Instituto de Física de Cantabria

Consejo Superior de Investigaciones Científicas

Departamento de Física Moderna

Universidad de Cantabria

Exploring statistical isotropy in *Planck* Cosmic Microwave Background temperature and polarisation data

Explorando la isotropía estadística con los datos de temperatura y polarización del Fondo Cósmico de Microondas de *Planck*

Memoria presentada para optar al título de Doctor en Ciencia y Tecnología por la Universidad de Cantabria

por

Christian Gimeno Amo

Bajo la supervisión de

Enrique Martínez González y Rita Belén Barreiro Vilas

Enrique Martínez González, Doctor en Ciencias Físicas y Profesor de Investigación del Consejo Superior de Investigaciones Científicas

y

Rita Belén Barreiro Vilas, Doctora en Ciencias e Investigadora Científica del Consejo Superior de Investigaciones Científicas

CERTIFICAN que la presente memoria

Explorando la isotropía estadística con los datos de
temperature y polarización del Fondo Cósmico de
Microondas de *Planck*

ha sido realizada, bajo su dirección, por Christian Gimeno Amo, y constituye su Tesis para optar al grado de Doctor por la Universidad de Cantabria. Asimismo, emiten su conformidad para que la presente Memoria sea depositada y se celebre, ulteriormente, la correspondiente Lectura y Defensa.

En Santander, a 21 de Julio de 2025,

Fdo.: Prof. Enrique Martínez González   Fdo.: Dra. Rita Belén Barreiro Vilas

Agradecimientos

Qué mejor forma de empezar esta tesis que agradeciendo a todas las personas que me han apoyado durante esta aventura de cuatro años. Sin vosotros, esto no habría sido posible.

Primero, a mis dos directores, Belén y Kike. Gracias por haberme guiado durante este tiempo. Puedo decir con total seguridad que he aprendido mucho con vuestras enseñanzas y consejos, y que he ganado más confianza en mí mismo. Esta tesis tampoco habría sido posible sin la ayuda de dos investigadores que han sido unos referentes para mí: Frode K. Hansen y Anthony Banday. Muchas gracias.

También me gustaría agradecer a la gente del IFCA, y en especial a mis compañeros de despacho: Guille, el primer amigo que hice en Santander y mi compi de juegos de mesa y pádel, espero que te vaya todo bien por Japón; Miguel, cuántos viajes al *Lupa* y risas en el despacho habremos compartido; y Felice, mi italiano favorito. Sois increíbles, nunca cambiéis.

Cómo olvidarme de Guillermo, Celia, Rubén, Patricia, Mikel y Elena. Si ese comedor pudiera hablar... mejor que *holdee*. Y, por último, pero no menos importante, mis compis de piso: Chema y Sergio. Ha sido un verdadero placer conoceros y convivir con vosotros. Me quedo con un par de anécdotas: como aquella vez que casi perdemos el taxi para ir a Tenerife, o lo “agradable y pacífica" que era nuestra calle.

A mis amigos de Oriñon: Lantxo, Bentx, Gegun, Rou, Perez, Gue, Jony, Amoriza, Gari, Ane, Olatz eta Amaia. No sé cómo, pero año tras año os superáis. No recuerdo un verano sin vosotros. ¡Por muchas más fiestas juntos!

Por último, a mis padres, a mi yaya y a mi tita, mis pilares. ¿Qué puedo decir? Siempre han estado ahí, en los buenos y malos momentos, y soy muy afortunado por ello. Cuando el síndrome del impostor más daño hacía, ellos me recordaban todo lo que he conseguido hasta ahora por mí mismo... y lo que queda. Os quiero.

No puedo olvidarme de la personita que me ha acompañado este último año y medio: Penélope. De forma indirecta, pero la tesis me ha permitido conocerte. Estos últimos meses

has sido mi mayor motivación para acabar de redactar este documento. Gracias por aguantarme, que no es poco. Estaré ahí para lo que necesites. Tampoco me olvido de Chary y Luis. Gracias por portaros tan bien conmigo.

Abstract

> *"Somewhere, something incredible is waiting to be known."*
>
> Sharon Begley

The Cosmic Microwave Background (CMB) constitutes one of the most valuable sources of information for understanding the Universe. Among others, measurements from the COBE, WMAP, and *Planck* missions have enabled the establishment of a cosmological model in which the Universe is consistent with Euclidean geometry, is undergoing accelerated expansion, and is composed of baryonic matter, dark matter, and dark energy. Furthermore, the density perturbations that gave rise to the large-scale structure observed today are compatible with purely adiabatic perturbations with an almost scale-invariant power spectrum. All of this is encapsulated in the ΛCDM model, a six-parameter framework that provides an excellent fit to both CMB and large-scale structure observations.

Despite its remarkable success, the model exhibits some tensions, among which are the so-called CMB anomalies. These include a lack of power and angular correlation on large angular scales, a hemispherical asymmetry in the distribution of power, an anomalous alignment of the lowest multipoles, and the presence of a “cold spot" in the southern hemisphere with an anomalous amplitude and curvature. These deviations have been identified in the temperature data with individual significances that, although not conclusive, motivate the exploration of explanations beyond the hypothesis that they are statistical fluctuations within the standard model.

Polarisation data could provide nearly independent information. However, current measurements are limited by large-scale systematic effects and a low signal-to-noise ratio, which prevent the extraction of robust conclusions. The forthcoming high-precision polarisation measurements, such as those expected from LiteBIRD, will be crucial for determining whether these unexpected features are indicative of new physics.

This PhD thesis presents my three contributions to the study and characterisation of some of these anomalies. Chapter 4 focuses on a comprehensive analysis of the hemispherical power asymmetry (HPA) in both temperature and polarisation E-mode data from the most recent *Planck* data release. An HPA is confirmed in temperature with a p-value below 1%, while the E-modes exhibit a potential hint of asymmetry at the $1\% - 3\%$ level, with an axis closely aligned with that found in temperature. Chapter 5 is dedicated to the development of an image-processing technique for filling masked regions of the sky with a realisation statistically compatible with the cosmological model and the observed data. This method is implemented in a publicly available software package, `CMB-PAInT`, which offers a valuable alternative for studies using masked skies. Finally, Chapter 6 includes two complementary analyses. The first one investigates the directional dependence of the temperature angular power spectrum by fitting a dipole to the spatial distribution of power in each bandpower. The resulting directions exhibit anomalous clustering around an axis that closely matches the HPA direction, an effect not observed in the polarisation data. The second analysis explores the spatial distribution of cosmological parameters, finding consistency with ΛCDM expectations for most parameters, except for A_s. Prior to these contributions, Chapters 1 and 2 provide an overview of modern cosmology and the CMB, while Chapter 3 offers a review of the current state of the art of the CMB anomalies. Finally, Chapter 7 (or 8 for Spanish readers) wraps up with the general conclusions of my PhD thesis.

Resumen

> *"Somos una forma que tiene el cosmos de conocerse a sí mismo."*
>
> Carl Sagan

El fondo cósmico de microondas (FCM) constituye una de las principales fuentes de información para conocer el Universo en el que vivimos. Las mediciones realizadas, entre otras, por las misiones COBE, WMAP, y *Planck* han permitido establecer un modelo cosmológico en el que el Universo es compatible con una geometría euclídea, se encuentra en expansión acelerada, y está compuesto por materia bariónica, materia oscura, y energía oscura. Además, sabemos que las perturbaciones de densidad que dieron origen a las estructuras cósmicas que observamos son compatibles con perturbaciones puramente adiabáticas con un espectro casi invariante de escala. Todo ello viene recogido en el modelo ΛCDM, un modelo descrito únicamente por seis parámetros, y que proporciona un ajuste notablemente preciso a los datos del FCM y de la estructura a gran escala.

Sin embargo, a pesar de su éxito, el modelo presenta ciertas tensiones entre las que se encuentran las denominadas anomalías del FCM. Estas incluyen un déficit de potencia y correlación angular escalas grandes, una asimetría hemisférica en la distribución de potencia, una alineación anómala de los multipolos más bajos, y la presencia de una mancha fría en el hemisferio sur con una amplitud y curvatura anómalamente elevadas. Estas desviaciones han sido observadas en los datos de temperatura con significancias individuales que, aunque no concluyentes, motivan la exploración de explicaciones más allá de la hipótesis de que se trata de meras fluctuaciones estadísticas dentro del modelo estándar.

El estudio de estas anomalías en los datos de polarización podría aportar información prácticamente independiente sobre su origen. No obstante, las mediciones actuales presentan limitaciones debido a la presencia de efectos sistemáticos a gran escala y al bajo nivel de señal-ruido, lo que impide sacar conclusiones robustas. Es por ello, que los datos de

polarización de alta precisión que obtendremos en la próxima decada, gracias a misiones como LiteBIRD, serán cruciales para dar respuesta a estas cuestiones.

Esta tesis doctoral recoge mis tres contribuciones al estudio y caracterización de algunas de estas anomalías. En particular, el capítulo 4 se centra en un análisis exhaustivo de la asimetría hemisférica de potencia (HPA, por sus siglas en inglés) en los datos de temperatura y los modos E de polarización del último procesado de datos de *Planck*. Se confirma la presencia de una HPA en temperatura con un *p-value* inferior al 1%, mientras que en los modos E de polarización se observa un posible indicio, con una significancia entre el 1% y el 3%, y un eje de asimetría muy próximo al identificado en temperatura. El capítulo 5 está dedicado al desarrollo de una técnica de procesado de imagen que permite rellenar los datos enmascarados con una realización estadísticamente compatible con el modelo cosmológico y los datos observados. Se presenta el software de acceso público `CMB-PAInT`, una herramienta eficaz para estudios que utilizan regiones del cielo enmascaradas. Por último, el capítulo 6 incluye dos análisis complementarios. En el primero, se estudia la distribución espacial de potencia en cada *bandpower* del espectro angular de potencias, ajustando un dipolo sobre cada uno de ellos y cuantificando el agrupamiento de las direcciones resultantes. Este análisis aporta evidencia de una dependencia direccional del espectro angular de potencias de temperatura, con un alineamiento anómalo de las direcciones en torno a un eje próximo al de la HPA, mientras que los datos de polarización no muestran nada similar. El segundo análisis examina la distribución espacial de los parámetros cosmológicos. Se observa que la mayoría de parámetros son compatibles con las expectativas del modelo ΛCDM, con la excepción de A_s. Previamente a estas contribuciones, los capítulos 1 y 2 ofrecen una visión general de la cosmología y el FCM, mientras que el capítulo 3 contiene un breve repaso a las diferentes anomalías. Finalmente, los capítulos 7 y 8 contienen las conclusiones generales de la tesis en inglés y español, respectivamente.

