\ell > 100$) [75].

3.1 CMB and Foreground Maps

Input CMB

In the analysis conducted for this thesis work, we have used the theoretical power spectrum obtained assuming the best-fit ΛCDM cosmological parameters as inferred by Planck's 2018 analysis. The CMB anisotropies are assumed to be Gaussian distributed with the harmonic variance given by the power spectrum $C_\ell = \left\langle \left| a_{\ell m}^{CMB} \right|^2 \right\rangle$. The CMB theoretical power spectrum is computed with *CAMB* [2] with a tensor-to-scalar ratio $r = 0$ (no primordial gravitational waves). Input CMB maps are then generated from the theoretical power spectrum using the `snyfast` function of the *Healpy*[3] package.

Synchrotron Models

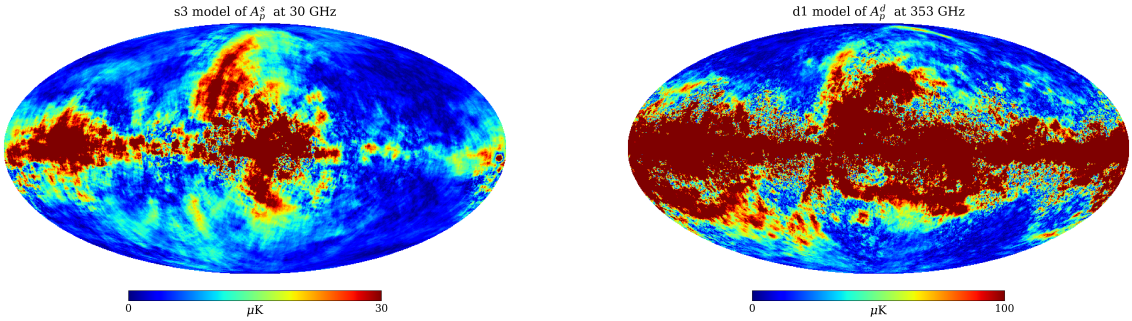
The WMAP map at a reference frequency of 23 GHz is taken as a template of the synchrotron emission [76]. Then, it is extrapolated at other frequencies according to the following models:

1. Model `s0`: It follows a power law with $\beta_s = -3$.
2. Model `s1`: A frequency independent but spatially varying β_s is assumed. The spectral index map was derived by combining the Haslam 408 MHz data with the 7-year WMAP 23 GHz data [77].
3. Model `s3`: This model accounts for spatial variations and the curvature of β_s above a certain frequency ν_c, given as $\beta_s + C_s \ln(\nu/\nu_c)$. Here, β_s is the same

[1] https://github.com/galsci/pysm
[2] Code for Anisotropies in the Microwave Background
[3] https://healpy.readthedocs.io/en/latest/

as the one used in model `s1`, and the curvature parameter C_s has a constant value of 0.052 with $\nu_c = 23\,\text{GHz}$ as found by [78]. Figure 3.1a shows the simulated `s3` polarization intensity map at 30 GHz.

Both the template and spectral index map have an intrinsic resolution of FWHM 5 degrees due to the limiting sensitivity of the experiments from which they have been estimated, and then Gaussian fluctuations are added at smaller scales by extrapolating the power spectrum.



(A) (B)

FIGURE 3.1: Polarization intensity template maps simulated with *PySM3* are shown for synchrotron (left) and dust (right).

Dust Models

The reference dust emission template is the Planck observed sky at 353 GHz and then extrapolated at other frequencies according to the following models:

1. Model `d0`: a MBB with $\beta_d = 1.54$ and $T_d = 20\,K$ constant across the sky.
2. Model `d1`: a MBB with β_d and T_d which vary across the sky. These parameters are derived from the Planck data using the Commander code [79]. Figure 3.1a shows the simulated `d1` polarization intensity map at 353 GHz.
3. Model `d4`: a generalization of model `d1` to multiple dust populations. This option utilizes the two-component model from [80]. In this case, the spectral index maps are assumed constant across the sky.
4. Model `d7`: physical dust model developed by Hensley and Draine in 2017 [81].
5. Model `d12`: 3D polarized dust emission model, where total dust emission is obtained by co-adding emission from different layers of dust, each of them assuming a MBB [82].

In polarization, all the above models are developed to match the 353 GHz Planck data which is dominated by thermal dust.

3.2 Beam Correction

In CMB observations the finite resolution of a telescope impacts the measurement of the CMB's angular power spectrum. The angular resolution of a telescope can be approximated by a Gaussian axisymmetric profile characterized by its full width at half maximum (FWHM). This limited resolution has a significant impact at small angular scales, as features smaller than the telescope's resolution are smoothed out and cannot be accurately measured. This effect is described by the beam transfer function in harmonic space:

$$b_\ell = \exp\left(-\frac{\ell(\ell+1)\sigma_b^2}{2}\right), \tag{3.1}$$

where σ_b represents the beam width related to FWHM by:

$$\sigma_b = \frac{\mathrm{FWHM}}{\sqrt{8\ln 2}}. \tag{3.2}$$

For this work, simulated CMB and foreground signals of each frequency map of a given experiment are convolved with a beam whose FWHM is reported in the tables of section 3.5. Afterwards, all the input maps (including noise, as described in section 3.4) are brought to a common resolution of 70 arcmin. This is required for the component separation analysis we are going to perform (see section 4.2). Then, the observed power spectrum C_ℓ^{obs} is the one convolved with the beam response:

$$C_\ell^{\mathrm{obs}} = b_\ell^2 C_\ell^{\mathrm{true}}. \tag{3.3}$$

Therefore, the true power spectrum can be derived by dividing the observed one by the squared beam transfer function:

$$C_\ell^{\mathrm{true}} = \frac{C_\ell^{\mathrm{obs}}}{b_\ell^2}. \tag{3.4}$$

3.3 Thermodynamic Units

PySM by default generates sky emission maps in Rayleigh-Jeans units. Using RJ units is useful in the microwave range, especially when analyzing the CMB and foreground emissions. However, at higher frequencies where the RJ approximation no longer applies, it is necessary to convert RJ units into thermodynamic temperature (CMB) units before implementing separation methods to accurately interpret the data within the context of the full Planck satellite frequency coverage, as CMB spectrum is flat in those units. In the RJ limit, where $h\nu \ll k_B T$, the spectral radiance I can be approximated as:

$$I \approx \frac{2\nu^2 k_B T}{c^2}, \tag{3.5}$$

we invert this formula to find the RJ temperature T_{RJ} from the spectral radiance:

$$T_{RJ} = \frac{c^2 I}{2\nu^2 k_B}. \tag{3.6}$$

For small deviations from the average temperature of the CMB, of the CMB, the relationship between changes in the CMB thermodynamic temperature and changes in the RJ temperature is given by:

$$dT_{CMB} = \frac{x^2 e^x}{(e^x - 1)^2} dT_{RJ}, \tag{3.7}$$

where $x = \frac{h\nu}{k_B T}$.

3.4 Simulated Noise

The analyses presented in this thesis use two types of simulated noise models, to account for the instrumental noise for the relevant experiments. The first type represents the isotropic Gaussian, white noise, uncorrelated among pixels and frequencies. In each pixel, noise is sampled from a Gaussian distribution with dispersion σ_b, which is estimated from the ratio of sensitivity values given in the tables in section 3.5 and the linear size of the pixel in arcmins. The noise maps are then smoothed by an effective kernel according to the following formulae:
1. When $\mathrm{FWHM_{in}} \leq \mathrm{FWHM_{out}}$:

$$b_\ell = \exp\left(-\frac{\ell(\ell+1)}{2} \left(\frac{\sqrt{\mathrm{FWHM_{out}^2} - \mathrm{FWHM_{in}^2}}}{\sqrt{8\ln 2}} \right)^2 \right) \tag{3.8}$$

2. When $\mathrm{FWHM_{in}} > \mathrm{FWHM_{out}}$:

$$b_\ell = \frac{1}{\exp\left(-\frac{\ell(\ell+1)}{2} \left(\frac{\sqrt{\mathrm{FWHM_{in}^2} - \mathrm{FWHM_{out}^2}}}{\sqrt{8\ln 2}} \right)^2 \right)} \tag{3.9}$$

where FWHM_{in} is the resolution of frequency channels given in tables of section 3.5. Multi-frequency maps need to be brought at a common target resolution ($FWHM_{out}$) before being processed by some of the component separation pipelines as those considered in this thesis work. Therefore, also the noise is smoothed accordingly with the procedure described above and then combined with input signal maps.

The second category of simulated maps, incorporates more realistic noise property assumptions [20]. Specifically, these maps include non-white noise components, like $1/f$ noise, which are responsible for the loss of large-scale modes due to filtering required to reduce the impact of atmospheric contamination. The

noise power spectrum N_ℓ is estimated as:

$$N_\ell = N_w(\ell)\left[1+\left(\frac{\ell}{\ell_{knee}}\right)^{\alpha_{knee}}\right], \tag{3.10}$$

where N_w is the baseline white noise and can be derived from the sensitivity values of the tables of section 3.5. α_{knee} and ℓ_{knee} characterise the contribution from $1/f$. α_{knee} ranges from -3.0 to -2.4. ℓ_{knee} is the angular scale above which the contribution from $1/f$ is dominant. The characteristic multipole ℓ_{knee} lies between 15 and 40 for this noise model of the SATs polarization experiment [20]. This model aims to capture key features of the SO noise environment, as large angular scales suffer from contamination by $1/f$ noise at low frequencies in the detector time stream, primarily due to atmospheric and electronic noise.

3.5 Simulated Data-sets

In the present thesis work, we have simulated realistic sky maps using the specifications and anticipated observational capabilities of the SO [20], LiteBIRD [26] and also Planck [37]. We start by testing component separation methods for B-modes on simulated Planck data and then we focus on the ongoing SO experiment. However, we also consider the impact of including in the analysis of this experiment high-frequency data that cannot be mapped from the ground but can only be obtained from space missions, and we consider the addition of data already available from Planck at 353 GHz and future data from LiteBIRD:

- Planck (PL):

Frequency (GHz)	30	44	70	100	143	217	353
FWHM (arcmin)	32.29	27.94	13.08	9.66	7.22	4.90	4.92
Sensitivity (μK arcmin)	210	240	300	118	70.2	105	439

TABLE 3.1: Instrumental characteristics of Planck mission observation channels used for the data analysis

- SO:

Frequency (GHz)	27	39	93	145	225	280
FWHM (arcmin)	91	63	30	13	17	11
Sensitivity (μK arcmin)	49.5	29.7	3.7	4.7	8.9	22.6

TABLE 3.2: Experimental configuration for the Simons Observatory (SO) simulated data set

- SO with the addition of the last two highest frequency channels of LiteBIRD (SL):
- SO with the addition of the Planck's highest frequency channel in polarization (SP):

Frequency (GHz)	27	39	93	145	225	280	337	402
FWHM (arcmin)	91	63	30	13	17	11	20.9	17.9
Sensitivity (μK arcmin)	49.5	29.7	3.7	4.7	8.9	22.6	21.95	47.45

TABLE 3.3: Experimental configuration for the SL

Frequency (GHz)	27	39	93	145	225	280	353
FWHM (arcmin)	91	63	30	13	17	11	4.92
Sensitivity (μK arcmin)	49.5	29.7	3.7	4.7	8.9	22.6	439

TABLE 3.4: Experimental configuration for the SP

Each simulated frequency map is obtained by adding the CMB and foreground maps, smoothed with the corresponding Gaussian beam, with instrumental noise, as described in previous sections. All the maps are simulated using the HEALPix [83] (Hierarchical Equal Area isoLatitude Pixelation)[4] pixelization scheme with a resolution parameter of `nside = 128`, corresponding to a pixel resolution of 27.50 arcmin and a maximum multipole moment `lmax=2×nside`. Since we are targeting primordial B-modes for which the expected signal is confined at $\ell < 100$, this resolution is adequate to capture all the relevant scales.

These simulations provide a realistic representation of future data. Component separation methods, essential for isolating the primordial B-mode signal from other sources with polarized emission are applied to these simulated datasets allowing the testing and refinement of these techniques in preparation for the analysis of actual data from these experiments. The results demonstrate the potential effectiveness of these techniques in identifying the faint B-mode signal amidst the dominant foregrounds, thus contributing to the advancement of methodologies in CMB analysis. In the next chapter, different component separation techniques used in this thesis to separate the CMB signal from the observational data are discussed.

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

Chapter 4

Component Separation Methods

4.1 Introduction

Component separation in CMB data analysis consists of distinguishing the CMB signal from various foreground emissions. This can be challenging because the CMB signal is subdominant compared to the foreground emission, making it difficult to cleanly extract this signal. Nevertheless, this can be done because the CMB radiation has a nearly perfect blackbody spectrum, which means Planck's law can precisely predict its intensity at different frequencies. The non-cosmological signals such as synchrotron and thermal dust emission have non-thermal spectra with varying frequency dependence. Collecting observations at several frequencies makes disentangling the CMB from foregrounds possible by tracing their spectral properties across the sky.

This explains why all CMB experiments observe the sky in multiple frequency bands. For example, the space missions to date have been:

- **COBE-DMR (Cosmic Background Explorer - Differential Microwave Radiometers)**: Utilized three frequency bands between 31 and 90 GHz [84].
- **WMAP (Wilkinson Microwave Anisotropy Probe)**: Designed with five frequency bands between 23 and 94 GHz for improved extraction of the CMB from foreground emission [7].
- **Planck**: Operated with nine frequency bands ranging from 30 to 857 GHz, providing improved component separation and better tracing of dust emission [85].

Various techniques have been used to extract the CMB signal based on different assumptions. Some adopt an empirical modelling of the foreground SEDs across the sky. They are known as parametric methods and include *Commander* [86], *FGBuster* [87] and others [88]. Alternative approaches do not assume models of the foreground emissions. Among these, there is the template fitting [89, 90] and Internal Linear Combination (ILC) [91, 92, 93, 94]. The latter consists of removing foreground contamination by means of statistics and is the method we are going to specifically investigate in this thesis. This approach has been adopted to extract the CMB signal from the multifrequency data for WMAP [7] and Planck [90]. In table 4.1, we have summarized various algorithms and methods used for component separation in CMB data analysis.

Class	Methods/Algorithms	Description
Linear Combination Methods	Internal Linear Combination (ILC)	Combine different frequency maps to minimize output variance and by preserving the full blackbody signal. Used for WMAP and Planck data analysis [89, 90].
	Needlet ILC (NILC)	Apply ILC in the needlet (wavelet) domain to improve subtraction of contaminants locally in pixel space and at different angular scales. Used for Planck data analysis [95].
Template-based methods	Template fitting	Subtract templates of foreground components derived from observed data. Used for WMAP and Planck data analysis [89, 90].
Blind Source Separation	Independent Component Analysis (ICA)	Separate mixed signals into statistically independent components without prior knowledge of their properties. Used for WMAP and Planck data analysis [96, 97].
	FastICA	The independent component analysis that effectively separates independent sources by maximizing non-Gaussianity, making it ideal for separating Gaussian CMB from non-Gaussian foregrounds. Used for WMAP and Planck data analysis [98].
	Spectral Matching Independent Component Analysis (SMICA)	Extract the CMB signal by using maximum likelihood to separate sources in the harmonic domain, combining spectral and spatial information to isolate the CMB from foregrounds. Used for Planck data analysis [99].
Bayesian Methods	Bayesian Component Separation (e.g., Commander)	Use Bayesian inference to model and separate CMB and foregrounds, incorporating prior information. Used for Planck data analysis [100, 90].
Entropy-Based Methods	FastMEM	A harmonic space maximum entropy method that estimates component maps by employing frequency scaling models and external power spectra. Used for earlier CMB experiments and some Planck data [101].

TABLE 4.1: Different component separation algorithms used for CMB data analysis

4.2 Internal Linear Combination (ILC)

ILC is a blind component separation method that does not make any prior assumptions about the frequency dependence or spatial properties of the foregrounds. It combines multifrequency maps linearly using a set of weights to minimize the variance of the final map while maintaining a unit response to the CMB. This technique is easy to implement and computationally inexpensive. The ILC technique is based on two fundamental assumptions:

1. CMB has a black-body spectrum,
2. CMB fluctuations are uncorrelated with the foreground signals.

The observed sky X_i at frequency ν_i is composed of CMB component X_{CMB}, foregrounds ${X_i}^f$, and instrumental noise n_i:

$$X_i = X_{\rm CMB} + X_i^f + n_i\,. \tag{4.1}$$

We are assuming CMB to be frequency independent, this holds only in thermodynamic units. Using ILC, the CMB signal can be re-constructed as:

$$\hat{X}_{\rm ILC} = \sum_{i=1}^{N_\nu} \omega_i \cdot X_i = \sum_{i=1}^{N_\nu} \omega_i \cdot (X_{\rm CMB} + X_i^f + n_i), \tag{4.2}$$

where the summation is over the number of frequency channels N_ν and ω_i is a set of coefficients estimated by minimizing the variance of the final map while retaining the full CMB signal. All input maps must be smoothed to match the beam of the channel with the lowest angular resolution. Maintaining consistent resolution across all input maps in ILC is essential, as the method relies on the assumption of a frequency-independent CMB signal, requiring uniform smoothing and scaling. Since CMB is independent of the frequency in thermodynamics (i.e. CMB) units, we impose that:

$$\sum_{i=1}^{N_\nu} \omega_i = 1 \tag{4.3}$$

which results in:

$$\hat{X}_{\rm ILC} = X_{\rm CMB} + \sum_{i=1}^{N_\nu} \omega_i \cdot (X_i^f + n_i). \tag{4.4}$$

Then, the variance of the final map can be computed by imposing statistical independence of the CMB component from foregrounds and noise:

$$\mathrm{Var}(\hat{X}_{\rm ILC}) = \mathrm{Var}\left(X_{\rm CMB}\right) + \mathrm{Var}\left(\sum_{i=1}^{N_{\nu_i}} w_i (X_i^f + n_i)\right) \tag{4.5}$$

$$\mathrm{Var}(\hat{X}_{\rm ILC}) = \left\langle \hat{X}_{\rm ILC}^2(p) \right\rangle - \left\langle \hat{X}_{\rm ILC}(p) \right\rangle^2 = \boldsymbol{\omega}^T \mathbf{C} \boldsymbol{\omega}, \tag{4.6}$$

where the $<>$ brackets denote the average over all pixels p in the map, $\boldsymbol{\omega}$ is a column vector with all the coefficients ω_i, and the covariance matrix $\mathbf{C}$ can be

expressed as:

$$C_{ij} = \left\langle X_i(p) X_j(p) \right\rangle - \left\langle X_i(p) \right\rangle \left\langle X_j(p) \right\rangle , \tag{4.7}$$

where ij runs over all the frequency channels. To derive the set of coefficients which preserves the CMB signal and minimizes the output variance, we use the method of Lagrange multipliers [92]:

$$\begin{pmatrix} 2\mathbf{C} & -\mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\omega} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \tag{4.8}$$

where λ is the Lagrange multiplier, and $\mathbf{1}$ and $\mathbf{0}$ denote column arrays of ones and zeros, respectively. The coefficients of the linear combination are finally determined as:

$$\omega_i = \frac{\sum_{j=1}^{N_\nu} C_{ij}^{-1}}{\sum_{i,j=1}^{N_\nu} C_{ij}^{-1}} . \tag{4.9}$$

If the properties of the foregrounds vary significantly across the sky due to spatially dependent spectral indexes, the ILC method may not be the optimal approach. In such a case, to improve the performance of the method, the sky can be divided into different patches and a separate set of weights for each patch can be estimated [89]. Afterwards, the total map is constructed by combining all these individual results.

An alternative approach to this minimization is applying it in harmonic space, referred to as HILC [102]. This method operates on spherical harmonic coefficients rather than pixel values, enhancing the effectiveness of foreground removal. However, it does not account for the fact that noise is the dominant source of CMB measurement error at high galactic latitudes, while foreground signals dominate at low galactic latitudes. Conversely, the ILC in pixel space does not consider that noise predominates at high angular frequency (small scales) while foreground emission dominates on large scales. This problem requires a refined version of the ILC introduced in the next section.

4.3 Needlet Internal Linear Combination (NILC)

NILC performs variance minimization locally both in pixel and multipole space. To do so, it employs *needlets*, a specific type of spherical wavelets [103] (discussed in [104, 105]), for component separation in pixel and harmonic space. The multi-frequency maps are filtered using these wavelets and then they are linearly combined to obtain a minimum variance solution at the different needlet scales. Needlets enable localized filtering in both pixel and harmonic space. This means that the linear combination weights can adjust to local conditions of foreground and noise contamination in the pixel domain while prioritizing foreground removal at large angular scales and noise subtraction on smaller ones. The weights can vary smoothly on large scales and rapidly on small scales using needlets. This would not be possible by dividing the sky into patches in the standard ILC pipeline [106].

A set of filters, also known as the needlet harmonic windows $b_j(\ell)$, is defined over a range of multipoles ℓ serving for localization in the harmonic space, such that:

$$\sum_{j=1}^{\infty} \left[b(\ell/B^j) \right]^2 = 1, \tag{4.10}$$

where j specifies the needlet scale and B is a fixed parameter (usually ranging between 1 and 2) which sets the width of the wavelet. Spherical needlets in the real domain can be derived from $b_j(\ell)$[1] as:

$$\Psi_{jk}(\hat{n}) = \sqrt{\lambda_{jk}} \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} b_j(\ell) Y^*_{\ell m}(\hat{n}) Y_{\ell m}(\hat{\xi}_{jk}) \,. \tag{4.11}$$

The set of points $\{\xi_{jk}\}$ represents cubature points on the sphere for the scale j. These points in pixel space correspond to the pixel centres in the HEALPix pixelization scheme. Each index k corresponds to a specific HEALPix pixel at a resolution parameter `nside`, which may vary for each needlet scale j. The weights λ_{jk} are inversely proportional to the number N_j of pixels used for the needlet decomposition, i.e., $\lambda_{jk} = \frac{4\pi}{N_j}$. The input field $X(\hat{n})$ can be decomposed in the Ψ_{jk} basis. The coefficients of this expansion, β_{jk}, are called needlet coefficients:

$$\beta^X_{jk} = \int_{S^2} X(\hat{n}) \Psi_{jk}(\hat{n}) \mathrm{d}\Omega_{\hat{n}} = \sqrt{\lambda_{jk}} \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} b_j(\ell) X_{\ell m} Y_{\ell m}(\hat{\xi}_{jk}) \,. \tag{4.12}$$

The needlet decomposition is linear. Therefore, the needlet coefficients β^X_{jk}, which are derived from the harmonic coefficients $a^X_{\ell m}$, can be expressed as a linear combination of the needlet coefficients of the different components (CMB, foregrounds and noise) at HEALPix pixel centre ξ_{jk}:

$$\beta^X_{jk} = \beta^{\mathrm{CMB}}_{jk} + \beta^{\mathrm{f}}_{jk} + \beta^{\mathrm{n}}_{jk} \,. \tag{4.13}$$

In practice, the needlet coefficients of a map $X(\nu_i)$ at scale j and pixel k are obtained, by applying the needlet filter $b_j(\ell)$ on its harmonic coefficients $a^X_{\ell m}(\nu_i)$:

$$\beta^X_{jk}(\nu_i) = \sum_{\ell,m} a^X_{\ell m}(\nu_i) \cdot b_j(\ell) \cdot Y_{\ell m}(\xi_{jk}) \,. \tag{4.14}$$

The input needlet coefficients maps are then linearly combined in such a way as to obtain a minimum-variance map β^{NILC}_j on each scale j :

$$\beta^{NILC}_{jk} = \sum_i w_{i,jk} \beta^X_{i,jk} = \sum_{\ell,m} a^{NILC}_{\ell m} \cdot Y_{\ell m}(\xi_{jk}), \tag{4.15}$$

$w_{i,jk}$ are the weights of the combined input signal. The NILC method constructs a weighted sum of needlet coefficients at different frequencies ν. The variance of the

[1] $b_j(\ell) = b(\ell/B^j)$

combined map is given by:

$$\mathrm{Var}(\beta_{jk}^{NILC}) = \sum_{\nu,\nu'} w_{\nu,jk} w_{\nu',jk} \mathrm{Cov}(\beta_{\nu,jk}, \beta_{\nu',jk}) \tag{4.16}$$

$$= \mathbf{w}_{jk}^T \mathbf{C}_{jk} \mathbf{w}_{jk} \,. \tag{4.17}$$

The solution to this minimization problem is given by:

$$\mathbf{w}_{jk} = \frac{\mathbf{C}_{jk}^{-1}\mathbf{e}}{\mathbf{e}^T \mathbf{C}_{jk}^{-1}\mathbf{e}} \tag{4.18}$$

with $\mathbf{C}$ being the covariance matrix of multifrequency needlet coefficients and $\mathbf{e}$ a vector of ones in the case of maps in CMB temperature units. The final NILC map is then reconstructed by filtering again the harmonic coefficients $a_{\ell m,j}^{NILC}$ and summing them all for each ℓ and m :

$$X_{NILC}(\hat{n}) = \sum_{\ell,m} a_{\ell m}^{NILC} \cdot Y_{\ell m}(\hat{n}) = \sum_{\ell,m} \left(\sum_j a_{\ell m,j}^{NILC} \cdot b_j(\ell) \right) \cdot Y_{\ell m}(\hat{n}). \tag{4.19}$$

Minimizing the variance (and hence the contamination) separately on different needlet scales leads to more effective cleaning because it allows the method to account for the distinct characteristics of foreground contamination across various angular and spatial scales.

Different kinds of needlets exist in the literature. The *standard* needlets, first introduced in Narcowich et al. (2006) [103], are constructed using filter functions that enable good localization in both spatial and harmonic domains providing compact support in the harmonic space. On the other hand, *mexican* needlets, also known as Mexican hat [107] wavelets, are another specific type of needlet that utilizes a Gaussian window function to achieve better localization properties. They offer improved spatial localization compared to standard needlets. Another type of bases includes *cosine*-shaped needlets [108, 58], where a cosine function modulates the window function to ensure localization in both frequency and pixel space.

4.3.1 NILC Bias

When estimating the covariance matrix within a limited domain, the NILC method can introduce a bias in the reconstructed CMB B-mode map [109]. This bias is a result of spurious empirical correlations between the CMB modes and residual contaminants. These correlations can distort the recovered CMB signal, potentially leading to a negative bias and a loss of power in the reconstructed CMB angular power spectrum, especially on the largest scales.

The physical quantity of interest is the power spectrum of the cleaned map, $C_\ell^{BB,out}$. This is given by:

$$C_\ell^{\mathrm{out}} = C_\ell^{\mathrm{cmb}} + C_\ell^{\mathrm{fgds}} + C_\ell^{\mathrm{noise}} + 2 \cdot C_\ell^{\mathrm{c-f}} + 2 \cdot C_\ell^{\mathrm{c-n}} + 2 \cdot C_\ell^{\mathrm{n-f}}, \tag{4.20}$$

where the first three terms are the angular power spectra of CMB, residuals of foregrounds and noise, while the other terms represent the corresponding correlations among these components. An inaccurate estimate of the covariance switches the correlation terms on. To decrease the statistical uncertainty on large angular scales one can measure the covariance across a sufficiently large region with enough modes to ensure that random fluctuations do not influence the covariance matrix estimation. To do so, multiple Needlet bands are combined at low multipoles:

$$b^{\text{new}}(\ell) = \sqrt{\sum_{j=j_{\text{min}}}^{j_{\text{max}}} b_j^2(\ell)}, \tag{4.21}$$

to sample additional modes. If NILC is correctly implemented, the correlation terms should be very low, and the CMB should be well reconstructed. Then the bias in the final map can be computed as:

$$C_\ell^{\text{bias}} = C_\ell^{\text{out}} - C_\ell^{\text{cmb}} - C_\ell^{\text{fgds}} - C_\ell^{\text{noise}} \tag{4.22}$$

and includes all the cross-correlation terms. This bias should be approximately zero across all angular scales. In the following thesis, we assess the level of bias in our analysis by leveraging a series of simulated skies, each incorporating a distinct CMB and instrumental noise realization. The bias is computed using eq. (4.22) from the averaged NILC outputs. The uncertainty associated with the average of the bias spectra can be obtained by dividing the standard deviation of the bias angular power spectra across different CMB reconstructions by the square root of the number of simulations, given by $\frac{\sigma(C_\ell^{bias})}{\sqrt{N_{sims}}}$ [110].

4.4 Multi-Clustering Approach

As mentioned in Section 4.3, the blind NILC technique is particularly valuable for isolating the weak primordial B-mode signal. However, spatial variation in foreground spectral parameters presents a challenge for NILC in effectively eliminating Galactic contamination. To mitigate this issue, we employ an extension of the NILC technique: the Multi-Clustering NILC (MC-NILC) [111]. This method improves the NILC framework by performing independent variance minimization on different sky regions (clusters or patches) that are specifically chosen to have similar spectral characteristics of B-mode Galactic emission within them.
To implement the MC-NILC pipeline, we need to concentrate on analyzing two aspects:

- Build a blind tracer map to track the spectral properties of the Galactic foreground in B-modes.
- Select an appropriate clustering method to group pixels with similar tracer values into the same patch.

Foregrounds Tracer

As originally proposed in [111], in B-modes, the spectral properties of dust and synchrotron emissions can be blindly traced by computing a ratio of B-mode maps at two separate frequencies. These maps are selected to ensure that the spectral dependencies of dust and synchrotron are fully embraced across the entire frequency range of the experiment. To a first approximation, this ratio can be seen as a map of $(\nu_1/\nu_2)^{b^B_{eff}}$, where ν_1 and ν_2 refer to the selected frequencies and b^B_{eff} denotes the effective spectral index of B-modes foreground emission when modelled using a power-law.

It is preferable to use B-mode maps in the ratio rather than Q and U maps or their corresponding spectral indices, as the spatial distribution of Q and U spectral properties varies significantly from those of B-modes due to the non-local nature of the QU-B transformation.

To assess the effectiveness of the approach, two distinct cases can be considered:

1. *Ideal case*: where a distinct ratio is computed for each needlet scale j. The needlet coefficients are extracted from 'noiseless' foreground simulations at two different frequencies.

2. *Realitic case:* where a single ratio is applied across all needlet scales, constructed from templates (estimated from the input data) of foreground B modes at two different frequencies. These templates are filtered with the first needlet band b_0 to reduce the more significant contamination from the CMB and noise at smaller angular scales.

Clustering Approach

Once a tracer of the B-mode foreground spectral properties is constructed, we proceed by partitioning the sky into different patches that group pixels with similar values of this tracer. This clustering is achieved by applying thresholding to the histogram of the selected ratio.

In [111], the thresholding process is executed through two distinct approaches:

- **Clusters of Equal Area (CEA)**: In this approach, the sky is divided into clusters all with the same area. Specifically, the clusters are defined to include subsets of N_p/K pixels, where N_p represents the total number of pixels, and K is the predetermined number of clusters. These clusters are organized such that they encompass pixels with progressively increasing values of the tracer ratio. This method ensures that each cluster covers an equal portion of the sky, facilitating a balanced comparison across regions.

- **Random Partitions (RP)**: In this method, multiple different partitions of the sky are generated. Each partition consists of clusters with a random number of pixels leading to clusters of varying sizes. The pixels within each cluster are chosen to have the nearest possible values of the tracer, preserving foreground information. Unlike the CEA method, the RP approach does not enforce uniformity in a cluster area, allowing for a diverse set of partitions

that can capture a wider range of variability in the foreground spectral properties.

In this work, we only consider the ideal approach to building a tracer map and utilize partitions of equal area.

4.5 Assessment of Pipeline Performance

To assess how well the component separation techniques perform, we estimate the level of residual contamination from foregrounds and noise separately on the 'cleaned' output CMB B-mode maps.

4.5.1 Residuals

Following [110], for each simulated multifrequency dataset, we estimate the foreground in the cleaned CMB map using eq. (4.18). The weights $\mathbf{w_i}$ of the solution are then used to combine the noise-only and foreground-only muti-frequency simulations to derive the residuals associated to the cleaned map. For each component separation approach, we will study the average power spectrum of residuals across all the simulations.

4.5.2 Likelihood Analysis

After applying component separation, we obtain a cleaned CMB map. To deduce cosmology from this map, assuming the field is Gaussian and undergoing linear evolution (as with the CMB), we can estimate the angular power spectrum of the map and fit a cosmological model to it. For primordial B-modes, the amplitude of the power spectrum depends solely on r, enabling us to constrain r for each simulated map.

In this thesis to do so, we estimate the likelihood of the 'observed' spectrum given a model which quantifies how well the observed binned spectrum matches the model spectrum for a given value of the parameter r. Specifically, we assume a Gaussian likelihood approximation [112]. This approximation introduces some information loss, but we bin the power spectrum (averaging over multiple multipoles) to mitigate this issue. According to the central limit theorem such averaging leads to a Gaussian distribution [113]. The log-likelihood is defined as:

$$\begin{aligned} -2\log\mathcal{L}(r) = \sum_{\ell_b\ell_b'} &\left(C_{\ell_b}^{out} - rC_{\ell_b}^{r=1} - C_{\ell_b}^{\rm lens} - C_{\ell_b}^{\rm noise}\right) \\ &M_{\ell_b\ell_b'}^{-1}\left(C_{\ell_b}^{out} - rC_{\ell_b}^{r=1} - C_{\ell_b}^{\rm lens} - C_{\ell_b}^{\rm noise}\right), \end{aligned} \tag{4.23}$$

where $C_{\ell_b}^{out} = C_{\ell_b}^{lens} + C_{\ell_b}^{fgds} + C_{\ell_b}^{noise}$ is the averaged binned CMB B-mode power spectrum extracted from the output 'cleaned' map. In which we assume the cross-correlation terms between different components to be zero and primordial B-modes are not included. Additionally, in $C_{\ell_b}^{out}$, we assume that the primary

systematic bias to the final angular power spectrum arises from the residual Galactic contamination, as we expect the usage of advanced de-noising techniques such as Monte-Carlo simulations can effectively reduce the noise bias. This justifies eq. (4.23). $C_{\ell_b}^{r=1}$ is the theoretical binned primordial CMB B-mode angular power spectrum for $r = 1$. $C_{\ell_b}^{lens}$ is the theoretical B-mode lensing power spectrum. $M^{-1}_{\ell_b\ell'_b}$ is the inverse covariance matrix for $C_{\ell_b}^{out}$ computed as:

$$M_{\ell_b\ell'_b} = \frac{1}{N-1}\sum_{i=1}^{N}\left(C_{\ell_b}^{(i,\mathrm{out})} - \overline{C_{\ell_b}^{\mathrm{out}}}\right)\left(C_{\ell'_b}^{(i,\mathrm{out})} - \overline{C_{\ell'_b}^{\mathrm{out}}}\right), \tag{4.24}$$

where i runs over N=100 simulations and $\overline{C_{\ell'_b}^{out}}$ is the averaged power spectrum of output cleaned map over 100 simulations. To evaluate the effect of foreground residuals on the estimation of r from C_ℓ^{out}, we fit r only on binned C_ℓ^{fgds} using eq. (4.23). By maximizing the likelihood function $\mathcal{L}(r)$, we can find the most probable value of r. In this work, we adopt a uniform binning with $\Delta\ell = 15$. Such a choice ensures the power spectrum is much more Gaussian distributed due to the central limit theorem.