Contents

Part I

Introduction

1

An Introduction to the Universe

> *"Evolution simply hasn't equipped the human brain with the ability to have any intuition for the vastness of the cosmos."*
>
> Daniel Baumann

Over the past century, cosmology has transformed from a largely philosophical endeavour into a precise, data-driven science. At the heart of this transformation lies the *standard model* of cosmology, the ΛCDM model, which provides a remarkably successful framework for describing the evolution of the Universe. This model rests on a set of key observational pillars: the expansion of the Universe, the precise measure of the temperature and polarisation anisotropies in the cosmic microwave background (CMB), the distribution of galaxies and large-scale structure, and the abundance of light elements as predicted by Big Bang nucleosynthesis. Each of these observations has contributed to the emergence of a coherent cosmological picture, one that includes a mysterious dark energy component, an unknown dark matter component, and an inflationary paradigm. The road to this understanding has been paved by a century of scientific breakthroughs, from Einstein's formulation of general relativity to the discovery of cosmic expansion and the detection of CMB anisotropies. This chapter offers an overview of our current understanding of the Universe and a historical recap of the key developments that have led us here.

At the beginning of the 20th century, a series of key theoretical and observational developments led by distinguished physicists such as Albert Einstein, George Lemaître, Alexander Friedmann, and Edwin Hubble, gave rise to what we know as modern cosmology. This young scientific discipline aims to understand the origin and evolution of the Universe using the fundamental laws of physics that include gravity and particle physics.The discovery of General Relativity by Albert Einstein in 1916 (Einstein, 1916) was one of the most important milestones as it provided for the first time a theoretical framework for describing gravity and the Universe. It is a geometric theory of gravity that arose from the incompatibility between Newton's classical theory, where gravity is an instantaneous force, and the special relativity (Einstein, 1905). Einstein proposed that gravity is no longer a force, but rather a consequence of the curvature of the spacetime. General Relativity relates the geometry of the spacetime to the constituents of the universe through Einstein's equations, which are summarised as the following tensor equality in natural units:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{1.1}$$

The left side of the equation encapsulates the geometric properties of spacetime through the metric, $g_{\mu\nu}$, and the Einstein tensor, $G_{\mu\nu}$, while the right side encodes information about the matter and energy content via the stress-energy tensor, $T_{\mu\nu}$. G is Newton's constant, and Λ is the famous cosmological constant, originally introduced by Einstein to balance the attractive force of gravity and achieve a static universe, in line with the observational evidence available at the time.

Soon after the discovery of the General Relativity, Alexander Friedmann (Friedman, 1922) found a solution to Einstein's equations describing a homogeneous, isotropic and expanding universe. He demonstrated for the first time that the General Relativity equations admit non-static solutions. The metric describing this universe is the Friedmann–Lemaître–Robertson–Walker (FLRW) metric[1] given in polar coordinates as

$$ds^2 = -dt^2 + a(t)^2 \left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right], \tag{1.2}$$

where a(t) is the scale factor, a dimensionless quantity that describes the expansion of the universe. It relates the comoving[2] and physical distances. It is set to be one at present time, so both distances match. Additionally, k is the intrinsic spatial curvature, which could be zero ($k = 0$), positive ($k = 1$) or negative ($k = -1$). The first case corresponds to the three-dimensional Euclidean space, in which two parallel lines do not intersect. The second case represents a space where two parallel lines eventually meet. In the last case, parallel lines diverge.

---

[1] Here, I follow the (-, +, +, +) signature for the metric.

[2] Comoving coordinates define a reference frame that remains fixed with respect to the expanding universe, as they do not change with time.

Sometimes, it is convenient to write the metric in terms of comoving coordinates. The line element takes this form,

$$ds^2 = a^2(\eta)[-d\eta^2 + d\chi^2 + S_k^2(\eta)d\Omega^2], \tag{1.3}$$

where the conformal time, $d\eta = dt/a(t)$, and the comoving radial distance, $d\chi = dr/\sqrt{1-kr^2}$, have been defined. $S_k(\chi)$ is the function relating comoving distances to angles subtended on the sky, and it depends on the geometry of the Universe,

$$S_k(\chi) = \begin{cases} R_0 \sin(\chi/R_0) & \text{for } k > 0, \text{ closed,} \\ \chi & \text{for } k = 0, \text{ flat,} \\ R_0 \sinh(\chi/R_0) & \text{for } k < 0, \text{ open.} \end{cases} \tag{1.4}$$

Note that, for a flat Universe, r and χ are equivalent, so $r_{\text{phys}} = a(t)r$. In the general case, we have $r_{\text{phys}} = a(t)\chi$.

In 1927 George Lemaître arrived independently to the same solutions (Lemaître, 1927) and predicted what would later be known as the Hubble-Lemaître law, showing that galaxies appear to recede from us at speeds proportional to their distances

$$v \approx H_0 d. \tag{1.5}$$

He also estimated the proportionality constant, the Hubble constant, two years before Edwin Hubble measured it using Vesto Slipher's galaxy radial velocities and distance measurements based on Cepheids[3]. Equation 1.5 corresponds to the Hubble flow, which is the velocity of the galaxy resulting from the expansion of space. Furthermore, Lemaître was the first to propose the Big Bang theory. If the universe is expanding, tracing its age backward eventually leads to the initial singularity what he called "primeval atom". In the 1940s and 1950s, George Gamow, along with Ralph Alpher and Robert Hearman, made significant contributions to the Big Bang theory (Alpher et al., 1948; Gamow, 1948). They proposed a mechanism for the formation of the light elements in the early universe, what is now known as Big Bang Nucleosynthesis (BBN), and predicted the existence of an old relic radiation, the Cosmic Microwave Background (CMB). In 1965 Arno Penzias and Robert Wilson detected an excess of 4K antenna temperature in their instrument (Penzias & Wilson, 1965), which was later determined to be due to the CMB (Dicke et al., 1965). This discovery was the definitive evidence for the Big Bang theory.

Coming back to the FLRW metric, when combined with Einstein's field equations the

---

[3] Henrietta Swan Leavitt discovered the period-luminosity relationship for Cepheid variables.

metric gives the two Friedmann equations,

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \tag{1.6}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P), \tag{1.7}$$

where $\rho = \rho_r + \rho_m + \rho_\Lambda$, is the energy density and accounts for all the radiation, matter, and cosmological constant[4] contributions. In order to derive these equations, the following energy-momentum tensor has been assumed

$$T_{\mu\nu} = (\rho + P)\, u_\mu u_\nu + P g_{\mu\nu}, \tag{1.8}$$

where u^μ is fluid's four-velocity relative to a comoving observer. This is the energy-momentum tensor of a perfect fluid, which is completely characterised by its rest frame energy density, ρ, and pressure, P. It arises from the homogeneous and isotropic assumptions that restrict the possible values of the energy-momentum tensor elements. Apart from the two Friedman equations, a third equation[5] can be derived from the relativistic version of the continuity and Euler equations, which determines how ρ and P evolve over time, ensuring that energy and momentum are conserved,

$$\nabla_\mu T^\mu_\nu = \frac{\partial T^\mu_\nu}{\partial x_\mu} + \Gamma^\mu_{\mu\lambda} T^\lambda_\nu - \Gamma^\lambda_{\mu\nu} T^\mu_\lambda = 0. \tag{1.9}$$

If we focus on the conservation equations for the energy density ($\nu = 0$)

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0. \tag{1.10}$$

Considering that most cosmological fluids can be described by a constant equation of state, $\omega = P/\rho$, Eq. 1.10 can be rewritten as

$$\frac{\dot{\rho}}{\rho} = -3(1+\omega)\frac{\dot{a}}{a}, \tag{1.11}$$

which has the following solutions that depends on the equation of state,

$$\rho \propto a^{-3(1+\omega)}. \tag{1.12}$$

This provides information on the scaling of different fluids with the expansion. Depending on their equation of state, we distinguish three types of fluids: matter, radiation and dark energy.

---

[4] Note that the Λ term in Eq. 1.1 can be moved to the right-hand side.

[5] Only two of them are needed to determine the dynamics of the universe.

- **Matter** refers to a gas of non-relativistic particles which has effectively zero pressure, i.e. its pressure is much smaller than its energy density. Setting $\omega = 0$ in Eq. 1.12 gives $\rho \propto a^{-3}$. This reflects the fact that energy within a given volume stays constant, while the density dilutes as the volume expands like $V \propto a^3$. The "matter" term refers to ordinary and dark matter (DM). Fritz Zwicky first coined this term in 1933 after observing a discrepancy in the Coma cluster between the gravitational mass inferred from the virial theorem and the much lower mass estimated from its brightness, finding the former to be at least 400 times greater (Zwicky, 1933). He was forced to introduce an extra form of matter named "dunkle materie", i.e., dark matter. Further evidence for the existence of a non-visible matter came in 1970s when Vera Rubin and collaborators measured the rotation curve of spiral galaxies (Rubin & Ford, 1970). Contrary to the expected decreasing[6], the rotation curves flatten, which could only be explained if galaxies were embedded in halos of DM. Today, the clearest evidence for DM comes from the gravitational lensing of the CMB (Planck Collaboration et al., 2020c).