In the next chapter, the ILC and NILC data analysis pipelines are applied to simulated Planck data.

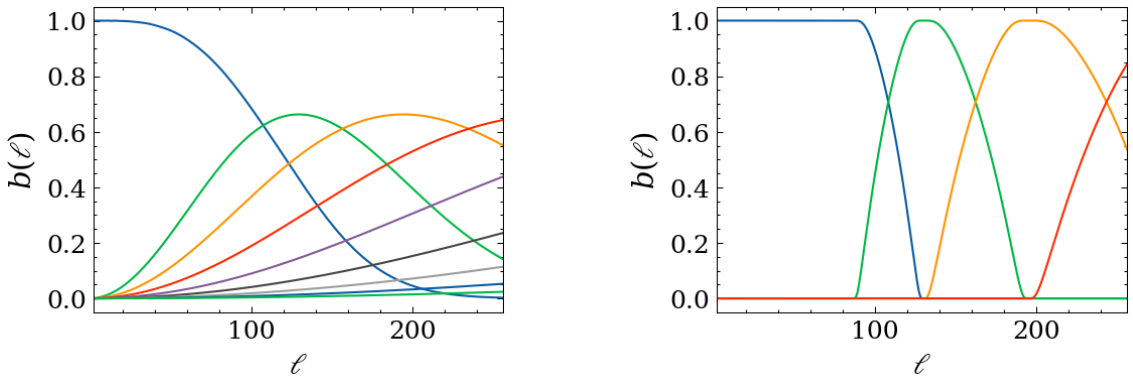
Chapter 5

Planck Data Analysis

In this chapter, a base-level model of ILC in addition to NILC is applied to the simulated multifrequency data set, obtained by combing CMB, noise and foreground signals across the seven Planck polarization frequency channels considering the experimental configuration in table 3.1. We perform the analysis only on the B-modes. Although Planck was not specifically designed for B-mode studies, the goal is to assess the performance of the ILC and NILC techniques, in preparation for future higher-sensitivity measurements.

5.1 Reconstructed CMB B-modes

For the NILC analysis, the adopted configurations for *standard* and *mexican* needlets are shown in Figure 5.1. The shape of the needlet bands is determined by the harmonic function b in equation (4.10), while their width by the parameter B is defined in eq. (4.10). Lower values of B indicate more localization in harmonic space with fewer multipoles contributing to any needlet coefficient, while larger values result in wider harmonic bands. Standard needlet-based NILC pipeline will be referred to as NILC (s), and Mexican needlet-based NILC will be denoted as NILC (m). This terminology will be used uniformly throughout this work.



(A) (B)

FIGURE 5.1: (A) Mexican needlet bands with $B = 1.5$ in harmonic domain. The first low-multipole filter is obtained by merging the first 11 bands using eq. (4.21). (B) Standard needlets with the first 12 bands merged with $B = 1.5$.

Both ILC and NILC are applied to the entire sky. Their results are reported in terms of angular power spectra. The spectra are estimated using the MASTER (Monte Carlo Apodized Spherical Transform Estimator) algorithm [114, 115] utilizing the *pymaster* python package[1] to correct for masking and beam effects. A bin size of $\Delta\ell = 15$ is employed to perform the power spectrum extraction [116].

The averaged foreground residual maps are shown in Figure 5.2 for *PySM* models s3d1, s3d7, and s3d12, showing residuals are stronger along the Galactic plane. Therefore, a Galactic plane mask, illustrated in figure 5.3, has been applied to estimate angular power spectra, excluding 40% and 60% of the sky. This allows us to extract the power spectrum on regions with lower foreground contamination. This reduces the impact of strong foregrounds at low multipoles, resulting in a slightly improved reconstruction of the CMB power spectrum.

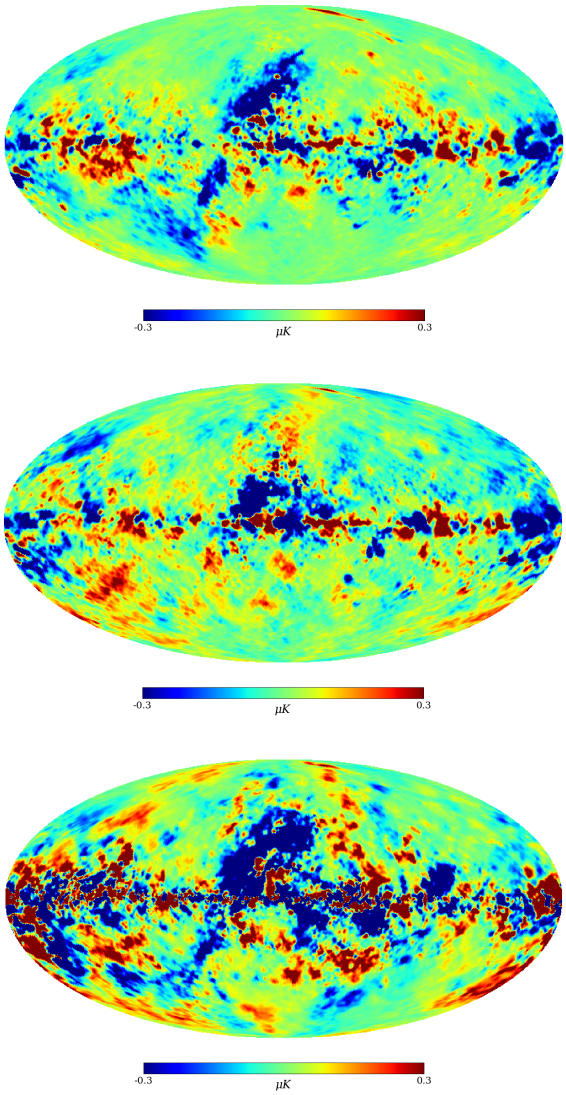


FIGURE 5.2: Average foreground residual B-mode maps from ILC pipeline over 100 simulations for three different foreground models: s3d1 (top), s3d7 (middle), and s3d12 (bottom).

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

For the extraction of the angular power spectra, we specifically consider the publicly released Planck Gal60 and Gal40 masks (shown in figure 5.3), retaining 60% and 40% of the sky, respectively.

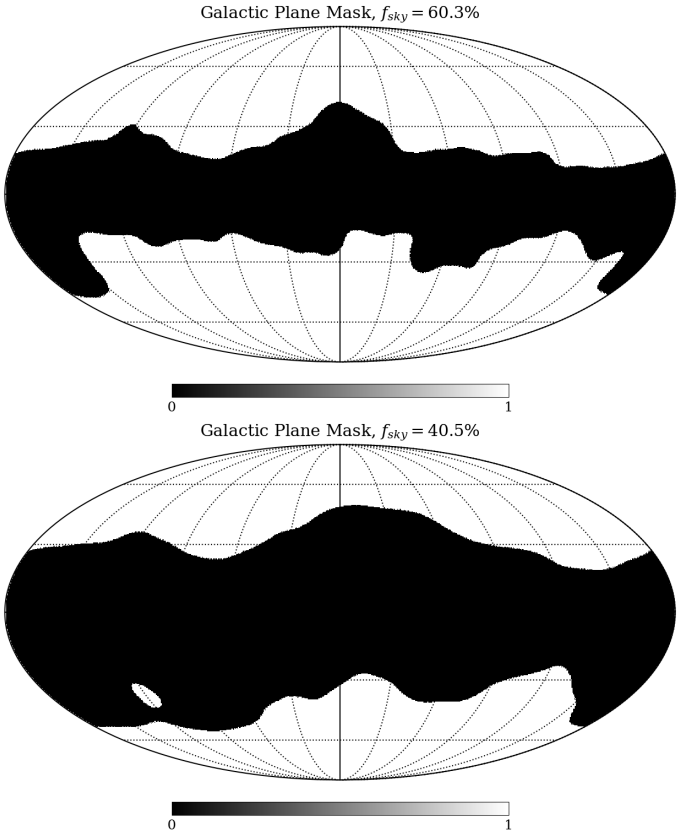
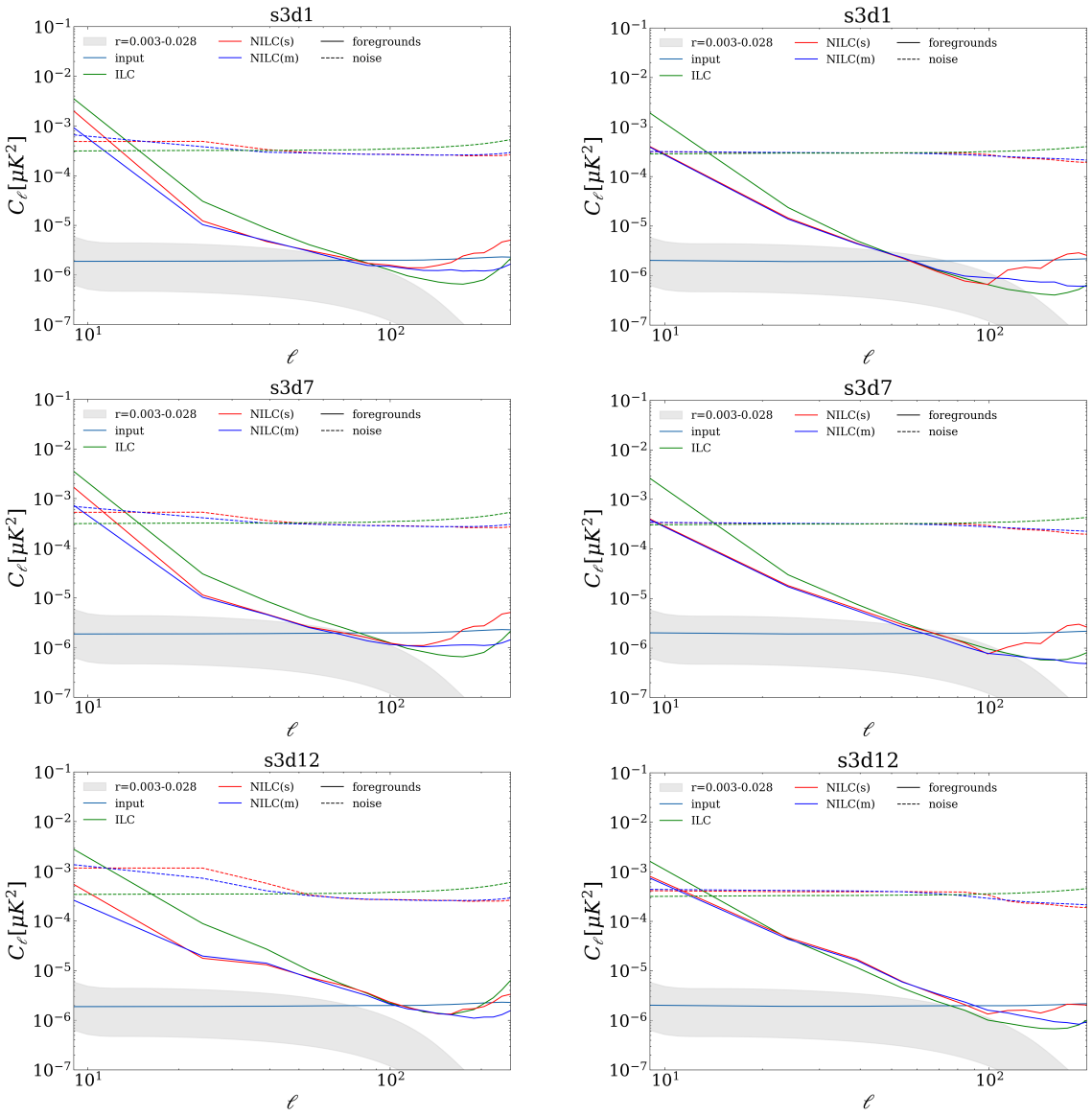


FIGURE 5.3: Planck masks adopted in the analysis of Planck component separated CMB maps.

5.1.1 Results

To apply the ILC and NILC techniques introduced in sections 4.2 and 4.3, we generated 100 distinct realizations of the CMB Q and U maps (see section 3.1), and similarly 100 simulations of Gaussian white and isotropic noise Q and U maps for each of the Planck experiment's frequency channels, for which the sensitivity levels are reported in table 3.1. Dust and synchrotron Q and U foreground maps at the seven Planck polarization frequency bands are also generated. The Q and U maps obtained by combining all these components are then brought to a common angular resolution corresponding to a Gaussian beam with $FWHM = 70$ arcmin accounting for the input angular resolution of each channel as given in eq. (3.8) and eq. (3.9). Corresponding B-mode frequency maps, for use in the analysis, are then obtained through a QU-to-B full-sky transformation. The component separation is applied to Planck simulations assuming three different foreground models (`s3d1`, `s3d7`, `s3d12`) to encapsulate the variability and uncertainty in foreground emissions under varied scenarios. Each combination of models comprises a unique set of parameters and assumptions regarding the behaviour of astrophysical foregrounds.



(A) $f_{sky} = 60\%$ (B) $f_{sky} = 40\%$

FIGURE 5.4: B-mode angular power spectra of noise and foreground residuals for ILC (green), NILC with standard needlets (NILC(s), red) and NILC with Mexican needlets (NILC(m), blue) when applied to the Planck simulated data set. Foreground residuals are displayed in solid lines, while noise residuals in dashed lines. Both are obtained as averages over 100 simulations. Power spectra are computed after the application of Galactic masks retaining f_{sky}=60% (A) and f_{sky}=40% (B). The binning size of $\Delta\ell = 15$ is adopted. Results are reported for three different foreground models: s3d1 (top), s3d7 (middle), and s3d12 (bottom). The input CMB (with r=0) power spectrum averaged over 100 spectra, obtained after applying the same galactic mask to 100 simulated CMB-only maps is shown with the light-blue solid line. The grey shaded area indicates the amplitude range targeted by SO for the primordial tensor signal: $r \in [0.003, 0.028]$.

Figure 5.4 shows the averaged power spectrum of foreground residuals (solid

lines) from the different pipelines, estimated as outlined in section 4.5.1. The ILC foreground residuals are considerably lower than NILC(s) and NILC(m) at small angular scales (high multipoles), indicating the effectiveness of the ILC method in reducing foreground contamination at these scales. This is because ILC minimizes variance in pixel space, where variance is mainly sourced by modes on small scales allowing it to effectively trace foregrounds at these scales. NILC minimizes variance at separate scales. Therefore, it works better at identifying and removing foregrounds at large scales compared to ILC. This feature is crucial for detecting the tensor-to-scalar ratio r on large scales.

For the same reasons mentioned above ILC noise residuals are lower at low multipoles because ILC weights are calibrated mainly on small-scale features, which are mainly noise dominated. On the other hand, NILC performs slightly better at lowering the noise residuals at small angular scales as it implements variance minimization separately at different angular scales allowing it to disentangle the noise dominance at small scales and remove it effectively. These findings are consistent with those presented in [110].

As `d12` is a more complex model compared to `d1` and `d7`, minimizing its impact is more challenging, therefore it shows relatively higher residual levels at power spectrum level than `d1` and `d7`. The residual levels for `s3d1`, `s3d7` and `s3d12` models do not show significant variations at the power spectrum level. As the templates of dust `d1` and `d7` are derived from the Planck data. The `d12` model while not entirely derived from Planck's data, it is developed to match Planck's constraints. Thus, NILC cannot significantly discriminate among different models.

Results are quite stable with both masks. With the more conservative 40% mask, residuals for the two NILC implementations are more compatible, and we use this mask for deriving constraints on r. The posterior distribution for the effective foreground contamination to r is shown in Figure 5.5 and is estimated using eq. (4.23) on 40% of the sky. Since our simulated data have no primordial B-modes ($r = 0$), the only contribution to r comes from residual foregrounds. In particular, we also assume the use of advanced de-noising techniques, such as Monte-Carlo simulations, that can effectively reduce the noise, allowing us to neglect it in our analysis (see section 4.5.2). Thus, we fit the foreground residuals spectrum C_ℓ^{fgds} with a primordial tensor spectrum with varying r ranging from 10^{-4} to 1. The large foreground residuals in the case of Planck lead to highly biased posterior distributions. The detection at 68% confidence level (CL) is $r \sim 10^{-1}$, which is much higher than the target detection level of future CMB experiments.

5.2 Conclusion

The ILC approach does not perform well with the Planck data set, leading to a clear bias on the value of r due to residual foregrounds. This is because ILC does not make any assumptions about the foreground models and minimizes the variance of the final map. However, since the polarization data from Planck is heavily dominated by noise, ILC's ability to mitigate the foreground contribution in B-modes, which is sub-dominant at all frequency channels, is limited. Additionally, excluding 60% of the sky fraction shows only a slight reduction in the residual

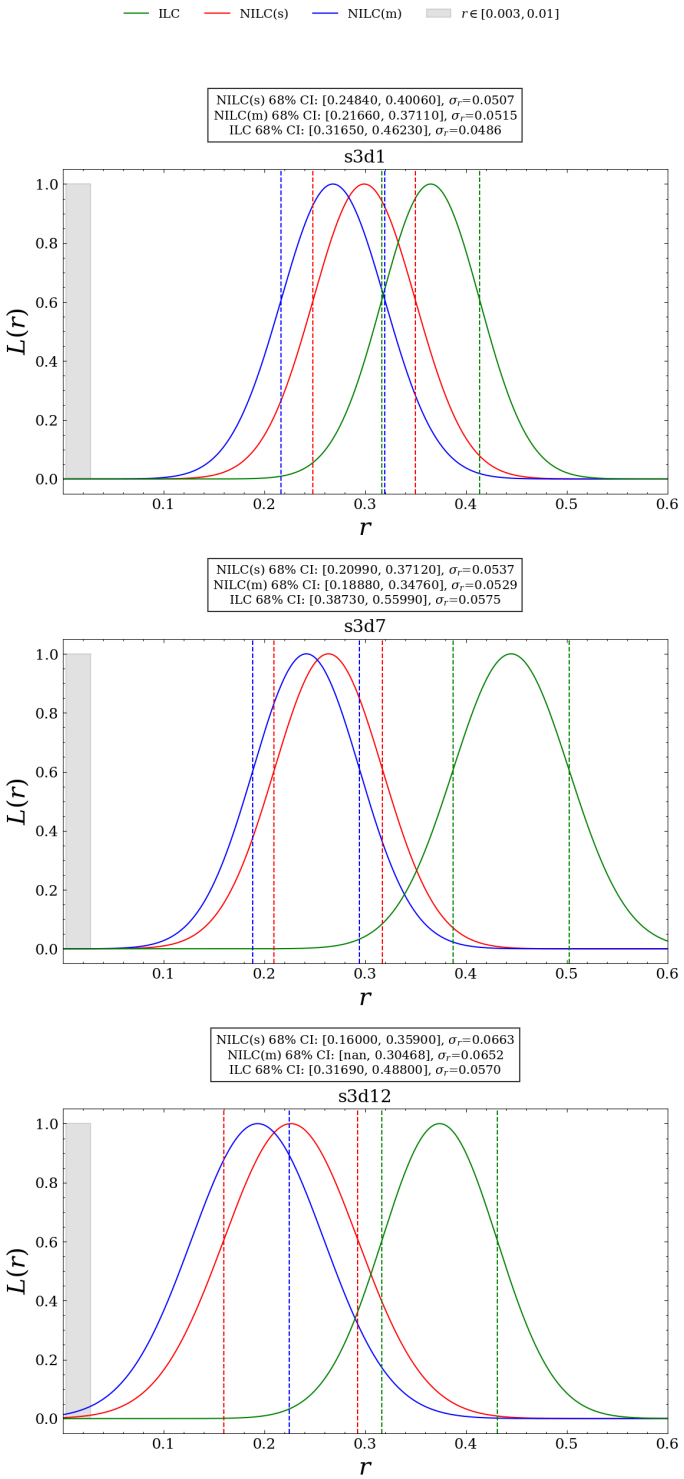


FIGURE 5.5: The posterior distribution of an effective tensor-to-scalar ratio fitted to the NILC and ILC foreground residuals over $f_{sky} = 40\%$. The 68% CI are marked with dashed lines. The grey shaded band indicates the amplitude range for the primordial tensor signal target by future experiments: $r \in [0.003, 0.028]$.

levels, but it is still above the required sensitivity level. As the Planck satellite has limited sensitivity to polarization, the data collected by Planck is not sufficient to

reach an accurate cosmological measurement of the tensor-to-scalar ratio.

The uncertainty $\sigma_r(r = 0)$ on r, we have found from our simulated data ranges from 0.04 to 0.06. This is not compatible with the upper bound on r of 0.028, obtained by combining data from Planck and BICEP2/Keck [23, 24] as expected. Instead, it is comparable with $\sigma_r = 0.069$, found by the analysis of B-modes in [51].

Although both ILC and NILC give bias on r due to residual foreground contamination, we find that NILC outperforms ILC. NILC is more effective in reconstructing r, given the fact that most of the constraining power is on large scales where NILC better removes foregrounds. Therefore, in the next chapter, we will investigate the performance of the NILC method on higher sensitivity SO, SP and SL datasets to constrain the value of r.

Chapter 6

SO Data Analysis

The Simons Observatory (SO) Small Aperture Telescopes (SATs) aim to measure the primordial tensor-to-scalar ratio r with high precision, targeting an uncertainty of less than 0.003 by observing the polarized CMB over a sky region with $f_{sky} \sim 34\%$ [20]. In this chapter, we present the analysis of the NILC and MC-NILC pipelines introduced in sections 4.3 and 4.4 respectively, and compare their performance on a set of realistic SO sky simulations. Their effectiveness is evaluated in terms of bias (see section 4.3.1), foreground and noise contamination, and statistical uncertainty on r using a Gaussian likelihood as already outlined in section 4.5. Due to limited sky coverage and large-scale atmospheric contamination, SO will not access information on scales larger than $\ell < 30$, thus being able to constrain r only through the observations of the recombination peak. We utilize the same needlet bands for both NILC and MC-NILC illustrated in Figure 5.1, but the separation methods exclude modes with multipoles below $\ell = 32$, as SO-SATs cannot detect multipoles $\ell < 30$ due to atmospheric contamination. In this chapter, we consider the same models for Galactic foreground emission as in Planck's analysis. This analysis excludes the Simons Observatory's Large Aperture Telescope (LAT), as it is not intended for primordial B-mode studies.

SO Sky Coverage

The SO SAT is designed to cover approximately 30-40% of the sky. However, due to the uneven scanning strategy of the telescope, certain regions of the sky will be observed more frequently than others. This concentrated observation in specific areas results in an effective sky fraction of about 10% [20], as depicted in Figure 6.1. Since these regions will be scanned multiple times, they will consequently have lower noise levels, leading to higher data quality and more precise measurements than the less frequently observed areas. Thus, we perform the subsequent analysis for the Gaussian noise on this 10% of the sky as shown in Figure 6.1(top). Figure 6.1(bottom) shows the observed 34% fraction of the sky by SO-SAT over which we have implemented the NILC method using the realistic noise model.

6.1 Gaussian Noise vs Realistic Optimistic Noise

In this section, we present the analysis of NILC on simulated data set of SO experiment with both Gaussian and Realistic noise, even considering the combination of SO data with high-frequency data from Planck (SP) and LiteBIRD (SL) according

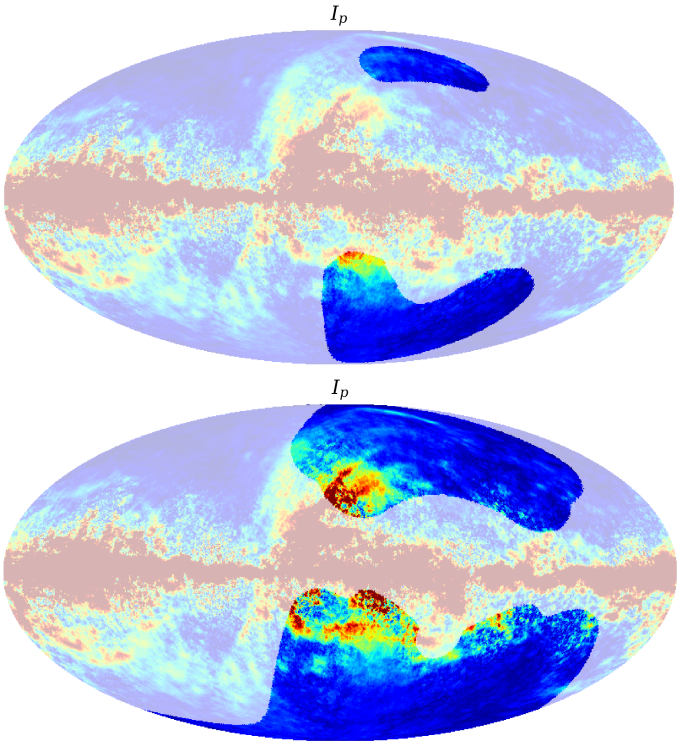


FIGURE 6.1: Top: Coverage of the sky by the SO-SAT experiment mapped in Galactic coordinates. The unshaded area represents the most observed 10% fraction of the sky. Bottom: Full Coverage of the sky by SO-SAT experiments. The unshaded area represents the observed 34% part of the sky.

to the experimental configurations outlined in section 3.5. These data sets are generated in the same way as presented in section 5.1.

6.1.1 NILC Results

NILC is implemented on the SO, SP and SL simulated maps. Gaussian isotropic noise model indicating the optimistic level of white noise is simulated with the baseline senstivity levels given in section 3.5 for each frequency channel, and smoothed with an effective kernel according to eq. (3.8) and eq. (3.9) (see section 3.4). The results for this model are obtained considering the component separation of the most observed 10% region of the sky as shown in Figure 6.1(top). The realistic noise model is simulated according to assumptions introduced in section 3.4. The cleaning procedure is performed, in this case instead, on $f_{sky} = 34\%$ as shown in Figure 6.1(bottom), as it is the complete region observed by SO. However, for a direct comparison of both noise models, the angular power spectra are computed, also in this case, over the patch of most observed pixels with $f_{sky} = 10\%$ shown in Figure 6.1(top). The angular power spectra (averaged over 100 simulations) of foreground and noise residuals for Gaussian, isotropic noise and realistic noise are illustrated in Figure 6.2. They are reported within the multipole range employed for the analysis of SO-SATs simulations: ($32 < \ell < 256$). From Figure 6.2 (bottom pannel), we notice that noise residual levels are higher

in the case of realistic noise for the SO because the realistic noise model accounts for the actual more complex structure of the instrumental noise across the sky. This includes the inhomogeneous scanning strategy of the telescopes and the $1/f$ term accounting for correlated noise and atmospheric contamination. In contrast, the Gaussian white noise model represents a highly optimistic scenario where the noise is assumed to be purely statistical, and uniform across the sky for each frequency channel. This idealized model does not capture some of the additional sources of noise that are present in a real experimental setup. Other conclusions that can be drawn from Figure 6.2 are:

- For *PySM* model `s3d1`, the foreground residuals are slightly higher than `s3d7`. However, the noise levels are relatively consistent among the two models. The `s3d12` model shows higher foreground residuals, largely because of the increased complexity due to multiple dust components along the line of sight which causes much more significant SED distortions and variations, if compared to the other models. This result highlights that NILC's scale-dependent approach struggles to locally track multiple distinct SEDs simultaneously, leading to higher residual dust contamination.

- For s3d12 we observe a clear improvement from introducing LiteBIRD. Apart from this, we do not observe any significant improvement among SO, SP and SL data sets. These results suggest that the addition of Planck or LiteBIRD high-frequency channels does not introduce beneficial additional information in the SO patch for this less complex sky model.

- For the SO and SP datasets, the foreground and noise residuals are generally comparable across all combinations of PySM models. However, a notable reduction in noise residuals is observed for the SL dataset, which combines data from SO (Simons Observatory) and LiteBIRD with the experimental configuration listed in Table 3.5. This improvement is primarily due to the inclusion of two additional high-frequency channels at 337 GHz and 402 GHz. These channels not only extend the frequency coverage but also feature decreased instrumental noise and enhanced sensitivity with respect to SO and SP, particularly for detecting foreground components, like dust, which are more prominent at higher frequencies. The improved sensitivity of SL high-frequency observations in these channels leads to lower overall noise residuals. The finer angular resolution (FWHM) of the higher frequency channels further contributes to the precision of the measurements, enhancing the overall data quality. In the case of `s3d12`, the residuals are lower when the SL data set is used.

- The foreground residuals in the case of realistic noise are similar to those in the case of Gaussian white, isotropic noise, with some improvement at intermediate angular scales especially for the sky models `s3d7` and `s3d1`. Since the foregrounds are the dominant signal in many regions of the sky, the presence of either realistic noise or Gaussian white noise has a limited impact on the overall residuals after applying NILC.

- The NILC method achieves foreground residuals below the SO target detection level of $r \sim 10^{-2}$ with comparable NILC performance for standard and

mexican needlets, except for the `s3d12` model, where NILC weights attempt to locally track multiple distinct SEDs simultaneously, resulting in noisier and less efficient removal of Galactic emission in each pixel.

- Noise residual levels are higher in the case of realistic noise than Gaussian isotropic, white noise model.

An assessment of the NILC bias in reconstructed CMB B-modes, as introduced in section 4.3.1 and computed with eq. (4.22) for both noise models, is reported in figure 6.3. For all data sets, sky models and noise assumptions, the NILC bias is well within the 1σ dispersion of the mean power spectrum, computed as $\frac{\sigma(C_\ell^{bias})}{\sqrt{N_{sims}}}$, with $N_{sims} = 100$ the number of simulations. These results indicate that the configuration employed in this analysis for both mexican and standard needlet bands (see Figure 5.1) does not introduce any significant bias in the partial sky reconstruction of the CMB signal.

6.1.2 Likelihood Estimation

In this section, we analyse how the estimate of r is impacted by foreground and noise residuals in the B-mode solutions obtained for the SO patch (see section 4.5.2). In eq. (4.20), the primary systematic bias to the final angular power spectrum arises from the residual Galactic contamination. To evaluate the effect of this residual term on the estimation of r from C_ℓ^{out}, we fit r on binned C_ℓ^{fgds} using a Gaussian likelihood as described in eq. (4.23). We assume the use of advanced de-noising techniques such as Monte-Carlo simulations that can effectively reduce the noise, allowing us to neglect it in our analysis. For the estimation of the likelihood, a binning scheme of $\Delta\ell = 15$ has been used to make the angular power spectrum Gaussian distributed (see section 4.5.2).

Likelihood results for the three data sets and three *PySM* models are reported in Figure 6.4 for Gaussian white noise and in Figure 6.5 for realistic SO noise. The likelihood is consistently estimated by excluding the multipole range of $\ell < 32$. The posterior on r is mildly biased for `s3d1` and `s3d7`, while significantly biased for `s3d12`, due to larger foreground residuals. As expected, the posteriors on r for the case with realistic noise are broader since the level of the noise residual is higher.

The posterior distribution for all datasets demonstrates NILC effectiveness in removing the majority of foreground contamination, leading to an upper limit on r from foreground residuals of $r \sim 4 \cdot 10^{-3}$ at 68% CL with $\sigma_r \sim 10^{-3}$ across the `s3d1` and `s3d7` models. For `s3d12`, instead, the detection remains statistically very significant due to increased foreground contamination. To overcome these biased distributions and enhance measurement accuracy, we therefore study the application of the MC-NILC pipeline. This advanced approach aims to further refine foreground signal subtraction and reduce uncertainties on r by utilizing more sophisticated subtraction methods to address more complex foreground emission, as outlined in section 4.4.

The same analysis is repeated with a more conservative mask excluding the 0.2% of brightest regions shown in Figure A.1 using the Gaussian white, isotropic noise model. The masking strategy and results are reported in Appendix A.

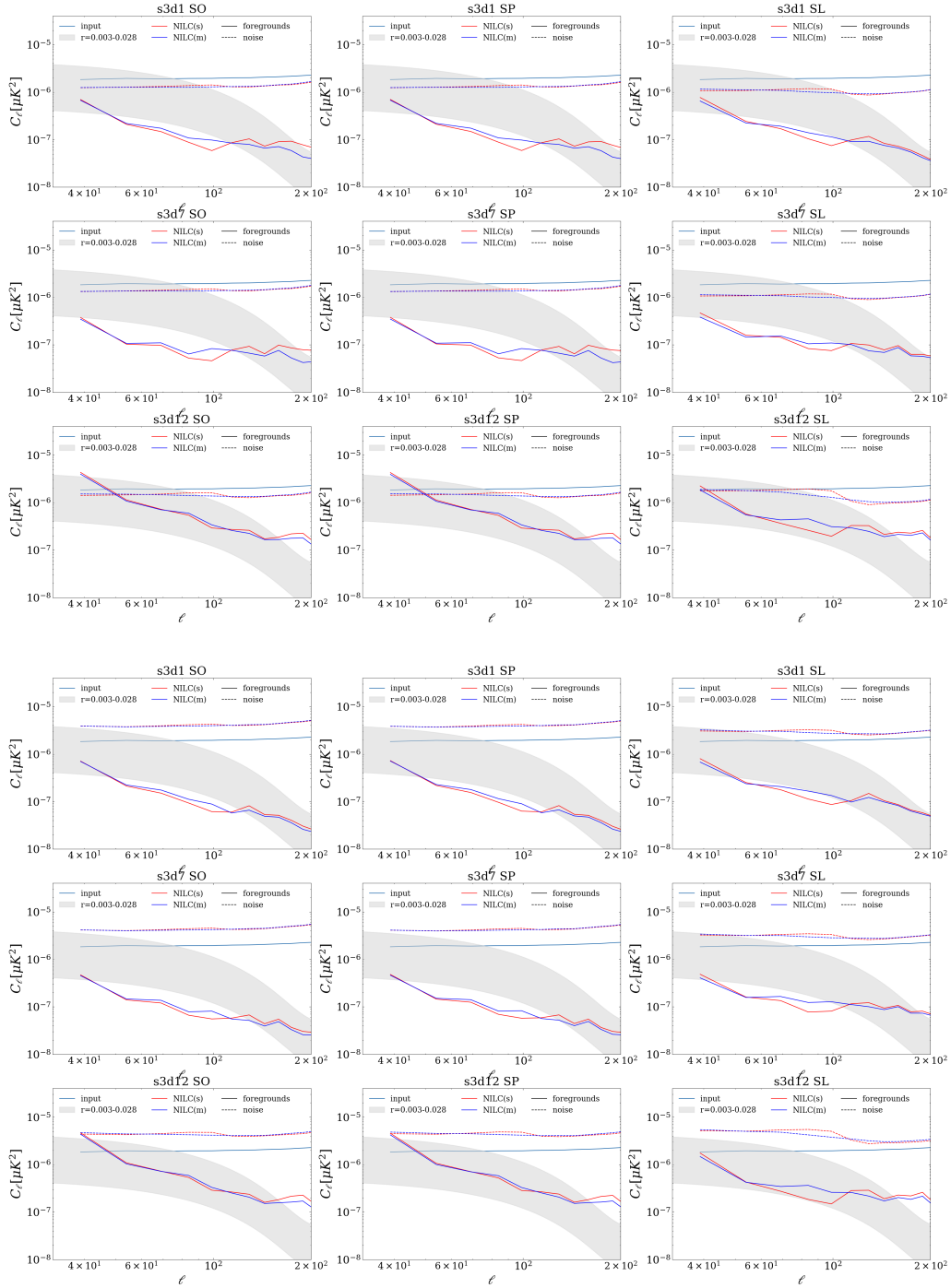


FIGURE 6.2: B-modes angular power spectra averaged over 100 simulations for foreground and noise residuals, respectively shown with solid and dashed lines. Two needlet bases are considered i) Standard (red) and ii) Mexican (blue). Power spectra are computed within the SO 10% most observed region (shown in Figure 6.1) and they are binned with an interval of $\Delta\ell = 15$. In the top panel, results are given for Gaussian white, and isotropic noise, and in the bottom panel for realistic noise. The input CMB (with r=0) power spectrum, averaged over 100 spectra, obtained after applying the same SO mask to 100 simulated CMB-only maps is shown with the light-blue solid line. The grey-shaded region indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$. Results are reported for three different foreground models: s3d1 (top), s3d7 (middle), and s3d12 (bottom) and three data sets: SO (left), SP (middle) and SL (right).