- **Radiation** refers to a gas of relativistic particles for which the energy density is dominated by the kinetic energy. It can be shown that radiation has $P = \rho/3$, which gives $\rho \propto a^{-4}$. In an expanding universe, the energy density of radiation decreases faster than the inverse of the volume, because the energy of each photon is also redshifted like $E \propto a^{-1}$. The redshift is defined as the fractional shift in the wavelength,

  $$z \equiv \frac{\lambda_0 - \lambda_1}{\lambda_1}, \tag{1.13}$$

  where λ_0 is the observed wavelength and λ_1 is the emitted one. Redshift originated by the expansion of the universe is also called the cosmological redshift. It occurs because the space itself is expanding and stretching the wavelength of light travelling through it. The redshift can also be expressed in terms of the scale factor,

  $$1 + z = \frac{1}{a(t_1)}. \tag{1.14}$$

  This implies that if we observe a galaxy at redshift $z = 1$, we are detecting light emitted when the size of the Universe was half its current size. Except for photons, which are relativistic as they are massless, the inventory of relativistic particles changes throughout the history of the universe. At early times, all the particles were relativistic as the temperature was so high that the rest mass of all of them was negligible. Later on, the temperature dropped below the masses of many particles and they started to behave like non-relativistic matter. Neutrinos, on the other hand, had remained relativistic

---

[6] For a spherically symmetric mass distribution where all the mass is concentrated near the centre, the orbital velocity in the outer regions is expected to decrease proportionally to $r^{-1/2}$.

until late times.

- **Dark energy** At the end of the twentieth century, studies of Supernovae Type Ia found that the Universe was expanding in an accelerated way (Riess et al., 1998; Perlmutter et al., 1999). This was an unexpected result for a matter-dominated Universe, where a cosmic deceleration is expected. To explain this, a fluid with negative pressure is required. As a consequence, the cosmological constant, Λ, which had been discarded many years before following the observations that the Universe is not static, was reintroduced. With ω = -1, the energy density remains constant, $\rho_\Lambda = \Lambda/(8\pi G)$. The vacuum energy associated to the empty space is a natural candidate for a fluid with a constant energy density and it is predicted by the quantum field theory. Unfortunately, the value predicted by the theory is much larger than the observed value. The two values differ by 120 orders of magnitude (Adler et al., 1995). Despite the term dark energy is often used to refer to the cosmological constant, it describes a more general fluid whose equation of state could be different from ω = -1, or even it can vary in time, as suggested by recent observations (Adame et al., 2025; DESI Collaboration et al., 2025).

Figure 1.1 shows the evolution of the energy density for each component. The Universe started with a radiation-dominated (RD) epoch. Then, at some point, the matter and radiation densities equalled, a moment referred to as matter-radiation equality, which occurred roughly at $z_{\rm eq} \approx 3400$, when the Universe was about $50,000$ years old. From that point onward, the energy density became dominated by matter (MD). This lasts until $z \approx 0.3$ ($t_{\Lambda\rm D} \approx 10.3$ Gyr), when the cosmological constant started to dominate over the other components (ΛD). In general, it is more convenient to rewrite the densities in dimensionless form and work with the following density parameters,

$$\Omega_{i,0} \equiv \frac{\rho_{i,0}}{\rho_{\rm crit,0}}, \tag{1.15}$$

where subscript "i" refers to different components, and $\rho_{\rm crit}$ is the critical density, which corresponds to the energy density required for the Universe to be spatially flat. The subscript '0' indicates that the quantities refer to their present-day values. From Eq. 1.6 and substituting H with its present-day value, the Hubble constant,

$$k = 0\ ;\ H_0 = 100h \text{ km s}^{-1}\text{Mpc}^{-1} \ \rightarrow\ \rho_{\rm crit,0} = \frac{3H_0^2}{8\pi G} = 1.9 \times 10^{-29}\, h^2\text{g cm}^{-3}. \tag{1.16}$$

Using dimensionless densities, Eq. 1.6 can be rewritten as

$$H^2 = H_0^2[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda], \tag{1.17}$$

where the curvature "density" parameter is also included, $\Omega_k \equiv -k/H_0^2$. Evaluating this

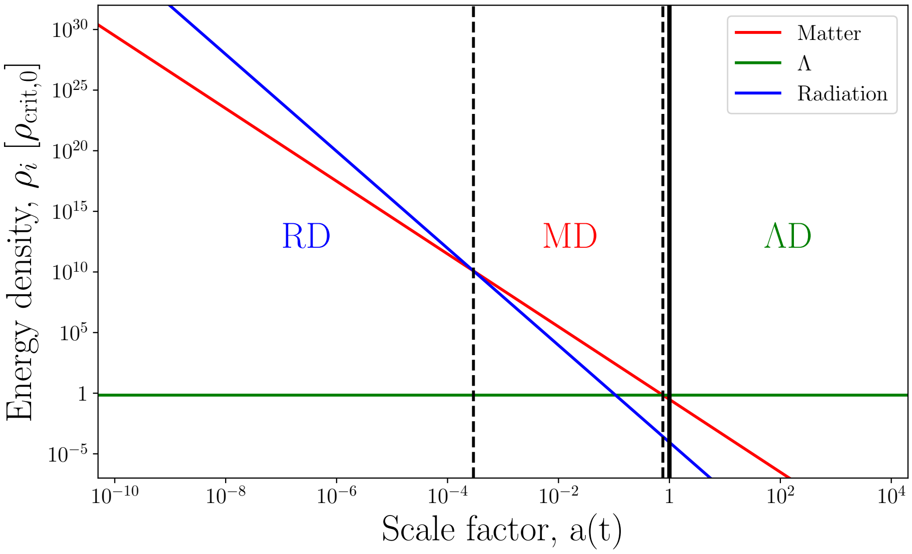

FIGURE 1.1: Evolution of the radiation (blue), dark energy (green), and matter (red) energy densities, in units of critical density, during Universe history. The solid vertical line indicates the present time, while the dashed lines mark the matter–radiation equality ($a_{\mathrm{eq}} \approx 2.9 \times 10^{-4}$) and the matter–dark energy equality ($a_{\Lambda\mathrm{D}} \approx 0.76$). This is computed for $\Omega_m = 0.3071$, $\Omega_r = 9 \times 10^{-5}$, $\Omega_k = 0$, and $\Omega_\Lambda = 0.693$ (Tristram et al., 2024).

equation at the present time we get

$$1 = \Omega_r + \Omega_m + \Omega_\Lambda + \Omega_k = \Omega_0 + \Omega_k, \tag{1.18}$$

from where it is straightforward to see that when the curvature is zero, the density equals the critical density, $\sum_i \rho_{i,0} = \rho_{\mathrm{crit},0}$.

The discovery of the accelerated Universe and dark energy helped to solve some tensions in cosmology. In particular, the matter-only model predicted an age for the Universe that was shorter than the ages of the oldest observed stars. Moreover, observations of the large-scale structure showed that the total matter content accounted for only about 30% of the critical density, which was in tension with strong indications that the Universe is nearly spatially flat. Including dark energy not only solved the age discrepancy, but also brought the total energy density in line with a flat geometry.

In general, determining how the scale factor, a(t), evolves in time requires solving numerically the Friedmann equation 1.17, which is a complicated non-linear differential equation. Figure 1.2 shows the expansion history for *Planck* cosmology. However, if we assume the single component case, we can solve for it analytically.

$$a(t) \propto \begin{cases} t^{2/3} & \mathrm{MD} \\ t^{1/2} & \mathrm{RD} \\ e^{H_0\sqrt{\Omega_\Lambda}t} & \Lambda\mathrm{D} \end{cases} \tag{1.19}$$

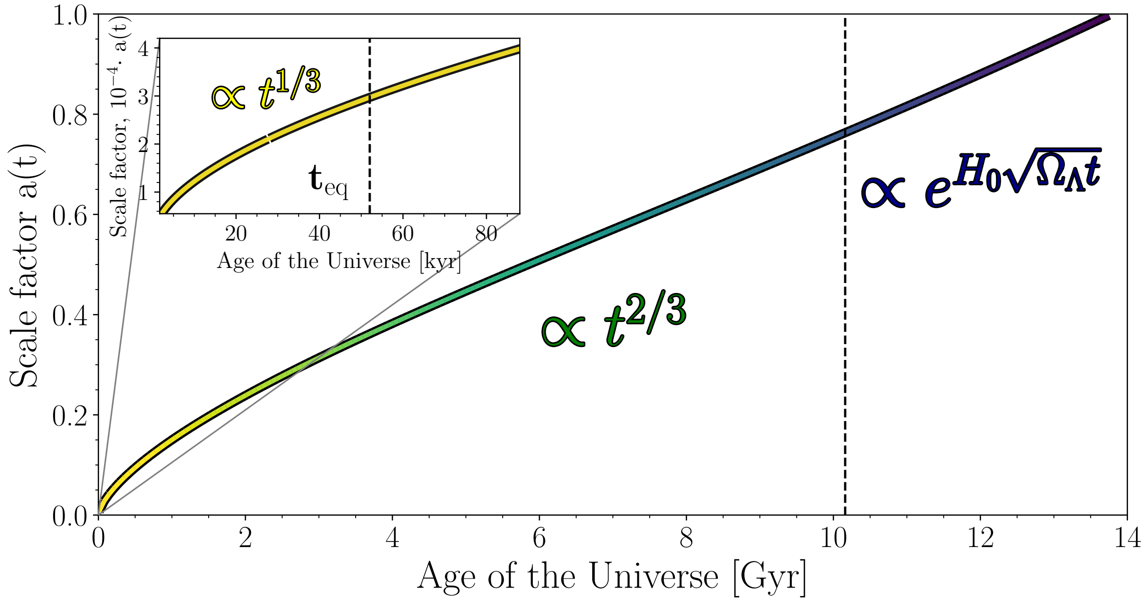

FIGURE 1.2: Expansion history of the Universe. Age of the Universe was calculated using `cosmology` sub-package from `astropy` for a Universe like ours with $\Omega_r = 9 \times 10^{-5}$, $\Omega_k = 0$, $\Omega_m = 0.307, \Omega_\Lambda = 0.693$, and $H_0 = 67.81\ \mathrm{km\,s^{-1}Mpc^{-1}}$. The black dashed line time at which dark energy starts to dominate ($t_{\Lambda\mathrm{D}} \approx 10.2$ Gyr).

From these solutions it is straightforward to see how the expansion is decelerated for MD case, while for ΛD case the Universe expands exponentially. In particular, for a MD case, we have

$$H(t) = \frac{\dot{a}}{a} = \frac{2}{3}\frac{1}{t} \quad \rightarrow \quad H_0 = \frac{2}{3}\frac{1}{t_0}, \tag{1.20}$$

and therefore, we can infer the age of the Universe (t_0) from the Hubble constant. If we set $H_0 \approx 70\ \mathrm{km\ s^{-1}Mpc^{-1}}$, we obtain $t_0 \approx 9.3 \times 10^9$ years, which is substantially smaller than the value of 13.8×10^9 years estimated from *Planck* data.

The rest of the chapter is dedicated to outlining the current state-of-the-art. This includes a more detailed description of the ΛCDM model in Section 1.2, an overview of the most relevant observational probes in Section 1.3, and a brief summary of the unresolved issues in Section 1.4. However, we first review the thermal history of the Universe in Section 1.1 (see Figure 1.3 for a schematic illustration).

1.1 Thermal history of the Universe

The Hot Big Bang model starts at the Planck time $\approx 10^{-43}$ s. Beyond this, the current theory breaks down and cannot make any prediction. It is also known as the *"quantum gravity limit"* because a theory beyond it requires an understanding of gravity at the quantum level[7].

---

[7] It is believed that above this limit all the forces (gravity, strong and weak nuclear forces, and electromagnetic force) were unified.

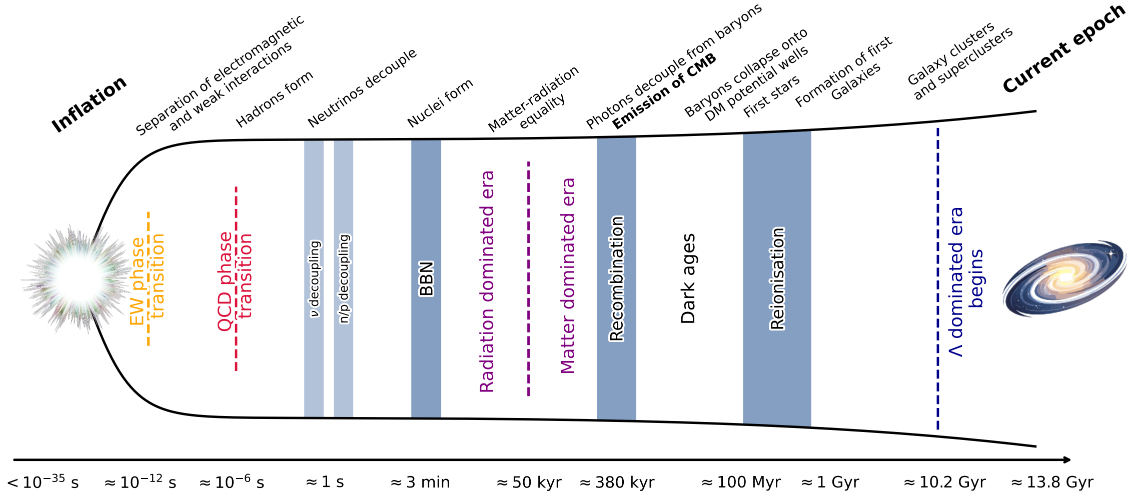

FIGURE 1.3: Schematic timeline of the thermal history of the Universe. Credits: Txemi.

Immediately afterward, from $\sim 10^{-43}$ to 10^{-36} seconds after the Big Bang, the universe is theorised to have entered the Grand Unified Theory (GUT) epoch, a hypothetical period in the early universe where strong, weak, and electromagnetic forces are believed to be unified in a single fundamental interaction. This period remains highly speculative, as it involves energy scales far beyond the current experimental capabilities. Toward the end of the GUT epoch, it is believed that the Universe experienced an abrupt quasi-exponential expansion where the scale factor grew by at least a factor of 10^{26}. This is cosmic inflation (Starobinsky, 1980; Guth, 1981; Linde, 1982), which was proposed to explain the observed homogeneity, isotropy, and flatness of the cosmos. A more detailed discussion is presented in Section 1.2. After GUT transition (t $\sim 10^{-35}\ s$, T $\sim 10^{27}\ K \simeq 10^{14}$ GeV)[8,9], the electroweak and strong forces emerged as independent fundamental forces, and quarks, leptons, and their respective anti-particles, filled the universe. At this moment the Universe was a hot and dense plasma, where all particles were interacting in thermal equilibrium[10]. As the Universe expanded, its temperature decreased, leading to significant changes in the composition of the thermal bath. At high temperatures, pair of particles and antiparticles were spontaneously created and annihilated in thermal equilibrium. However, as the temperature dropped below the rest mass of a given particle species, the production of that species were exponentially suppressed, and annihilation dominated. As a result, the particles and their antiparticles gradually disappeared from the thermal bath. This process occurred sequentially for different species, depending on their masses. The Universe continued cooling and at $t \sim 10^{-12}$ s ($T \sim 10^{15}$ K $\simeq 100$GeV) the electroweak transition took place, where the

---

[8] Temperature and energy are related by the Boltzmann constant, so that $1\,\mathrm{eV} \approx 1.16 \times 10^4\,\mathrm{K}$.

[9] During the radiation-dominated epoch, the temperature scales as $\frac{T}{\mathrm{MeV}} \propto t^{-1/2}$, where t is the cosmic time.

[10] A system of particles is in thermodynamical equilibrium if the particles exchange energy and momentum in an efficient way. It is established when the rate of interactions for a given process is much larger than the expansion rate, H, so $t_c = \frac{1}{\Gamma} << t_H = \frac{1}{H}$.

electromagnetic and weak forces separated. This transition occurred as a consequence of the Higgs mechanism. At this point, Higgs field acquired a non zero vacuum expectation value, giving their mass to the Standard Model (SM) particles. In particular, the massive gauge bosons mediating the weak interaction appeared, W^+, W^-,and Z^0, along with the massless photon, which mediates the electromagnetic force. At $t \sim 10^{-6}$ s ($T \sim 10^{12}$ K $\simeq 200$ MeV) the quantum chromodynamic (QCD) transition happened. Free quarks could not longer exist, and they became bound with gluons through strong force into hadrons (baryons and mesons). Heavier hadrons decayed immediately to lighter particles.

Until now, we have omitted an important process that must have happened in the early Universe, the baryogenesis. This is the physical process that generated the observed asymmetry between matter and antimatter. If, during the Big Bang, the Universe had produced the same amount of baryons and antibaryons, they would have annihilated each other, and after freeze-out, the remanent baryon-to-photon ratio would be of the order of 10^{-20} (Kolb & Turner, 1990). Freeze-out refers to the moment when a certain species' interaction rate becomes slower than the expansion rate of the Universe, preventing it from keeping thermal equilibrium with the plasma. At this point, the particles decoupled from the thermal bath and started evolving independently, preserving their temperature and number density, relic abundance, and propagating freely. For instance, if DM were composed of a weakly interacting massive particles (WIMPs), this would provide an explanation of how it acquires its relic abundance. Observations from the Big Bang Nucleosynthesis (BBN) and the CMB have constrained the baryon-to-photon ratio to be around 10^{-10} (Dodelson, 2003). The observed asymmetry requires a mechanism satisfying the three Sakharov conditions (Sakharov, 1991). Baryogenesis must have happened at early steps, when temperature was above the rest-mass energy of protons ($m_p = 938.3$ MeV from Navas et al. (2024)) and neutrons ($m_n = 939.6$ MeV from Navas et al. (2024)). Accordingly, the current baryonic content of the Universe was established during the first microsecond after the Big Bang.

Between $t \sim 10^{-6}$ s and $t \sim 1$ s, the Universe entered in the Lepton era as the electrons, neutrinos and their antiparticles, and photons, were the only remaining relativistic species. In particular, neutrinos were coupled to the plasma through weak interactions. These reactions started to be inefficient for $T \sim 10^{10}$ K and neutrinos decoupled from the plasma, generating the cosmic neutrino background (CνB). Shortly after, the temperature dropped below the electron rest mass ($m_e = 0.511$ MeV from Navas et al. (2024)), electron-positron pairs started to annihilate, leaving a small amount of electrons and injecting energy in the photon gas, but not in the recently decoupled neutrino background[11]. This is the reason why the blackbody spectrum of CνB is predicted to have a lower temperature than the CMB, $T_\nu \approx 1.95$ K (Weinberg, 2008). From $t \sim 1$ s, nuclear reactions between baryons started. As the temperature dropped below a critical threshold, the number of high-energy

---

[11] This is not completely true. The weak interactions provide still some thermal contact between neutrinos and the plasma, and electron-positron annihilation slightly increases neutrino energy density. As a result, the effective number of neutrino species increases to $N_{\text{eff}} = 3.045$ (Weinberg, 2008).

photons capable of breaking nuclear bonds decreased significantly. This allowed for the formation of the first light elements. Most of the free neutrons ended up bound in ^{4}He, with a few residual abundances of D, ^{3}He, ^{7}Li, and ^{7}Be. This process is known as *Primordial Nucleosynthesis*, or more commonly, Big Bang Nucleosynthesis (BBN). Further details on BBN are provided in Section 1.3.1.

An important event in the history of the early Universe was the photon decoupling. As the Universe cooled below approximately 0.3 eV, electrons and nuclei started to combine into neutral hydrogen and helium atoms in a process known as recombination[12] (Zeldovich et al., 1968; Peebles, 1968). This transition drastically reduced the density of free electrons, allowing the mean free path of photons to increase rapidly. By the time the temperature dropped to around 0.25 eV, roughly $380,000$ years after the Big Bang, the photon scattering rate dropped below the Hubble expansion rate. At this point, photons decoupled from matter[13], and the Universe became transparent for the first time. These primordial photons have since travelled and are now observed as the Cosmic Microwave Background (CMB). The epoch of recombination thus marks the surface of the last scattering, setting the ultimate horizon of the observable Universe through electromagnetic radiation.

After recombination, the Universe entered the Dark Ages phase, where the only photons filling the Universe were the CMB photons and 21 cm emission from neutral hydrogen (Scott & Rees, 1990). During this epoch, baryonic matter began to fall into the gravitational potential wells that DM already started to form. After matter–radiation equality, the initial inhomogeneities in the dark matter density field grew through gravitational instability, leading to the formation of dark matter halos. Gas condensed within these halos, eventually giving rise to the first stars, known as Population III stars (Schaerer, 2002), and the earliest galaxies (Dayal & Ferrara, 2018). These luminous objects emitted intense ultraviolet radiation, which ionized the surrounding neutral hydrogen in a process known as *cosmic reionisation* (Barkana & Loeb, 2001). As a result, photons began to interact with free electrons via Thomson scattering again, making the Universe partly opaque. Population III stars are thought to be extremely massive, burning their hydrogen rapidly and exploding in powerful supernovae that enrich the interstellar medium with heavy elements synthesized through nuclear fusion. Structure formation continued throughout cosmic history, with galaxies grouping into clusters and superclusters.