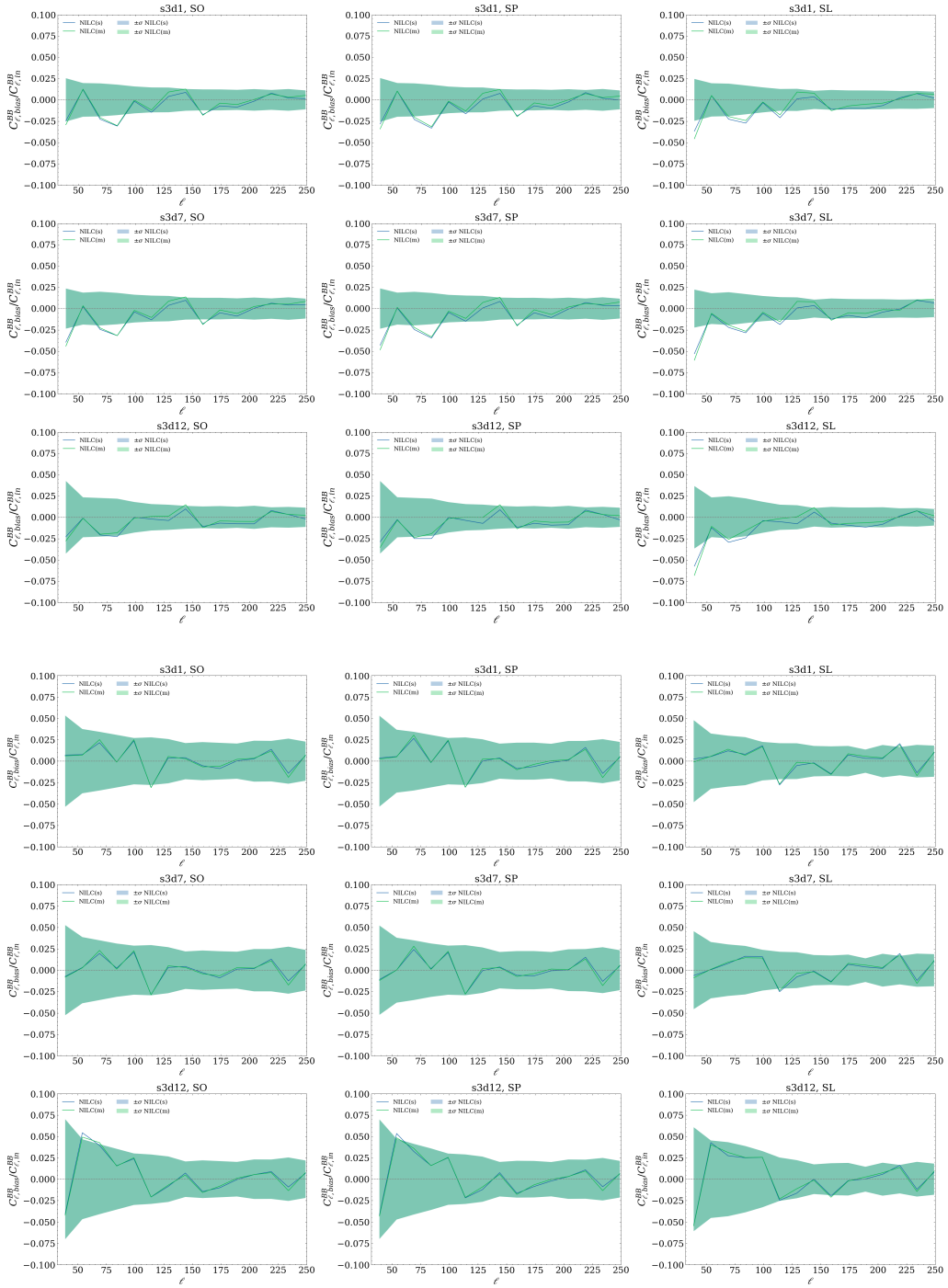


FIGURE 6.3: NILC reconstruction bias $C_{\ell,bias}^{BB}$ estimated with eq. (4.22) over $f_{sky} = 10\%$, whereas $C_{\ell,in}^{BB}$ is the averaged binned CMB B-mode angular power spectra estimated from CMB B-mode simulations on the same fraction of the sky. The uncertainty on the power spectrum of the average bias is estimated as $\frac{\sigma(C_\ell^{bias})}{\sqrt{N_{sims}}}$. In the top panel, results are given for Gaussian white, and isotropic noise, and in the bottom panel for realistic noise. Results are reported for three different foreground models: s3d1 (top), s3d7 (middle), and s3d12 (bottom) and three data sets: SO (left), SP (middle) and SL (right), and for both standard and mexican needlets

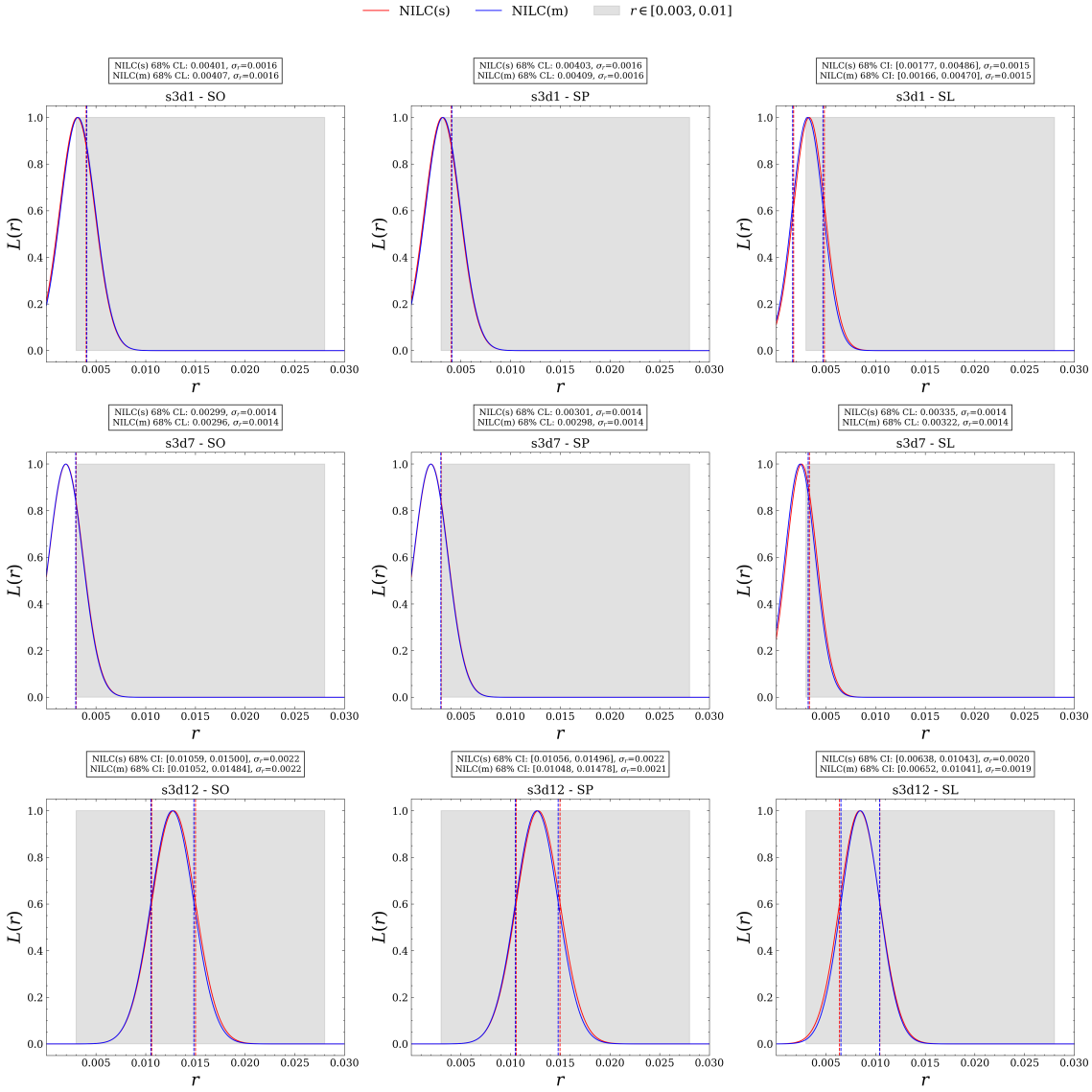


FIGURE 6.4: $L(r)$ (described in section 4.5.2 obtained with eq. (4.23)) obtained from NILC application for Gaussian white, isotropic noise simulations, computed considering angular power spectra (averaged over 100 simulations) estimated in $f_{sky} = 10\%$ SO patch and with $\Delta\ell = 15$. 68% CI for mexican (in blue) and standard (in red) are shown with dashed vertical lines. The grey-shaded region indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$. Results are reported for three different foreground models: d1s3 (top), d7s3 (middle), and d12s3 (bottom), and for three different data sets: SO (left), SP (middle) and SL (right).

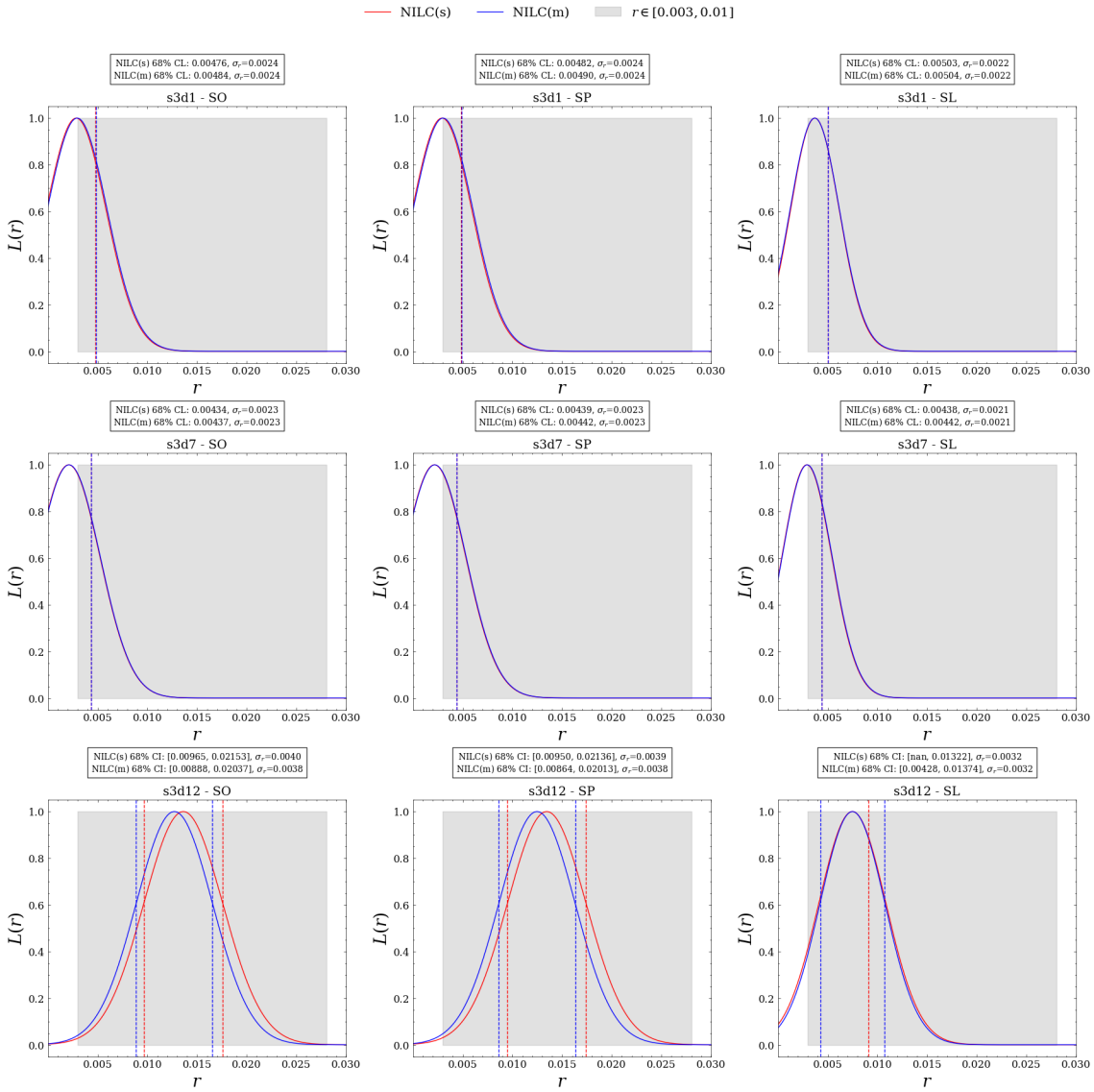


FIGURE 6.5: $L(r)$ from NILC application in the case of realistic SO noise simulations, computed considering angular power spectra (averaged over 100 simulations) computed in the SO patch with $f_{sky} = 10\%$ and $\Delta\ell = 15$. 68% CI for mexican (in blue) and standard (in red) are shown with dashed vertical lines. The grey-shaded region indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$. Results are reported for three different foreground models: s3d1 (top), s3d7 (middle), and s3d12 (bottom), and for three different data sets: SO (left), SP (middle) and SL (right).

6.2 MC-NILC Analysis

In the MC-NILC application presented in this thesis work, 280 GHz is selected as the higher frequency channel and 93 GHz as the lower frequency channel to build the effective spectral index map. This configuration enables the simultaneous tracing of spatial variations in polarized thermal dust and synchrotron emission in B-modes. In this analysis, the tracer is constructed in an ideal framework from foregrounds-only maps, as outlined in section 4.4. The adopted needlet configurations are those shown in Figure 5.1. For each needlet scale j, the foreground needlet coefficients $\beta_{f,j}$ are obtained by filtering the input foreground map with the corresponding needlet band and then the ratio $\beta^{280}_{f,j}/\beta^{93}_{f,j}$ is built. The ratio is then thresholded to obtain a partition with equal area patches in the SO footprint. An example of a partition with 5 different patches (for visualization purposes) is shown in Figure 6.6. This map provides in first approximation an estimation of

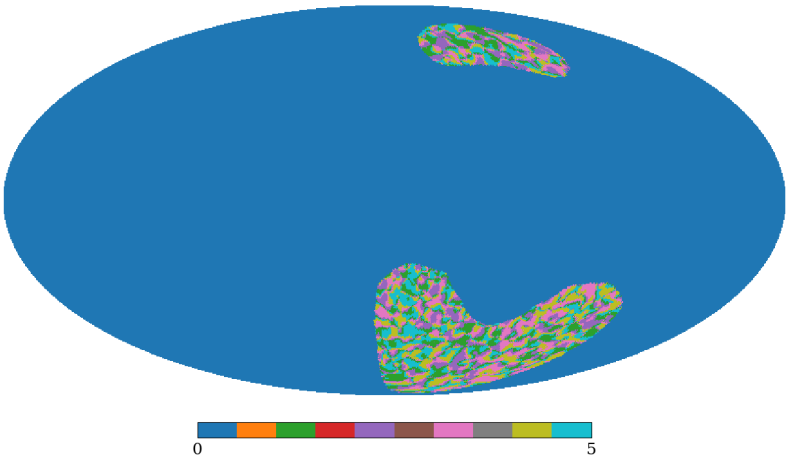


FIGURE 6.6: Partition of the SO patch in 5 clusters of equal area, constructed thresholding a tracer map obtained as the ratio of the foreground (s3d1) simulated maps at $\nu_1 = 280$ and $\nu_2 = 93$ and filtered with a needlet band for $j = 0$ (see section 4.4).

how the spectral characteristics of Galactic foregrounds in B-modes are distributed spatially across large and intermediate angular scales.

To apply the MC-NILC pipeline effectively, the optimal number of patches, denoted as N, is determined by evaluating the bias on r from foreground residuals in the reconstructed CMB data for different values of N. The obtained residuals and posteriors for the different cases are displayed in Figure 6.7 (top). This plot suggests that there is no sensitivity to different patch numbers after N=20. As illustrated in Figure 6.7 (bottom), the foreground residuals start to decrease as N increases, indicating enhanced separation of the foregrounds from the CMB signal. However, this reduction in residuals is not linear. After N=20, the decline in foreground residuals starts to plateau. Consequently, increasing the number of patches beyond this threshold does not significantly enhance the removal of foreground contamination. Therefore, for subsequent analysis, we have opted for N=20 clusters to prevent any bias in the CMB reconstruction. The variance of the output solution is then minimized separately within each patch.