---

[12] Although the hydrogen binding energy is 13.6 eV, the Saha equation (Baumann, 2022) shows that the free electron fraction decreases significantly only when the temperature drops well below this value, due to the abundance of high-energy photons capable of ionizing hydrogen.

[13] Dark matter decoupled much before and could start to collapse via gravity around the initial density perturbations. However, this process was highly suppressed during radiation dominated era.

1.2 Standard Cosmological Model: ΛCDM and inflation

Although the Hot Big Bang theory successfully describes the early thermal history of the universe, it does not by itself specify the dynamics of cosmic expansion or the precise composition of the universe today. The ΛCDM (Lambda Cold Dark Matter) model extends the Hot Big Bang framework by incorporating two new components that have been already discussed in the previous section; a cosmological constant (Λ) to account for the observed accelerated expansion, and cold dark matter (CDM) to explain large-scale structure (LSS) formation, gravitational lensing, and galaxy rotation curves, among others.

The ΛCDM model is the current standard model of cosmology, the simplest model that explains a wide range of observations, from CMB to LSS. The model makes some assumptions, i) the Universe is spatially flat ($k = 0$) governed by general relativity and composed of ii) a cosmological constant (Λ), ordinary matter, which is composed of all the particles of the Standard Model (SM) of particle physics[14], and mysterious DM which is assumed to be cold, i.e. non-relativistic DM. It also incorporates the iii) inflationary paradigm to solve some issues of the classical Big Bang model, and provides a framework to generate the seeds of the structures we observe today.

In this section, we briefly go through each of the components and describe the six cosmological parameters on which the model depends. The section concludes with a short overview of inflationary dynamics. This includes an introduction to motivate inflation and the equations that determine the dynamics of the background. A more detailed discussion of the initial perturbations is presented in Chapter 2.

- **Ordinary matter:** It refers to the matter that makes up stars, planets, and all visible structures in the Universe. It only accounts for 5% of the Universe content, which is matter well described by the SM of particle physics, primarily baryonic matter. It is the only form of matter that interacts via the electromagnetic force, absorbing, emitting, and reflecting light. Neutrinos are also included here. In the SM, they are assumed to be massless. However, it is already known from neutrino oscillations (Fukuda et al., 1998), first proposed by Bruno Pontecorvo, that at least two of the neutrinos should have a non-zero mass. The reason is because the flavor states (e, μ, τ) are different from the mass eigenstates, which have slightly different masses.

- **Cold dark mattter (CDM):** Around 26% of the energy content of the Universe is in the form of a non-visible type of matter, dark matter (DM), which either does not interact or interacts very weakly with electromagnetic radiation. Currently, its effect can only be observed through gravity force. It plays a vital role in the formation of galaxies and cosmic structures. The nature of this matter is still a mystery, and there are several

---

[14] SM of particle physics describes three of the known fundamental forces in the Universe and classifies the elementary particles, which are fermions (quarks and leptons) and bossons.

proposals, i.e. WIMPs, axions, self-interacting dark matter, fuzzy dark matter, (Mili-)charged dark matter, primordial black holes (PBH), etc. (Cirelli et al., 2024). DM must be *cold*, which means that DM must be non-relativistic. Hot DM models, including the SM neutrinos, are already excluded by structure formation. As alternatives to CDM, there are the so-called warm dark matter (WDM) models. In these scenarios, DM particles were produced with a velocity large enough to affect structure formation (Strigari, 2013), and in particular, they could suppress the formation of small-scale structures.

- **Dark energy:** It is a mysterious form of energy that dominates the energy density of the Universe. Accounting for 69% of the total content, dark energy counteracts gravity and is responsible for the accelerated expansion of the Universe. The simplest explanation is a cosmological constant (Λ) with an equation of state given by $\omega \simeq -1$, although alternative models with an evolving equation of state or modified gravity theories have also been proposed. The latest ones proposed that the accelerated expansion is a consequence of a change in the laws of gravity at large scales. These theories include f(R) gravity (Sotiriou & Faraoni, 2010).

1.2.1 Cosmological parameters

The base ΛCDM model compresses the dynamics of the Universe into a set of 6 independent cosmological parameters. The rest of the parameters can be derived from that set, which is not unique. The most common set of parameters are: $\Omega_b h^2$, $\Omega_c h^2$, θ_*, A_s, n_s, and τ. Observations of the CMB anisotropies[15], supernovae, and LSS can place tight constraints on these parameters. Table 1.1 shows the latest constraints from the combination of CMB anisotropies, CMB lensing, and Barionic Acoustic Oscillations (BAO). In this section, we briefly review each of them, including also a set of parameters that describe extensions of the ΛCDM model.

Geometry, expansion, and composition

This subsection describes the parameters characterising geometrical properties of the Universe, its evolution, and composition.

The angular size of the sound horizon at recombination, denoted as θ_*, is a key geometrical parameter, mainly constrained by the CMB. It is a robustly determined quantity, known with a precision better than 0.03%, and almost independent of late-time cosmology (Planck Collaboration et al., 2020c). It is defined as the ratio between the comoving size of the sound horizon at the epoch of recombination, $r_s(z_*)$, and the angular diameter distance[16] to that

---

[15] Small temperature fluctuations, of the order of 10^{-5}K, in the CMB that reflect primordial density fluctuations.

[16] The angular diameter distance relates the physical size of an object to its apparent angular size, and it depends on the underlying cosmological model.

epoch, $D_A(z_*)$,

$$D_A(z) = \frac{S_K(\chi)}{1+z} \quad \xrightarrow[k=0]{} \quad D_A(z) = \frac{\chi(z)}{1+z} = \frac{1}{1+z}\int_0^z \frac{\mathrm{d}z'}{H(z')}. \tag{1.21}$$

$$\theta_* = \frac{r_s(z_*)}{D_A(z_*)}. \tag{1.22}$$

The first quantity depends on the physics of the early Universe, while the second one is sensitive to the expansion history and the spatial geometry.

As already shown in previous sections, the evolution of the Universe depends on the dimensionless energy densities of its constituents (see Eq. 1.17), i.e., baryons (b), CDM (c), dark energy (Λ), radiation (r), and curvature (k). The energy density of radiation is well known and fixed by the CMB temperature, and the current observations are compatible with a spatially flat Universe, leaving only two quantities to be constrained. These are the physical densities $\Omega_b h^2$ and $\Omega_c h^2$, where $h = H_0/100\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ is the dimensionless Hubble constant[17]. In some analyses, the Hubble constant is used in place of θ_* as one of the six independent parameters. Although they are not strictly interchangeable, each can be derived from the other within the framework of a given cosmological model.

Initial conditions

The primordial power spectrum of density fluctuations, which is described later in Section 2.2.2, is parametrised by two quantities. These fluctuations, generated during inflation, are the initial seeds for the formation of structures in the late Universe. The spectrum is modelled as a power law characterised by its amplitude (A_s) and spectral index (n_s).

Astrophysical parameters

As previously discussed, the Universe experienced a reionisation epoch when the first stars and galaxies formed, reionising the hydrogen in the intergalactic medium. The optical depth to reionisation, τ, provides a measure of the line-of-sight free-electron opacity to the CMB photons,

$$\tau = \int_0^{\mathrm{z_{rei}}} \mathrm{dz}\sigma_\mathrm{T}\mathrm{n_e(z)}\frac{\mathrm{dr}}{\mathrm{dz}} = \int_0^{\mathrm{z_{rei}}} \frac{\sigma_\mathrm{T}\mathrm{n_e(z)}}{\mathrm{(1+z)H(z)}}\mathrm{dz}, \tag{1.23}$$

where σ_T is the Thomson cross section, $n_e(z)$ denotes the electron density as a function of redshift, z_rei represents the redshift at which reionisation begins, and $\mathrm{d}r/\mathrm{d}z$ is the line-of-sight proper distance per unit redshift, given by $\left[(1+z)H(z)\right]^{-1}$ in a flat Universe. Larger value of τ means, for a fixed electron density function, a larger period of reionisation, and thus an earlier star and galaxy formation.

---

[17] $\Omega_\Lambda h^2$ is derived from these two as the total energy density is equal to unity for flat Universe.

Parameter	Meaning	Value
$\Omega_b h^2$	Physical baryon density	0.02229 ± 0.00012
$\Omega_c h^2$	Physical cold dark matter density	0.1186 ± 0.0009
$100\theta_*$	Angular size of sound horizon	1.04111 ± 0.00024
H_0	Hubble constant [km/s/Mpc]	67.81 ± 0.38
τ	Optical depth to reionisation	0.0605 ± 0.0059
n_s	Scalar spectral index	0.9690 ± 0.0035
$\ln(10^{10} A_s)$	Amplitude of scalar perturbations	3.048 ± 0.012

Table 1.1: Best-fit parameters of the base ΛCDM model obtained using a combination of the *Planck* CMB temperature and polarisation power spectra for the *Planck* Data Release 4 (PR4), CMB lensing, and BAO (Tristram et al., 2024). In addition to the six standard cosmological parameters, H_0 is also included in this table.

Beyond ΛCDM

Several additional parameters are introduced in extensions of ΛCDM model. Each corresponds to a theoretical modification beyond the six base parameters. In the following, we provide an overview of some of them.

- **Neutrino and light relics** ($\sum m_\nu$, $N_{\rm eff}$)**:** The base ΛCDM model assumes a normal mass hierarchy[18] with a minimum mass of $\sum m_\nu = 0.06$ eV and an effective number of relativistic species fixed at $N_{\rm eff} = 3.045$ (Planck Collaboration et al., 2020c). Neutrinos leave detectable imprints on cosmological observables. In particular, the total neutrino mass, which determines the timing of the transition from relativistic to non-relativistic behaviour, affects the expansion rate and has consequences on gravitational clustering and the growth of structures. Cosmological probes such as the CMB and the BAO have placed upper bounds on $\sum m_\nu$, with current limits improving by nearly a factor of 20 over the past two decades. Regarding $N_{\rm eff}$, deviations from the standard value could indicate the presence of additional light particles, such as sterile neutrinos or light axions. Current observations remain fully consistent with the standard model prediction. See (Lesgourgues et al., 2013; Navas et al., 2024) for a more detailed review on neutrino impact in cosmology.

- **Dark energy modification** (ω_0, ω_a)**:** The base ΛCDM model fixes the equation of state of DE to be $\omega = -1$, a cosmological constant. However, if DE is actually a dynamical fluid, its equation of state would be time dependent, which is typically characterised by the so-called Chevallier-Polarski-Linder (CPL) parametrization (Chevallier & Polarski, 2001; Linder, 2003),

$$\omega(a) = \omega_0 + (1 - a)\omega_a. \tag{1.24}$$

---

[18] The neutrino mass hierarchy refers to the ordering of the three neutrino mass eigenstates, which can be either normal (NH) ($m_1 < m_2 < m_3$) or inverted (IH) ($m_3 < m_1 < m_2$). Neutrino oscillation experiments are sensitive only to differences in the squared masses, leaving the absolute scale and hierarchy undetermined.

- **Curvature (Ω_k):** The base ΛCDM model assumes spatially flat Universe, as predicted by the simplest inflationary models. Introducing a non-zero spatial curvature parameter, Ω_k, allows deviations from flat geometry, with $\Omega_k > 0$ corresponding to an open Universe and $\Omega_k < 0$ to a closed one. Curvature affects the angular diameter distance to the last scattering surface and modifies the expansion history (see Eq. 1.17).

- **Primordial spectrum extensions ($r_{0.002}$, dn_s/d ln k):** In the base ΛCDM model the primordial scalar perturbations are characterised by an amplitude (A_s) and a constant spectral index (n_s). In a more general case, the spectral index can depend on the scale as follows,

$$n(k) = n_s - 1 + (1/2)(dn_s/d\ln k)\ln(k/k_0) + (1/6)(d^2n_s/d\ln k^2)\left(\ln(k/k_0)\right)^2, \quad (1.25)$$

  where running, and running of the running, of the spectral index have been added.

  Apart from primordial scalar perturbations, inflation also predicts the generation of primordial gravitational waves, or tensor modes (see Section 2.2.2 for more details). The primordial power spectrum of tensor modes is also modelled by a power law characterised by its amplitude (A_t) and spectral index (n_t). In general, the amplitude of tensor perturbations is usually given in terms of the tensor-to-scalar ratio,

$$r = \frac{A_t}{A_s} \quad (1.26)$$

  The base ΛCDM model assumes no tensor modes, i.e., $r = 0$. Allowing r to vary defines an extended model, $\Lambda\text{CDM} + r$, for which the latest constraint is $r < 0.032$ at 95% confidence level (Tristram et al., 2022).

1.2.2 Inflationary paradigm

Cosmic inflation is expected to play a crucial role in the early Universe by driving the Universe towards the observed large-scale homogeneity and flatness state, even if it started from more generic initial conditions.

Although the inflationary period has not yet been confirmed, it is a widely accepted extension of the ΛCDM model. Inflation offers elegant solutions to fundamental issues of the Standard Big Bang model, i.e., horizon, flatness, and magnetic monopole problems, yields predictions that are in excellent agreement with observations, and provides a mechanism for generating the primordial perturbations that seeded the LSS (see Section 2.2.2). For these reasons, the confirmation of the inflationary paradigm is one of the current goals of modern cosmology.

The following pages provide a brief review of the horizon, flatness, and magnetic monopole problems, as well as the basic properties and dynamics of the simplest inflationary models. A detailed discussion of this topic can be found in D. Baumann's textbook

(Baumann, 2022) and lecture notes (Baumann, 2009).

Standard Big Bang problems

The standard Big Bang model cannot explain the observed large-scale homogeneity and flatness without imposing by hand very fine-tuned initial conditions[19]. To understand these problems, we must introduce a few concepts.

For two separated regions of spacetime to be causally-connected, light must have been able to travel between them. This condition is satisfied if both lie within the past lightcones defined by their respective particle horizons. At any given moment in the Universe's history, the particle horizon defines the maximum comoving distance from which light could have reached an observer since the beginning of the Universe,

$$d_h(\eta) = \eta - \eta_i = \int_{t_i}^{t} \frac{\mathrm{d}t}{a(t)} = \int_{\ln a_i}^{\ln a} (aH)^{-1} \mathrm{d}\ln a \tag{1.27}$$

where η is the conformal time[20]. Therefore, the particle horizon is related to the evolution of the comoving Hubble radius, $(aH)^{-1}$, which, for ordinary matter, always increases, and physically defines the distance at which objects recede from the observer at the speed of light due to the cosmic expansion. Integrating equation 1.27 from initial singularity up to recombination, which defines the particle horizon when the CMB was emitted, and comparing it to the comoving distance of the last scattering surface, one finds that two points in the CMB separated by more than approximately 2 degrees in the sky lie in causally disconnected regions (Baumann, 2022). This is shown in Figure 1.4. Then, how is it possible that CMB photons arriving from all directions exhibit the same temperature? This is the horizon problem.

Additionally, the standard Big Bang model suffers from the flatness problem. The time-dependent curvature parameter is proportional to the square of the comoving Hubble radius,

$$\Omega_k(a) = \frac{\rho_k(a)}{\rho_{\mathrm{crit}}(a)} = \frac{H_0^2}{(aH)^2}\Omega_{k,0}. \tag{1.28}$$

Taking into account that $(aH)^{-1}$ grows during the standard Big Bang model, this means that the curvature should be smaller in the past. The current measurements, which combines CMB anisotropies, lensing, and BAO, are consistent with our Universe being spatially flat, $\Omega_k = 0.0000 \pm 0.0016$ (Tristram et al., 2024). To ensure that $\Omega_k < 0.0016$ today, its value at the time of electroweak phase transition must have been $\Omega_k(t_{\mathrm{EW}}) < 10^{-29}$. This means that in the early Universe the energy density matched the critical density to at least 29 significant digits. Achieving such a fine-tuned match in the early Universe seems extremely unnatural.

---

[19] We are not considering that these conditions are outcomes from a still unknown quantum theory of gravity.

[20] Conformal time is defined as $d\eta = dt/a(t)$. In this parametrization of time, the expansion is "factored out", making light rays behave as if they move in flat spacetime, i.e., null geodesics are straight lines.

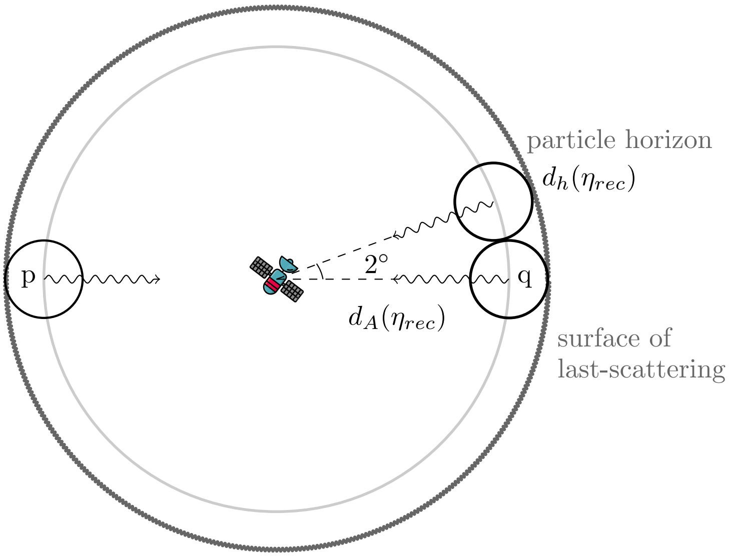

FIGURE 1.4: Illustration of the horizon problem in the standard Big Bang model. Any two points in the CMB separated by more than approximately 2 degrees have non-overlapping particle horizons, as there is not enough conformal time between the initial singularity and the surface of last-scattering for causal contact to be established. Consequently, in the standard Big Bang model the CMB consists of a few tens of thousands causally disconnected patches. Figure adapted from Baumann (2022).

Inflation emerges as a compelling solution to these problems by adding a short phase of decreasing Hubble radius during which the homogeneity and flatness conditions are established,

$$\frac{d}{dt}(aH)^{-1} < 0. \tag{1.29}$$

This condition leads to a period of accelerated expansion, $\ddot{a} > 0$. The key to solving the Horizon problem is that inflation makes the particle horizon much larger than the Hubble radius, allowing regions that appear causally disconnected today to have been in causal contact in the past. This is possible because inflation is driven by a fluid that violates the strong energy condition, $1 + 3\omega < 0$, as dark energy does. In particular, in inflationary cosmology, the initial singularity is pushed to negative conformal times[21], and $\eta = 0$ only marks the transition point, also known as reheating, between inflation and standard Big Bang theory. This is shown in Figure 1.5.

In inflationary cosmology, flatness is no longer a problem as $\Omega_k = 0$ is now an attractor. Consequently, if inflation lasts long enough, the initial curvature in the Big Bang model could be sufficiently suppressed to maintain its value below the observational constraint.

---

[21] The singularity is still at t = 0, but due to the abrupt expansion the light cones are stretched in such a way that they overlap before the end of inflation.

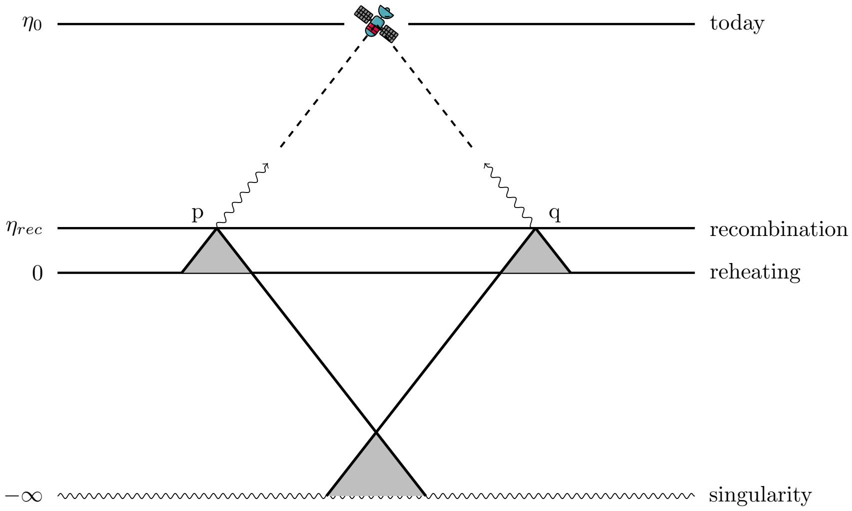

FIGURE 1.5: Illustration of the inflationary solution to the horizon problem. Inflation pushes the initial singularity to negative conformal times, which provides enough conformal time for all the points in the CMB to have overlapping past light cones. $\eta = 0$ now corresponds to the reheating period, a transition between inflationary cosmology and standard Big Bang theory. Figure adapted from Baumann (2022).