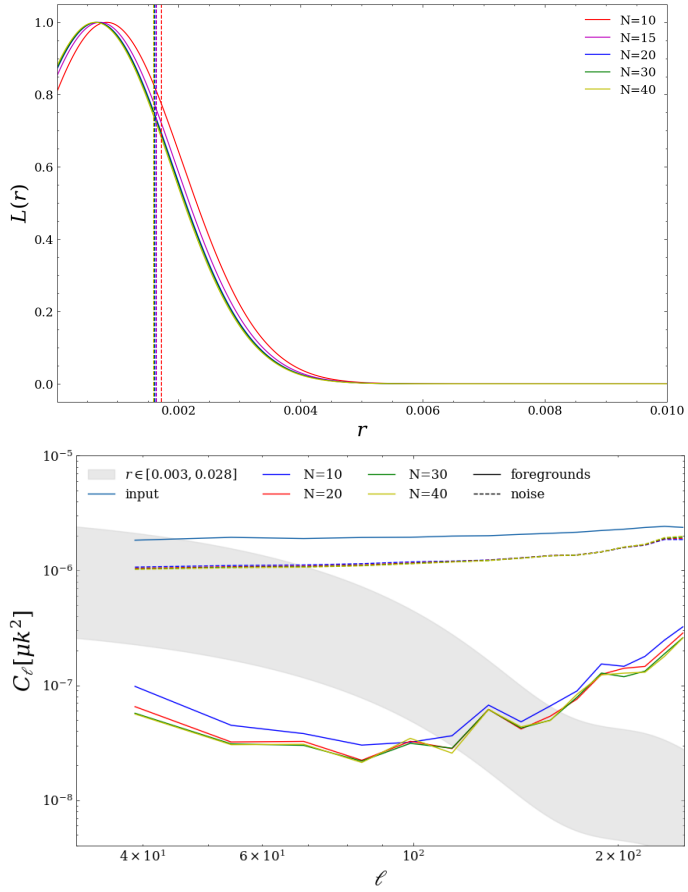


FIGURE 6.7: Top: Posterior distribution on r derived from fitting the average power spectrum of the MC-NILC foreground residuals C_ℓ^{fgds} for different partition numbers, N. The dashed lines indicate the 68% upper bound. The PySM model used for Galactic emissions is s3d1. The posterior distributions for the effective tensor-to-scalar ratios were obtained using a binning interval of $\Delta\ell = 15$. Bottom: Average angular power spectra of MC-NILC foreground (solid lines) and noise (dashed lines) residuals for various partition numbers N, from the analysis of 100 simulations of the SO data set. The input CMB (with r=0) power spectrum averaged over 100 spectra, obtained after applying the same SO mask to 100 simulated CMB-only maps is shown with the light-blue solid line. The grey shaded area indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$.

6.2.1 MC-NILC Results

The results shown in this section are obtained considering the *PySM* `s3d1` and `s3d12` models for the Galactic emission. We have used Gaussian white noise in this section. Foreground and noise residuals obtained with MC-NILC are shown in Figure 6.8. We observe that MC-NILC results in lower foreground residuals than

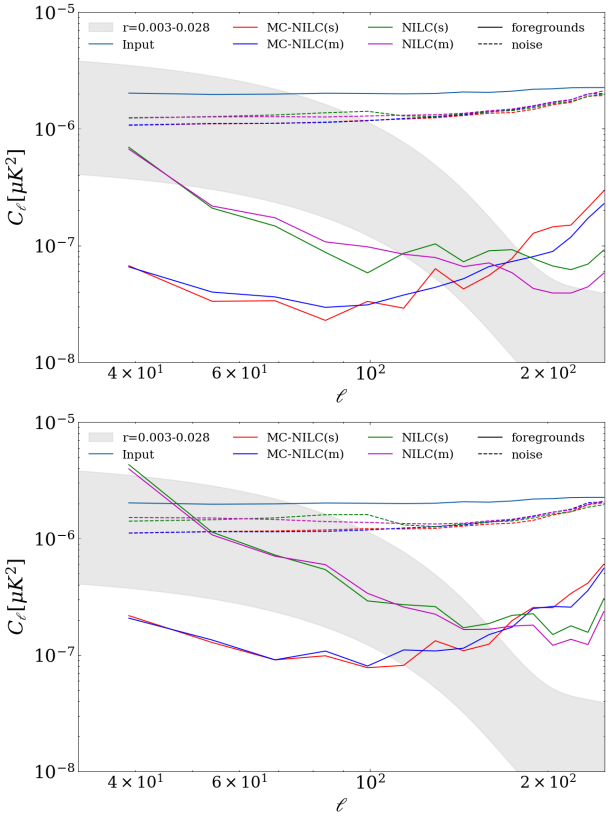


FIGURE 6.8: Averaged angular power spectra of foregrounds (solid lines) and noise residuals (dashed lines), binned with an interval of $\Delta\ell = 15$, obtained from the application of MC-NILC and NILC with standard needlets (MC-NILC (s), blue; NILC(s), green) and mexican ones (MC-NILC (m), red; NILC(m), purple). Spectra are estimated within the SO 10% most observed region (shown in Figure 6.1). The input CMB (with r=0) power spectrum averaged over 100 spectra, obtained after applying the same SO mask to 100 simulated CMB-only maps is shown with the light-blue solid line. The grey shaded area indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$. Top pannel: `s3d1` and Bottom pannel: `s3d12`

NILC across all the relevant angular scales for both *PySM* models. This reduction is particularly noticeable at the largest scales, where the power of foreground residuals is more than an order of magnitude lower than that of NILC. Aside from reducing foreground contamination, the MC-NILC solutions also show lower noise residuals than those of NILC. This improvement is expected, as the minimization process is conducted within each patch, where the B-mode foregrounds possess similar characteristics and the spectral properties of foregrounds remain constant

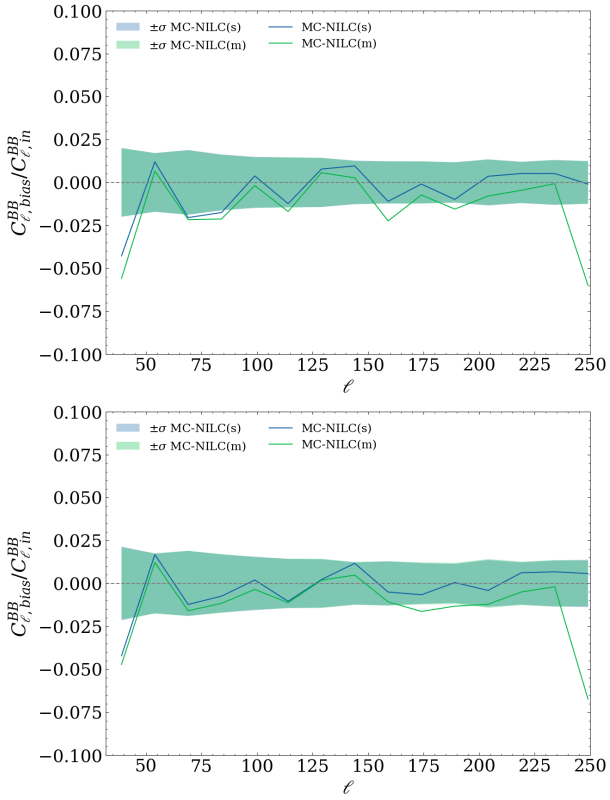


FIGURE 6.9: MC-NILC reconstruction bias $C_{\ell,bias}^{BB}$ is estimated with eq. (4.22) over $f_{sky} = 10\%$, whereas $C_{\ell,in}^{BB}$ is the averaged binned CMB B-mode angular power spectra estimated from CMB B-mode simulations on the same fraction of the sky. The uncertainty on the power spectrum of the average bias is estimated as $\frac{\sigma(C_\ell^{bias})}{\sqrt{N_{sims}}}$. Results are reported for `s3d1` (top) and s3d12 (bottom) models.

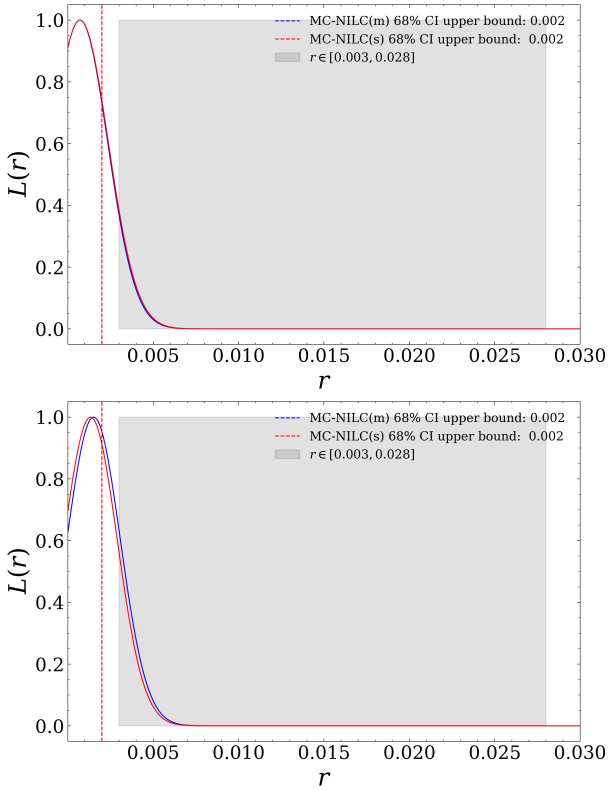


FIGURE 6.10: Posterior distribution on r for MC-NILC with mexican needlets (in blue) and standard needlets (in red). 68% CI for mexican (in blue) and standard (in red) are shown with dashed vertical lines. The grey shaded area indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$. Results are reported for `s3d1` (top) and s3d12 (bottom) models.

across the patch for each model. A simpler foreground sky is expected to lead to lower residuals from both contaminants as the weights are automatically better calibrated to subtract instrumental noise.

Figure 6.9 shows the amount of CMB reconstruction bias (see section 4.3.1) relative to the input CMB B-mode signal, given by the application of MC-NILC on a simulated dataset. We observe that the bias is under control, compatible with zero given the uncertainty of the average C_ℓ^{bias}. This quantity is estimated as $\frac{\sigma(C_\ell^{bias})}{\sqrt{N_{sims}}}$, where $N_{sim} = 100$.

The plot in Figure 6.10 shows that the effective tensor-to-scalar ratio fitted to the Galactic foreground residuals has a posterior distribution (computed according to section 4.5.2) with an upper limit at 68% CL of the order of $r \sim 2 \cdot 10^{-3}$ for MC-NILC (with mexican and standard needlets) with $\sigma_r = 0.001$ for both models. This upper bound is compatible with the target of SO. We have found this constraint by excluding the multipoles $\ell < 32$. This upper bound is mainly sourced by the contribution of the noise residuals to the covariance matrix in eq. (4.23).

6.3 Conclusion

NILC provides an alternative to parametric methods. Its needlet bases enable the linear combination weights to adapt to the local conditions of foreground and noise contamination in the pixel domain. This prioritizes foreground removal at large angular scales and noise subtraction at smaller ones. For the considered *PySM* models, NILC shows potential because it depends very little on predefined assumptions, providing an alternative to parametric component separation methods for analyzing B-mode data in forthcoming CMB experiments. However, substantial evidence indicates that the spectral energy distribution of polarized Galactic emissions significantly varies across the sky, making it challenging for NILC to accurately estimate and reduce foreground contamination. This is especially important for B-mode analysis, where the primordial signal is much weaker in comparison to the Galactic foregrounds.

For `s3d1` and `s3d12` models, we show that the MC-NILC partitioning effectively captures the varying characteristics of the foregrounds across the SO patch. This leads to lower residual contamination and tighter constraints on the tensor-to-scalar ratio compared to NILC.

MC-NILC analysis demonstrates significant improvements in the reduction of the noise and foreground residuals compared to the standard NILC method. By partitioning the sky into 20 regions based on the spatial distribution of Galactic foreground spectral properties, the MC-NILC pipeline achieves lower residuals, enhancing the separation of foregrounds from the CMB signal. NILC provides results that are marginally consistent with the SO targets for s3d1 and s3d7 (with an upper bound at 68% CL of $r \sim 4 \cdot 10^{-3}$ and $r \sim 3 \cdot 10^{-3}$), while we have verified that it is not able to properly reduce contamination for more complex models such as s3d12. The increased flexibility of MC-NILC allows better control of the foreground and increases sensitivity. We demonstrated this for the s3d1 model (upper bound at 68% CL of $r \sim 2 \cdot 10^{-3}$), and verified it for a complex foreground model as `s3d12` (upper limit at 68% CL of $r \sim 2 \cdot 10^{-3}$) and we expect comparable results for `s3d7`. Hence, MC-NILC indeed allows us to meet the SO targets and presents a valuable method for analyzing B-mode data in upcoming CMB experiments.

Conclusion

The primary objective of future CMB experiments will be to detect polarization B-modes created by primordial tensor perturbations. Such a detection would finally confirm the inflationary scenario. A major challenge in achieving this objective is Galactic foreground emission, which overshadows the primordial signal by several orders of magnitude. In this thesis, we investigated blind component separation methods to mitigate this contamination with minimal modelling assumptions. We first implemented the NILC pipeline on our simulated data sets using spherical wavelets, which enables the linear combination weights to adjust to the local conditions of the foregrounds at large scales and noise contamination on smaller angular scales. We then extended this approach to a multi-clustering method called MC-NILC as proposed in [111]. This approach improves the NILC technique by independently minimizing variance in various sky regions (clusters or patches) selected to have similar spectral characteristics of B-mode Galactic emission.

To test these techniques we generated realistic simulations of multi-frequency datasets for present and future experiments. In particular, we considered the ground-based Simons Observatory (SO), which started taking data at the beginning of 2024. In the analysis of the SO simulated set, we also considered a combination with high-frequency measurements from space, like those from Planck and those expected by LiteBIRD.

Utilizing the simulated Planck data, we demonstrated that due to the Planck satellite's limited sensitivity to polarization, both ILC and NILC give bias on r due to residual foreground contamination. Additionally, they introduce significant uncertainties in r of the order of $\sigma_r \sim 10^{-2}$. However, we found that NILC outperforms ILC given the fact that most of the constraining power is on large scales where NILC better removes foregrounds. Therefore, we applied the NILC pipeline to the SO simulated data by considering different levels of complexity for the foreground models: `s3d1`, `s3d7` and `s3d12`. Such analysis leads us to the following results:

- NILC proved to be effective in lowering the foreground and noise residuals for `s3d1` and `s3d7`. However, it resulted in a higher residual level for `s3d12` due to the higher complexity of the dust model.
- The results indicated that incorporating high-frequency channels from Planck and LiteBIRD does not result in any notable improvement in the SO patch, apart from a reduction in noise contamination for the case of the inclusion of LiteBIRD data.
- NILC provided results that are marginally consistent with the SO targets for `s3d1` and `s3d7` (with an upper bound at 68% CL of $r \sim 4 \cdot 10^{-3}$ and $r \sim 3 \cdot 10^{-3}$), while it is not able to properly reduce contamination for the more complex model `s3d12` (resulting in a bias on $r \sim 1 \cdot 10^{-2}$).

- We compared the performance of the NILC pipeline on two noise models: Gaussian white, isotropic model and realistic noise for the SO experiment. We concluded that the realistic noise model accounting for non-uniform sky coverage and $1/f$ noise component leads to higher noise residuals, highlighting the impact of atmospheric and instrumental realistic effects. Whereas, foreground residuals remain similar in both cases.

Finally, we applied MC-NILC, in this thesis, for the first time, to SO simulated data with Gaussian white and isotropic noise on two foreground models: `s3d1` and a more complex foreground model `s3d12`. NILC was not able to mitigate the contamination from `s3d12` and left a bias on r. However, with MC-NILC we were able to achieve considerably lower foreground and noise residuals than NILC and found an upper bound for both the models at 68% CL of $r \sim 2 \cdot 10^{-3}$, which is compatible with the SO science targets.

The work developed in this thesis places an important step in the extension of optimized minimum variance techniques (as MC-NILC) to partial-sky ground-based CMB observations. In the future, this study could be improved in several ways. First, we intend to extend the application of the MC-NILC method to the case that includes a more realistic SO noise model. In this thesis, we used an idealized approach to derive the tracer of the foreground emission that is needed to generate the sky partition. Specifically, we adopted the ratio of foreground maps simulated at different frequencies. However, in a real-case scenario, we cannot access this information. We aim to extend this study to a case where the MC-NILC tracer is obtained from the full simulated data through a realistic approach.

Continued advancements in component separation methods will be crucial for ensuring the success of next-generation experiments, such as the Simons Observatory [20], CMB-S4 [52], and LiteBIRD [26].

Appendix A

SO 3σ masking

Some pixels in a map often need to be masked even after a foreground reduction technique has been applied. The purpose of a foreground mask is to exclude regions in the sky which are supposed to be still significantly contaminated by foreground residuals so that the estimated power spectrum and the following cosmological analysis are not affected by such contaminants. As done in [111], such a mask is generated using the averaged foreground residual map smoothed to 3°, which is thresholded to obtain a binary mask shown in Figure A.1, with a slightly lower f_{sky} value of 9.8%. The threshold value is obtained by averaging the standard deviation of the 100 CMB B-mode maps generated with $r = 0.003$ and smoothed with the same resolution. The averaged foreground map is smoothed before applying the threshold to reduce small-scale fluctuations, which are primarily caused by noise and CMB signals in real data. This step helps prevent the resulting masks from having a patchy structure.

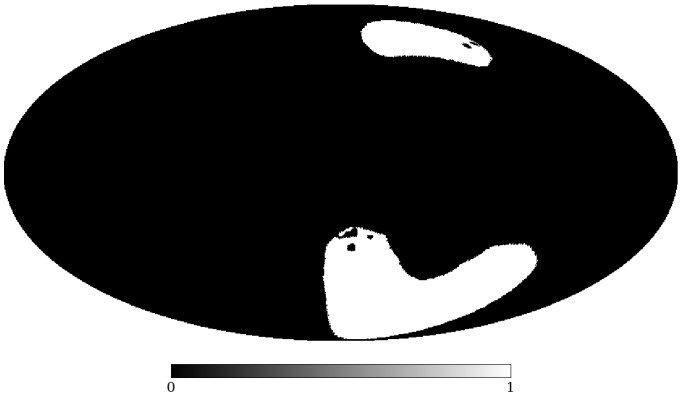


FIGURE A.1: Foreground mask obtained after setting the pixels above the threshold $3\sigma = 0.03$ equal to zero, resulting in $f_{sky} = 9.8\%$.

A.1 Masking Results

We report the results for NILC using Gaussian isotropic, noise model with the 3σ masking scheme as described above. Figure A.2 shows a reduction in the foreground residuals after applying the 3σ threshold with respect to the 10% SO patch. The reduction in foreground residuals is mainly because we excluded the most contaminated regions in the sky with large angular scales. Thus by masking these areas, we minimized the contribution from the strong residual foregrounds

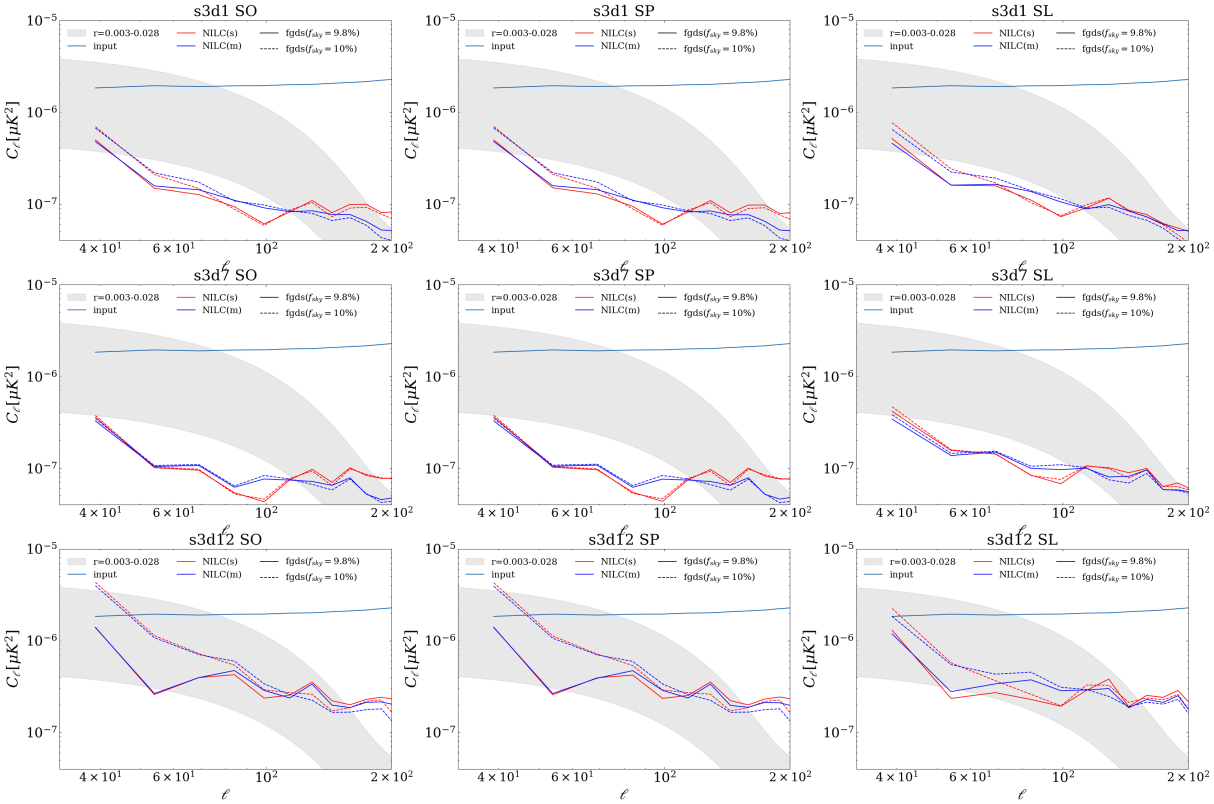


FIGURE A.2: B-modes angular power spectra averaged over 100 simulations for foreground residuals, respectively shown over $f_{sky} = 9.8\%$ (solid lines) and $f_{sky} = 10\%$ (dashed lines). Two needlet bases are considered i) Standard (red) and ii) Mexican (blue) binned with an interval of $\Delta\ell = 15$. The input CMB (with r=0) power spectrum averaged over 100 spectra, obtained after applying the same SO mask to 100 simulated CMB-only maps is shown with the light-blue solid line. The grey-shaded region indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$. Results are reported for three different foreground models: s3d1 (top), s3d7 (middle), and s3d12 (bottom) and three data sets: SO (left), SP (middle) and SL (right).

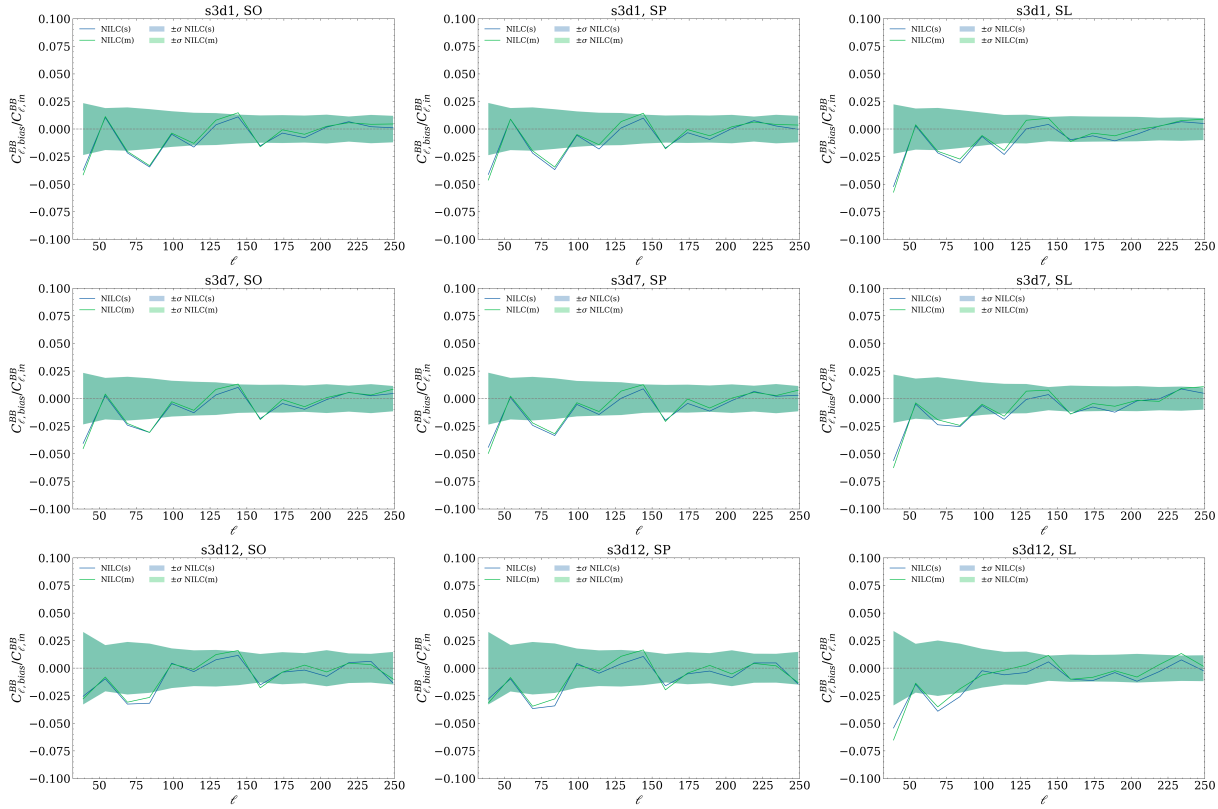


FIGURE A.3: NILC reconstruction bias $C^{BB}_{\ell,bias}$ is estimated with eq. (4.22) over $f_{sky} = 9.8\%$, whereas $C^{BB}_{\ell,in}$ is the averaged binned CMB B-mode angular power spectra estimated from CMB B-mode simulations on the same fraction of the sky. The uncertainty on the power spectrum of the average bias is estimated as $\frac{\sigma(C^{bias}_{\ell})}{\sqrt{N_{sims}}}$. Results are reported for three different foreground models: s3d1 (top), s3d7 (middle), and s3d12 (bottom) and three data sets: SO (left), SP (middle) and SL (right), and for both standard and mexican needlets

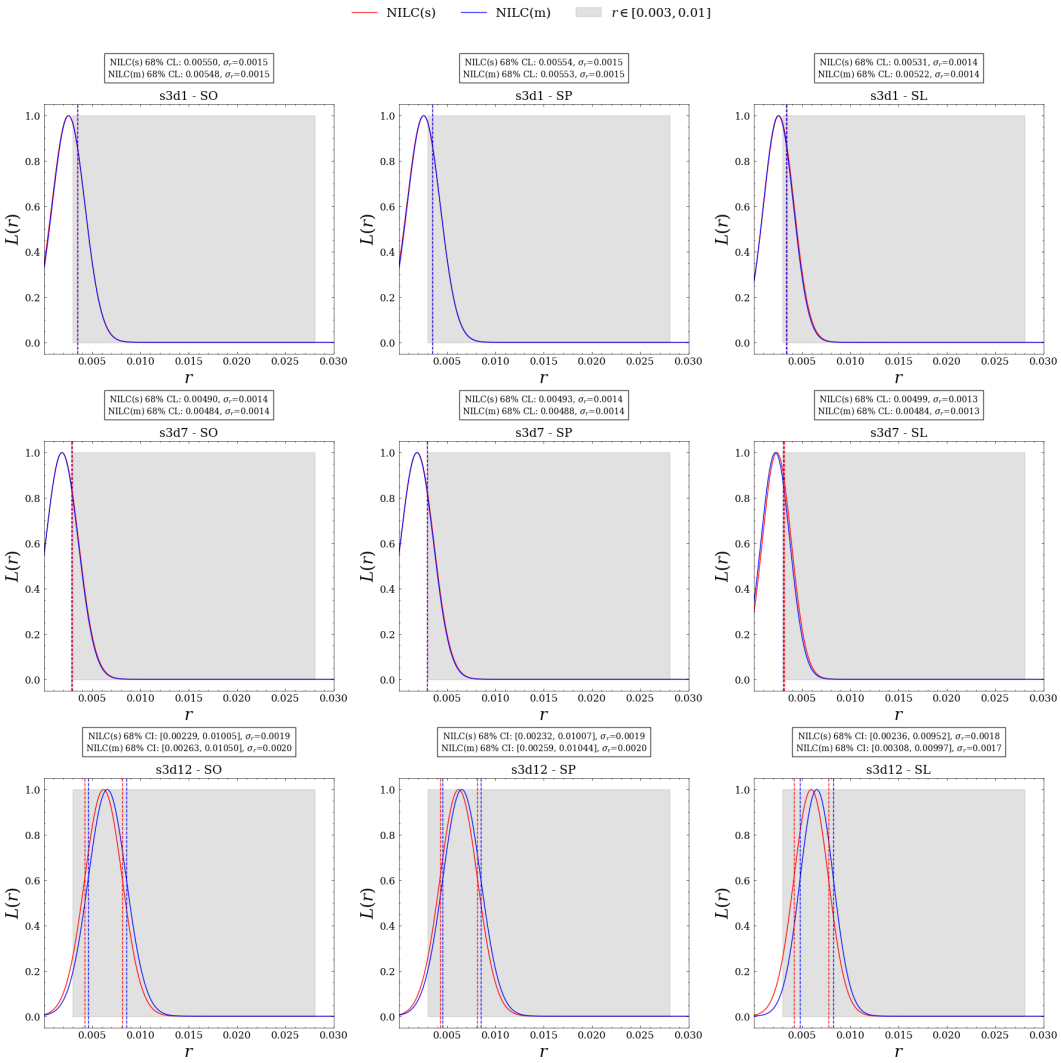


FIGURE A.4: $L(r)$ for Gaussian white noise on $f_{sky} = 10\%$ with $\Delta = 15\ell$. 68% CI for mexican (in blue) and standard (in red) are shown with dashed vertical lines. The grey shaded region indicates the amplitude range of the primordial tensor signal targeted by SO: $r \in [0.003, 0.028]$.

in the final analysis. However, the noise residuals remained consistent as they are largely uniform across the map and are not concentrated in the specific regions affected by the foregrounds. Therefore, masking the high-foreground areas had little impact on the overall noise level, as the noise is evenly distributed and does not correlate with the foreground structures in our ideal scenario. Posterior distribution with this making strategy is shown in Figure A.4, which shows lower 68% CL due to reduced foreground residuals

Bibliography

[1] A. G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009, 1998.

[2] S. Perlmutter et al. Measurements of omega and lambda from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565, 1999.

[3] F. Zwicky. The redshift of extragalactic nebulae. *Helvetica Physica Acta*, 6:110, 1933.

[4] V. C. Rubin and W. K. Jr. Ford. Rotation of the andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159:379, 1970.

[5] Planck Collaboration. Planck2018 results: Vi. cosmological parameters. *Astronomy amp; Astrophysics*, 641:A6, September 2020.

[6] P. J. E. Peebles. Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. *The Astrophysical Journal*, 263:L1–L5, 1982.

[7] C. L. Bennett et al. First-year wilkinson microwave anisotropy probe (wmap) observations: Preliminary maps and basic results. *The Astrophysical Journal Supplement Series*, 148(1):1, 2003.

[8] D. J. Eisenstein et al. Detection of the baryon acoustic peak in the large-scale correlation function of sdss luminous red galaxies. *The Astrophysical Journal*, 633(2):560, 2005.

[9] E. J. Copeland et al. Dynamics of dark energy. *International Journal of Modern Physics D*, 15(11):1753, 2006.

[10] G. Bertone et al. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405(5-6):279, 2005.

[11] A. G. Riess et al. Large magellanic cloud cepheid standards provide a 1% foundation for the determination of the hubble constant and stronger evidence for physics beyond cdm. *The Astrophysical Journal*, 876(1):85, 2019.

[12] B. Moore et al. Dark matter substructure within galactic halos. *The Astrophysical Journal*, 524(1):L19, 1999.

[13] A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23(2):347–356, 1981.

[14] A. D. Linde. A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108(6):389–393, 1982.

[15] Jerome Martin. The theory of inflation, 2018.

[16] S. W. Hawking. The development of irregularities in a single bubble inflationary universe. *Physics Letters B*, 115(4):295–297, September 1982.

[17] A. A. Starobinsky. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Physics Letters B*, 117(3-4):175–178, November 1982.

[18] A. H. Guth and S. Y. Pi. Fluctuations in the New Inflationary Universe. *Phys. Rev. Lett.*, 49(15):1110–1113, October 1982.

[19] Viatcheslav F. Mukhanov and G. V. Chibisov. Quantum Fluctuations and a Nonsingular Universe. *JETP Lett.*, 33:532–535, 1981.

[20] Simons Observatory Collaboration. The Simons Observatory: science goals and forecasts. *Journal of Cosmology and Astroparticle Physics*, 2019(2):056, February 2019.

[21] LiteBIRD Collaboration. LiteBIRD science goals and forecasts: improving sensitivity to inflationary gravitational waves with multitracer delensing. , 2024(6):010, June 2024.

[22] Jerome Martin. The theory of inflation, 2018.

[23] Giacomo Galloni, Nicola Bartolo, Sabino Matarrese, Marina Migliaccio, Angelo Ricciardone, and Nicola Vittorio. Updated constraints on amplitude and tilt of the tensor primordial spectrum. , 2023(4):062, April 2023.

[24] M. Tristram, A. J. Banday, K. M. Górski, R. Keskitalo, C. R. Lawrence, K. J. Andersen, R. B. Barreiro, J. Borrill, L. P. L. Colombo, H. K. Eriksen, R. Fernandez-Cobos, T. S. Kisner, E. Martínez-González, B. Partridge, D. Scott, T. L. Svalheim, and I. K. Wehus. Improved limits on the tensor-to-scalar ratio using BICEP and Planck data. *Phys. Rev. D*, 105(8):083524, April 2022.

[25] K. Wolz, S. Azzoni, C. Hervias-Caimapo, J. Errard, N. Krachmalnicoff, D. Alonso, C. Baccigalupi, A. Baleato Lizancos, M. L. Brown, E. Calabrese, J. Chluba, J. Dunkley, G. Fabbian, N. Galitzki, B. Jost, M. Morshed, and F. Nati. The simons observatory: pipeline comparison and validation for large-scale b-modes, 2023.

[26] LiteBIRD Collaboration. Probing cosmic inflation with the LiteBIRD cosmic microwave background polarization survey. *Progress of Theoretical and Experimental Physics*, 2023(4):042F01, April 2023.

[27] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. *The Astrophysical Journal*, 142:414–419, July 1965.

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

[29] G. Smoot et al. Structure in the COBE Differential Microwave Radiometer First-Year Maps. *The Astrophysical Journal Letters*, 396:L1, September 1992.

[30] Planck Collaboration. Planck 2013 results. XVI. Cosmological parameters. *Astronomy amp; Astrophysics*, 571:A16, November 2014.

[31] D. S. Swetz et al. Overview of the atacama cosmology telescope: Receiver, instrumentation, and telescope systems. *The Astrophysical Journal Supplement Series*, 194(2):41, June 2011.