Slow-roll inflation

Up to now, we have presented a way to solve the horizon and flatness problems. However, this generates a new question to answer: What is the physical mechanism driving the inflationary period?

In the simplest inflationary models, the energy density, and consequently the dynamics of the Universe, is dominated by a single scalar field, $\phi(t, \mathbf{x})$, known as the inflaton. In this section, we focus on the evolution of the background, $\phi(t)$, of this field[22] and the conditions under which it leads to an accelerated expansion of the Universe.

This field has a potential energy density, V(ϕ), and a kinetic energy, $\frac{1}{2}\dot{\phi}^2$, so its energy-momentum tensor takes the form of a perfect fluid with dynamical energy density and pressure,

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi); \qquad p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi); \qquad \omega_\phi = \frac{p_\phi}{\rho_\phi} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}. \tag{1.30}$$

The equation of state shows that the inflaton behaves as a cosmological constant, driving a period of exponentially accelerated expansion, when the potential energy dominates over the kinetic term. The dynamics of the Universe during this epoch are fully determined by

---

[22] It is assumed that the field is homogeneous except for small fluctuations, $\delta\phi$, which are studied in the next chapter.

the Friedmann and Klein-Gordon equations

$$H^2 = \frac{8\pi G}{3}\left[\frac{1}{2}\dot{\phi}^2 + V\right], \tag{1.31}$$

$$\ddot{\phi} + 3H\dot{\phi} = -\frac{dV}{d\phi}. \tag{1.32}$$

Both equations are coupled. The energy content of the field determines the expansion rate, which serves as a friction for the field. Inflation is sustainable as long as the kinetic energy does not contribute significantly to the total amount of energy. In particular, when the field slowly rolls down the potential, as occurs between $\phi_{\rm CMB}$ and $\phi_{\rm end}$ in Figure 1.6, inflation takes place within the so-called slow-roll regime. The conditions for it are satisfied when the speed and acceleration of the field are small, conditions that are controlled by the slow-roll parameters,

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = \frac{\frac{3}{2}\dot{\phi}^2}{\frac{1}{2}\dot{\phi}^2 + V}; \qquad \delta \equiv -\frac{\ddot{\phi}}{H\dot{\phi}}. \tag{1.33}$$

The slow-roll approximation is valid when $\{\epsilon, \delta\} << 1$. In this regime, the equations are simplified and the expansion is exponentially accelerated. Inflation lasts until $\epsilon = 1$, which is the moment when $\omega_\phi = -1/3$ and the accelerated expansion stops. The amount of inflation is quantify by the number of e-folds, which corresponds to the number of times the scale factor increases by a factor of e,

$$\mathcal{N}_{\rm tot} \equiv \int_{a_i}^{a_e} \mathrm{d}\ln a = \int_{t_i}^{t_e} H(t)\mathrm{d}t, \tag{1.34}$$

where a_e and a_i are the scale factors at the end and beginning of inflation, respectively, and t_i and t_e are defined so that $\epsilon(t_i) = \epsilon(t_e) = 1$. Even if inflation last outside the slow-roll approximation[23], most of the e-folds are generated in that regime.

At the end of inflation, the field begins to oscillate around the minimum of its potential, and the energy stored in the field is transferred to SM particles through a process known as reheating. For this to occur, the inflaton must be coupled to the SM fields. The new produced particles start interacting with each other and eventually they reach the thermal equilibrium that characterised the beginning of the Hot Big Bang model. The reheating temperature, T_R, which is the temperature of this new plasma, must be much larger than 1 MeV to allow for baryogenesis after inflation (Baumann, 2022).

Currently, the inflationary landscape contains hundreds of different scenarios (Martin et al., 2013). Many candidates for the shape of the potential $V(\phi)$ remain consistent with observations in the slow-roll single-field framework, and numerous additional models go beyond the single-field paradigm. Each of them makes different predictions about the prop-

---

[23] Between $\omega_\phi = -1$ and $\omega_\phi = -1/3$ (or equivalently, $\epsilon << 1$ and $\epsilon = 1$) the expansion is accelerated but deviates from exact exponential growth. It is a transition between slow-roll regime and the end of inflation.

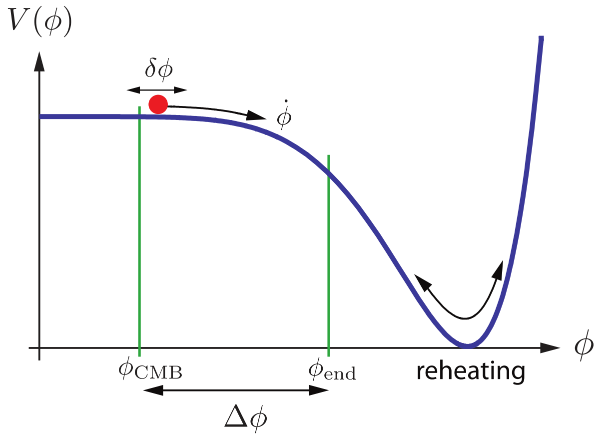

FIGURE 1.6: Schematic illustration of a slow-roll single-field potential. At the beginning, the field slowly rolls down the potential and the inflaton behaves as a dark energy source with $\omega_\phi \approx -1$ equation of state, driving an exponentially accelerated expansion. Inflation ends when the kinetic term starts contributing significantly to the total energy density. Then, the field starts oscillating and decaying to the SM particles in a process known as reheating. $\phi_{\rm CMB}$ corresponds to the value of inflaton at the time when the largest observable scales today exited the horizon, ensuring that all the regions of the CMB were initially causally connected. Figure taken from Baumann (2009).

erties of the primordial perturbations such as n_s, non-Gaussianity, and the tensor-to-scalar ratio.

1.3 Cosmological probes

This section provides a brief summary of the cosmological probes that have provided us precise measurements of the composition, geometry, and evolution of the Universe. These include the light element abundance (Section 1.3.1), the large-scale distribution of matter in the Universe (Section 1.3.2), and standard candles/sirens such as supernovae and gravitational waves (Section 1.3.3).

The CMB stands as one of the most compelling pieces of evidence for the Big Bang theory. This radiation, generated at the time of recombination, is an emission of uniform black body thermal energy that permeates the entire Universe, accounting for most of the radiation energy in the Universe. This section does not include any description about it, as a comprehensive and extensive discussion is provided in Chapters 2 and 3, with special attention to the CMB anomalies, the main topic of this thesis.

1.3.1 Big Bang Nucleosynthesis

As previously mentioned in Section 1.1, light elements are produced during BBN, while heavier elements are the result of nuclear reactions occurring in the stellar cores. The primordial abundance of light elements can be predicted using well-established nuclear physics, with errors primarily arising from uncertainties in the cross sections, which is the probability of a specific reaction to occur.

There are some key aspects to be taken into account in the BBN. One of the inputs for the BBN is the neutron abundance. Around $T \sim 1\mathrm{MeV}$, the weak interactions by which protons transformed into neutrons, and vice versa, became unlikely, and the neutron abundance is fixed to 1/6. The other important aspect is the baryon-to-photon ratio, η_b. The first step in the BBN is the deuterium production ($n + p \rightarrow D + \gamma$). However, due to small baryon-to-photon ratio the deuterium production is inhibited until the temperature drops well below its binding energy. This occurs essentially because, even if photons with energies comparable to the binding energy are suppressed, their abundance is so high that nuclei can still encounter such photons and dissociate (Dodelson, 2003). Therefore, helium production is delayed until enough deuterium is presented in the Universe. This is known as the deuterium bottleneck (Steigman, 2007). In the meantime, the neutron abundance decreased according to its lifetime ($\tau = 878.4 \pm 0.5$ s from Navas et al. (2024)). After deuterium, $^4\mathrm{He}$ is produced, with a small fraction of $^3\mathrm{He}$ and $^3\mathrm{H}$. Helium is favored over deuterium as the binding energy is larger. Finally, small amount of $^7\mathrm{Li}$ is produced, and $^7\mathrm{Be}$, but this eventually decays. Reaction chains stop and no heavier nuclei are produced, mainly because there are not stable nuclei with mass numbers equal to 5 or 8. In stars heavier elements are formed through reactions that involve more than two particles, the triple alpha process for example. However, during BBN there is not enough time and baryon density for this. Figure 1.7 shows the evolution of the abundances of different elements produced during BBN epoch.

Comparing the predictions of BBN with the observed relative abundances of light elements provides a powerful probe of particle physics and cosmology during the first few minutes after the Big Bang. The key challenge is to identify environments where these primordial abundances have remained largely unaltered by subsequent astrophysical processes (Cooke, 2024). The relative abundance of deuterium can be inferred by exploring the absorption lines that low metallicity gas clouds imprints in the spectra of high-redshift quasars (Cooke et al., 2018; Cooke, 2024). The deuterium abundance provides an alternative way to estimate the baryon density, and its remarkable agreement with the value inferred from the CMB, as shown in Figure 1.8, stands as one of the strongest confirmations of the standard cosmological model.