[32] J. E. Carlstrom et al. The 10 meter south pole telescope. *Publications of the Astronomical Society of the Pacific*, 123(903):568–581, May 2011.

[33] G. F. Smoot, M. V. Gorenstein, and R. A. Muller. Detection of Anisotropy in the Cosmic Blackbody Radiation. *Phys. Rev. Lett.*, 39(14):898–901, October 1977.

[34] Andrew R. Liddle and David H. Lyth. *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press, Cambridge, 2000.

[35] George B Rybicki and Alan P Lightman. *Radiative Processes in Astrophysics*. Wiley, New York, NY, 1985.

[36] Paolo Cabella and Marc Kamionkowski. Theory of Cosmic Microwave Background Polarization. *arXiv e-prints*, pages astro–ph/0403392, March 2004.

[37] Planck Collaboration. Planck 2018 results. I. Overview and the cosmological legacy of Planck. *Astronomy amp; Astrophysics*, 641:A1, September 2020.

[38] D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *The Astrophysical Journal Supplement Series*, 148(1):175–194, September 2003.

[39] Montroy et al. A Measurement of the CMB <EE> Spectrum from the 2003 Flight of BOOMERANG. *The Astrophysical Journal*, 647(2):813–822, August 2006.

[40] Uros Seljak and Matias Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *The Astrophysical Journal*, 469:437, October 1996.

[41] Matias Zaldarriaga and Uroš Seljak. Gravitational lensing effect on cosmic microwave background polarization. *Physical Review D*, 58(2), June 1998.

[42] Planck Collaboration, N. Aghanim, Y. Akrami, and et al Ashdown. Planck 2018 results. VIII. Gravitational lensing. *Astronomy amp; Astrophysics*, 641:A8, September 2020.

[43] Planck Collaboration, P. A. R. Ade, N. Aghanim, and et al Ashdown. Planck intermediate results. XLI. A map of lensing-induced B-modes. *Astronomy amp; Astrophysics*, 596:A102, December 2016.

[44] S Adachi, MAO Aguilar Faúndez, K Arnold, C Baccigalupi, D Barron, D Beck, S Beckman, F Bianchini, David Boettger, J Borrill, et al. A measurement of the degree-scale cmb b-mode angular power spectrum with polarbear. *The Astrophysical Journal*, 897(1):55, 2020.

[45] D. Hanson, S. Hoover, et al. Detection of mode polarization in the cosmic microwave background with data from the south pole telescope. *Physical Review Letters*, 111(14), September 2013.

[46] et al Choi, Steve K. The atacama cosmology telescope: a measurement of the cosmic microwave background power spectra at 98 and 150 ghz. *Journal of Cosmology and Astroparticle Physics*, 2020(12):045–045, December 2020.

[47] Wei-Tou Ni. From equivalence principles to cosmology: Cosmic polarization rotation, cmb observation, neutrino number asymmetry, lorentz invariance and cpt. *Progress of Theoretical Physics Supplement*, 172:49–60, 2008.

[48] Christian Fidler, Guido W. Pettinari, Martin Beneke, Robert Crittenden, Kazuya Koyama, and David Wands. The intrinsic b-mode polarisation of the cosmic microwave background. *Journal of Cosmology and Astroparticle Physics*, 2014(07):011–011, July 2014.

[49] Andrew Mack, Tina Kahniashvili, and Arthur Kosowsky. Microwave background signatures of a primordial stochastic magnetic field. *Physical Review D*, 65(12), June 2002.

[50] T. R. Seshadri and Kandaswamy Subramanian. Cosmic microwave background polarization signals from tangled magnetic fields. *Physical Review Letters*, 87(10), August 2001.

[51] M. Tristram, Banday, et al. Planck constraints on the tensor-to-scalar ratio. *Astronomy amp; Astrophysics*, 647:A128, March 2021.

[52] CMB-S4 Collaboration. CMB-S4: Forecasting Constraints on Primordial Gravitational Waves. *The Astrophysical Journal*, 926(1):54, February 2022.

[53] Ragnhild Aurlien et al. Foreground separation and constraints on primordial gravitational waves with the PICO space mission. , 2023(6):034, June 2023.

[54] N. Krachmalnicoff, C. Baccigalupi, J. Aumont, M. Bersanelli, and A. Mennella. Characterization of foreground emission on degree angular scales for CMB B-mode observations . Thermal dust and synchrotron signal from Planck and WMAP data. *Astronomy amp; Astrophysics*, 588:A65, April 2016.

[55] C Dickinson et al. The State-of-Play of Anomalous Microwave Emission (AME) research. , 80:1–28, February 2018.

[56] D. Herman et al. Beyondplanck: Xv. limits on large-scale polarized anomalous microwave emission fromplancklfi and wmap. *Astronomy amp; Astrophysics*, 675:A15, June 2023.

[57] R. A. Sunyaev and Ya. B. Zeldovich. Small-Scale Fluctuations of Relic Radiation. , 7(1):3–19, April 1970.

[58] Planck Collaboration. Planck 2018 results. IV. Diffuse component separation. *Astronomy amp; Astrophysics*, 641:A4, September 2020.

[59] M. A. Miville-Deschênes, N. Ysard, A. Lavabre, et al. Separation of anomalous and synchrotron emissions using wmap polarization data. *Astronomy and Astrophysics*, 490:1093–1102, 2008.

[60] Matias Vidal, C. Dickinson, R. D. Davies, and J. P. Leahy. Polarized radio filaments outside the Galactic plane. , 452(1):656–675, September 2015.

[61] Planck Collaboration. Planck 2015 results. XXV. Diffuse low-frequency Galactic foregrounds. *amp; Astrophysics*, 594:A25, September 2016.

[62] N. Krachmalnicoff, E. Carretti, C. Baccigalupi, G. Bernardi, S. Brown, B. M. Gaensler, M. Haverkorn, M. Kesteven, F. Perrotta, S. Poppi, and L. Staveley-Smith. S-PASS view of polarized Galactic synchrotron at 2.3 GHz as a contaminant to CMB observations. *Astronomy amp; Astrophysics*, 618:A166, October 2018.

[63] C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. A 408 mhz all-sky continuum survey. ii - the atlas of contour maps. *Astronomy and Astrophysics Supplement Series*, 47:1, 1982.

[64] W. Reich and P. Reich. A radio continuum survey of the northern sky at 1420 mhz. ii. *Astronomy and Astrophysics Supplement Series*, 63:205, 1986.

[65] J. L. Jonas, E. E. Baart, and G. D. Nicolson. The rhodes/hartrao 2326-mhz radio continuum survey. *Monthly Notices of the Royal Astronomical Society*, 297:977, 1998.

[66] G. Giardino, A. J. Banday, P. Fosalba, et al. Towards a model of full-sky galactic synchrotron intensity and linear polarisation: A re-analysis of the parkes data. *Astronomy and Astrophysics*, 387:82–97, 2002.

[67] P. Platania, C. Burigana, D. Maino, et al. Full-sky total intensity and polarization maps of galactic synchrotron emission. *Astronomy and Astrophysics*, 410:847–860, 2003.

[68] U. Fuskeland, K. J. Andersen, R. Aurlien, R. Banerji, M. Brilenkov, H. K. Eriksen, M. Galloway, E. Gjerløw, S. K. Næss, T. L. Svalheim, and I. K. Wehus. Constraints on the spectral index of polarized synchrotron emission from WMAP and Faraday-corrected S-PASS data. *Astronomy amp; Astrophysics*, 646:A69, February 2021.

[69] E. Krugel. *The Physics of Interstellar Dust*. CRC Press, 1st edition, 2002.

[70] A. Lazarian. Tracing magnetic fields with aligned grains. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 106:225–256, July 2007.

[71] Planck Collaboration. Planck 2018 results. XI. Polarized dust foregrounds. *Astronomy amp; Astrophysics*, 641:A11, September 2020.

[72] Planck Collaboration. Planck intermediate results. XXXVIII. E- and B-modes of dust polarization from the magnetized filamentary structure of the interstellar medium. *Astronomy amp; Astrophysics*, 586:A141, February 2016.

[73] Planck Collaboration. Planck intermediate results. L. Evidence of spatial variation of the polarized thermal dust spectral energy distribution and implications for CMB B-mode analysis. *Astronomy amp; Astrophysics*, 599:A51, March 2017.

[74] B. Thorne, J. Dunkley, D. Alonso, and S. Næss. The Python Sky Model: software for simulating the Galactic microwave sky. , 469(3):2821–2833, August 2017.

[75] G. Puglisi, V. Galluzzi, L. Bonavera, J. Gonzalez-Nuevo, A. Lapi, M. Massardi, F. Perrotta, C. Baccigalupi, A. Celotti, and L. Danese. Forecasting the Contribution of Polarized Extragalactic Radio Sources in CMB Observations. *The Astrophysical Journal*, 858(2):85, May 2018.

[76] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *The Astrophysical Journal Supplement Series*, 208(2):20, October 2013.

[77] M.-A. Miville-Deschênes, N. Ysard, A. Lavabre, N. Ponthieu, J. F. Macías-Pérez, J. Aumont, and J. P. Bernard. Separation of anomalous and synchrotron emissions using wmap polarization data. *Astronomy amp; Astrophysics*, 490(3):1093–1102, September 2008.

[78] A. Kogut. Synchrotron Spectral Curvature from 22 MHz to 23 GHz. *The Astrophysical Journal*, 753(2):110, July 2012.

[79] Planck Collaboration. Planck2015 results: X. diffuse component separation: Foreground maps. *Astronomy amp; Astrophysics*, 594:A10, September 2016.

[80] Douglas P. Finkbeiner, Marc Davis, and David J. Schlegel. Extrapolation of Galactic Dust Emission at 100 Microns to Cosmic Microwave Background Radiation Frequencies Using FIRAS. *The Astrophysical Journal*, 524(2):867–886, October 1999.

[81] Brandon S. Hensley and B. T. Draine. Modeling the anomalous microwave emission with spinning nanoparticles: No pahs required. *The Astrophysical Journal*, 836(2):179, February 2017.

[82] Ginés Martínez-Solaeche, Ata Karakci, and Jacques Delabrouille. A 3d model of polarized dust emission in the milky way. *Monthly Notices of the Royal Astronomical Society*, 476(1):1310–1330, January 2018.

[83] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *The Astrophysical Journal*, 622(2):759–771, April 2005.

[84] G. F. Smoot et al. Structure in the cobe differential microwave radiometer first-year maps. *The Astrophysical Journal*, 396:L1–L5, 1992.

[85] J. A. Tauber et al. Planck pre-launch status: The optical system. *Astronomy Astrophysics*, 520:A1, 2010.

[86] H. K. Eriksen, C. Dickinson, C. R. Lawrence, C. Baccigalupi, A. J. Banday, K. M. Gorski, F. K. Hansen, P. B. Lilje, E. Pierpaoli, M. D. Seiffert, K. M. Smith, and K. Vanderlinde. Cosmic microwave background component separation by parameter estimation. *The Astrophysical Journal*, 641(2):665–682, April 2006.

[87] Radek Stompor, Samuel Leach, Federico Stivoli, and Carlo Baccigalupi. Maximum likelihood algorithm for parametric component separation in cosmic microwave background experiments. *Monthly Notices of the Royal Astronomical Society*, 392(1):216–232, January 2009.

[88] E. de la Hoz, P. Vielva, R. B. Barreiro, and E. Martínez-González. On the detection of cmb b-modes from ground at low frequency. *Journal of Cosmology and Astroparticle Physics*, 2020(06):006–006, June 2020.

[89] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. First-yearwilkinson microwave anisotropy probe(wmap) observations: Foreground emission. *The Astrophysical Journal Supplement Series*, 148(1):97–117, September 2003.

[90] Planck Collaboration. Planck 2013 results. XII. Diffuse component separation. *Astronomy amp; Astrophysics*, 571:A12, November 2014.

[91] M. Tegmark and G. Efstathiou. A method for subtracting foregrounds from multifrequency cmb sky maps. *Monthly Notices of the Royal Astronomical Society*, 281(4):1297–1314, August 1996.

[92] H. K. Eriksen, A. J. Banday, K. M. Gorski, and P. B. Lilje. On foreground removal from thewilkinson microwave anisotropy probedata by an internal linear combination method: Limitations and implications. *The Astrophysical Journal*, 612(2):633–646, September 2004.

[93] R. Vio and P. Andreani. A statistical analysis of the "internal linear combination" method in problems of signal separation as in cosmic microwave background observations. *Astronomy amp; Astrophysics*, 487(2):775–780, June 2008.

[94] Soumen Basak and Jacques Delabrouille. A needlet internal linear combination analysis of wmap 7-year data: estimation of cmb temperature map and power spectrum: A nilc analysis of wmap 7-year data. *Monthly Notices of the Royal Astronomical Society*, 419(2):1163–1175, November 2011.

[95] J. Delabrouille and J.-F. Cardoso. Diffuse component separation in cmb observations. In *Proceedings of the XXth Rencontres de Blois*, 2009.

[96] C. Baccigalupi, F. Perrotta, G. de Zotti, G. F. Smoot, C. Burigana, D. Maino, L. Bedini, and E. Salerno. Extracting cosmic microwave background polarization from satellite astrophysical maps. , 354(1):55–70, October 2004.

[97] D. Maino, A. Farusi, C. Baccigalupi, F. Perrotta, A. J. Banday, L. Bedini, C. Burigana, G. De Zotti, K. M. Górski, and E. Salerno. All-sky astrophysical component separation with Fast Independent Component Analysis (FASTICA). , 334(1):53–68, July 2002.

[98] D. Maino, S. Donzelli, A. J. Banday, F. Stivoli, and C. Baccigalupi. Cosmic microwave background signal in Wilkinson Microwave Anisotropy Probe three-year data with FASTICA. , 374(4):1207–1215, February 2007.

[99] Jean-Francois Cardoso, Maude Martin, Jacques Delabrouille, Marc Betoule, and Guillaume Patanchon. Component separation with flexible models. Application to the separation of astrophysical emissions. *arXiv e-prints*, page arXiv:0803.1814, March 2008.

[100] H. K. Eriksen, J. B. Jewell, C. Dickinson, A. J. Banday, K. M. Górski, and C. R. Lawrence. Joint Bayesian Component Separation and CMB Power Spectrum Estimation. *The Astrophysical Journal*, 676(1):10–32, March 2008.

[101] M. P. Hobson, A. W. Jones, A. N. Lasenby, and F. R. Bouchet. Foreground separation methods for satellite observations of the cosmic microwave background. , 300(1):1–29, October 1998.

[102] Max Tegmark, Angélica de Oliveira-Costa, and Andrew J. Hamilton. High resolution foreground cleaned CMB map from WMAP. *Phys. Rev. D*, 68(12):123523, December 2003.

[103] Francis Narcowich, P. Petrushev, and Joseph Ward. Localized tight frames on spheres. *SIAM J. Math. Analysis*, 38:574–594, 01 2006.

[104] P. Baldi, G. Kerkyacharian, D. Marinucci, and D. Picard. Asymptotics for spherical needlets. *arXiv Mathematics e-prints*, page math/0606599, June 2006.

[105] D. Marinucci, D. Pietrobon, A. Balbi, P. Baldi, P. Cabella, G. Kerkyacharian, P. Natoli, D. Picard, and N. Vittorio. Spherical needlets for cosmic microwave background data analysis. , 383(2):539–545, January 2008.

[106] J. Delabrouille, J.-F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, and F. Guilloux. A full sky, low foreground, high resolution cmb map from wmap. *Astronomy amp; Astrophysics*, 493(3):835–857, November 2008.

[107] David Geller and Azita Mayeli. Continuous wavelets on compact manifolds. *Mathematische Zeitschrift*, 262(4):895–927, 2009.

[108] S. Basak and J. Delabrouille. A needlet ILC analysis of WMAP 9-year polarization data: CMB polarization power spectra. *Monthly Notices of the Royal Astronomical Society*, 435(1):18–29, aug 2013.

[109] J. Delabrouille, J.-F. Cardoso, M. Le Jeune, et al. The pre-launch planck sky model: a model of sky emission at submillimetre to centimetre wavelengths. *Astronomy and Astrophysics*, 493:835–857, 2009.

[110] Alessandro Carones, Marina Migliaccio, Domenico Marinucci, and Nicola Vittorio. Analysis of needlet internal linear combination performance on b-mode data from sub-orbital experiments. *Astronomy amp; Astrophysics*, 677:A147, September 2023.

[111] Alessandro Carones, Marina Migliaccio, Giuseppe Puglisi, Carlo Baccigalupi, Domenico Marinucci, Nicola Vittorio, and Davide Poletti. Multiclustering needlet ilc for cmb b-mode component separation. *Monthly Notices of the Royal Astronomical Society*, 525(2):3117–3135, August 2023.

[112] Samira Hamimeche and Antony Lewis. Likelihood analysis of cmb temperature and polarization power spectra. *Physical Review D*, 77(10), May 2008.

[113] J. R. Bond, A. H. Jaffe, and L. Knox. Estimating the power spectrum of the cosmic microwave background. *Phys. Rev. D*, 57(4):2117–2137, February 1998.

[114] Eric Hivon, Krzysztof M. Górski, C. Barth Netterfield, Brendan P. Crill, Simon Prunet, and Frode Hansen. MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets. *The Astrophysical Journal*, 567(1):2–17, March 2002.

[115] G. Polenta, D. Marinucci, A. Balbi, P. de Bernardis, E. Hivon, S. Masi, P. Natoli, and N. Vittorio. Unbiased estimation of an angular power spectrum. , 2005(11):001, November 2005.

[116] David Alonso, Javier Sanchez, and Anže Slosar. A unified pseudo-c framework. *Monthly Notices of the Royal Astronomical Society*, 484(3):4127–4151, January 2019.