The primordial $^4\mathrm{He}$ abundance, Y_p, is most commonly determined by analysing the emission lines from metal-poor H II regions in blue compact dwarf galaxies (Kurichin et al., 2021). The helium primordial abundance also affects the CMB, in particular the damping tail (Cooke, 2024), as the electrons recombine with helium before hydrogen. On the other hand, the measurement of the primordial abundant of $^3\mathrm{He}$ is nearly impossible as it is hard to find environments undisturbed by stellar activity. Finally, the best estimation of the primordial lithium abundance comes from the study of $^7\mathrm{Li}$ absorption in the atmospheres of

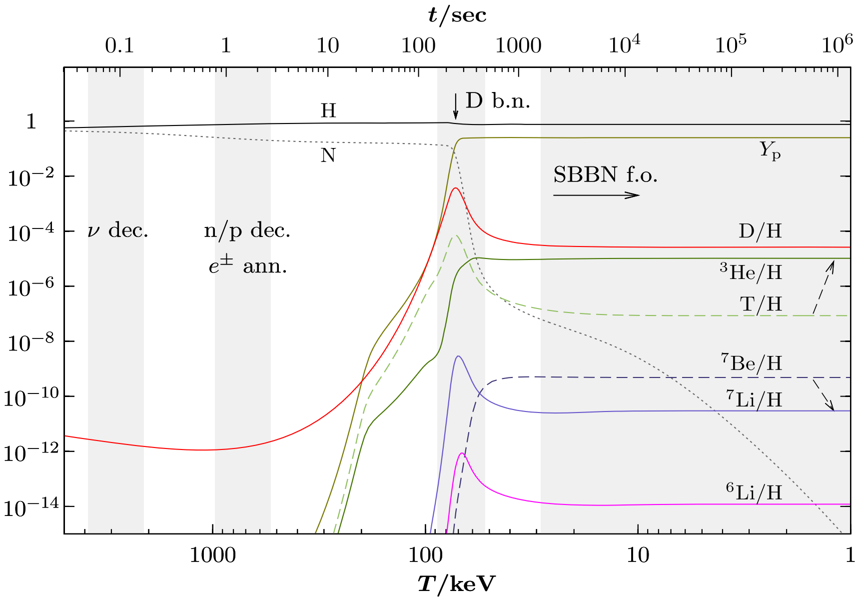

FIGURE 1.7: Evolution of light elements abundances during Big Bang Nucleosynthesis (BBN). Grey bands correspond to the main BBN stages. From left to right: neutrino (ν) decoupling, neutron-proton ratio freeze-out and electron-positron annihilation, deuterium bottleneck, and freeze-out of all nuclear reactions. Figure taken from Pospelov & Pradler (2010).

old metal-poor stars orbiting the Milky Way galaxy (Cooke, 2024). However, the inferred lithium abundance is significantly lower than the predicted amount from BBN.

1.3.2 Large-Scale Structure

LSS refers to the distribution of matter in the Universe on scales larger than individual galaxies, including clusters, filaments, and voids, formed by gravitational growth of initial density fluctuations. It has become an increasingly important source of cosmological information, particularly for constraining the nature of dark matter and dark energy. While the CMB mainly probes the early Universe at time of recombination, it provides limited insight into the late-time dynamics of these components. Therefore, the CMB needs to be complemented by high-quality measurements of the cosmic expansion history and the growth of the LSS (Euclid Collaboration et al., 2025).

One of the main observables of the galaxy distribution is the three-dimensional linear matter power spectrum, $P_m(k)$, which is defined as the variance of the Fourier modes of the three dimensional distribution of matter. In particular, Figure 1.9 shows the 3D linear $P_m(k)$ at $z = 0$. Reconstructing the three-dimensional distribution of matter requires surveys capable of accurately measuring both the angular position and redshifts of hundreds of thousands of galaxies. Examples of completed surveys that have mapped the LSS include the Sloan Digital Sky Survey (SDSS) (York et al., 2000), the 2dF Galaxy Redshift Survey (Colless et al., 2001), SDSS-III's Baryon Oscillation Spectroscopic Survey (BOSS) (Dawson

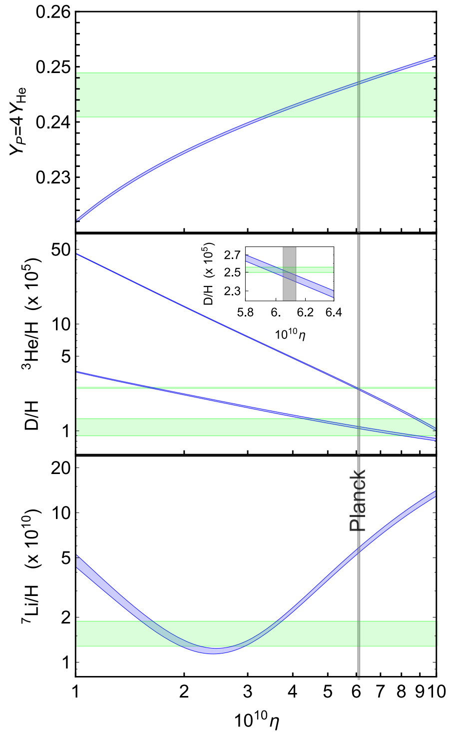

FIGURE 1.8: The primordial abundance of light elements predicted by BBN (blue curves) for different values of baryon-to-photon ratio, η. The green shaded regions represent the observational abundance, including their uncertainties. Finally, the grey vertical lines shows the current baryon-to-photon ratio constraints based on the CMB. Figure taken from Pitrou et al. (2018).

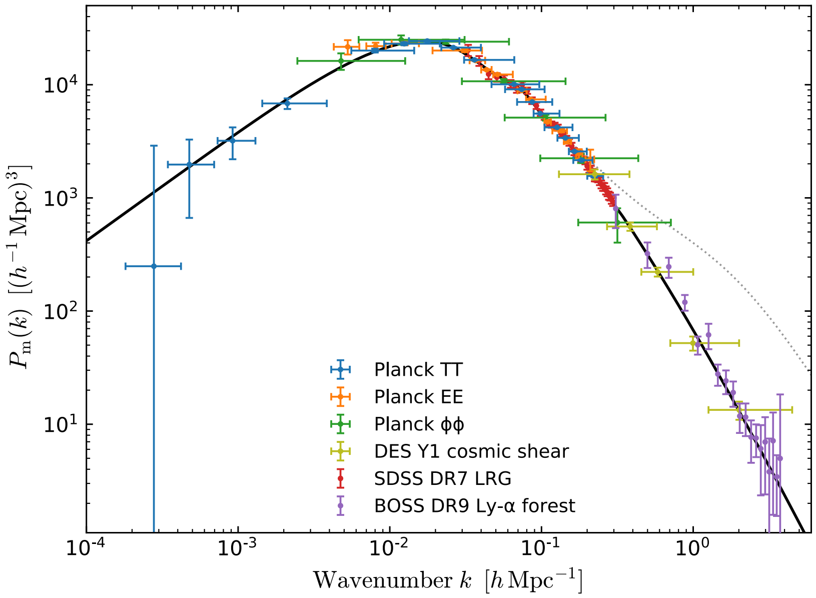

FIGURE 1.9: Linear matter power spectrum at $z = 0$ obtained from CMB and LSS surveys. The solid black line shows the $\Lambda\mathrm{CDM}$ predictions at linear-theory level, while the dotted line includes the non-linear effects at small scales. Figure adapted from Planck Collaboration et al. (2020a).

et al., 2013), eBOSS (Zhao et al., 2016), and Dark Energy Survey (DES) (Dark Energy Survey Collaboration et al., 2016). Future and ongoing surveys like J-PAS (Benitez et al., 2014), DESI (DESI Collaboration et al., 2016), Euclid (Euclid Collaboration et al., 2025), and Rubin Observatory's Legacy of Space and Time (LSST) (Ivezić et al., 2019) will improve our understanding of dark energy. In particular, BAO serve as a standard ruler in cosmology. It defines a preferred scale in the clustering of matter imprinted by acoustic waves that propagate in the photon-baryon plasma in the pre-recombination Universe (Dawson et al., 2013). Measuring this characteristic scale across different redshifts enables precise measurements of the expansion history. In addition, gravitational lensing offers a complementary probe of dark energy and the distribution of dark matter. As photons from distant galaxies cross the cosmic web, their paths are deflected by the gravitational potential generated by the matter distribution along the line of sight. In particular, this effect distorts the observed shapes of high-redshift galaxies, a phenomenon known as cosmic shear (Kilbinger, 2015), which can be measured at different redshifts to trace the evolution of LSS over the cosmic time.

On the other hand, it is well known that on small scales the matter fluctuations have become large and enter a non-linear regime, which is really hard to model. On top of that, baryonic feedback also provides an extra layer of complexity. This includes star formation, supernova explosions, and Active Galactic Nuclei (AGN) feedback. Taking into account that linear theory fails to capture these complex processes, cosmological simulations have emerged as an essential tool for understanding these effects (Aricò et al., 2021). MillenniumTNG (Pakmor et al., 2023) and BAHAMAS (McCarthy et al., 2017) projects are two examples of detailed N-body and hydrodynamical simulations.

1.3.3 Supernovae and Gravitational Waves

The present expansion rate of the Universe, H_0, can be determined in the local Universe by measuring distances and redshifts. Unfortunately, measuring cosmological distances is not an easy task. One of the solutions is based on the standard candles, astrophysical sources with a known brightness, which can be used to infer luminosity distances[24],

$$d_L(z) = (1+z)\frac{1}{H_0}\int_0^z \frac{dz'}{H(z')} = (1+z)\frac{1}{H_0}\int_0^z \frac{dz'}{\sqrt{\Omega_r(1+z')^4 + \Omega_m(1+z')^3 + \Omega_\Lambda}}. \tag{1.35}$$

Cepheids are excellent standard candles due to their luminosity-period relation, which can be calibrated using geometric distances from parallax measurements. However, inferring an accurate H_0 value requires observing sources far enough so that the redshift is not affected by peculiar velocities. The cosmic distance ladder is the solution. This term refers to a collection of techniques that allow one to determine the distance to faraway objects by using previous steps as calibrators. In this context, type Ia supernovae (SNIa) are excellent distance indicators, as their luminosity can be standardised. They are among the brightest phenomena in the Universe, making them detectable even at high redshifts. A SNIa arises from a thermonuclear explosion of a white dwarf in binary systems, although its mechanism is still under study (Liu et al., 2023).

The luminosity distance is computed for SNIa using the standarized distance moduli defined as

$$\mu = 5\log(d_L/10\,\mathrm{pc}). \tag{1.36}$$

One of the ingredients for the distance moduli is the absolute magnitude, which needs to be calibrated using distance ladders. The most common calibrators are the Cepheids (Riess et al., 2022) and the Tip of Red Giant Branch (TRGB) stars (Freedman et al., 2024; Hoyt et al., 2025), which are used to estimate the distance to galaxies hosting SNIa.

Recently, gravitational waves (GW) emitted during neutron star and black hole binary mergers have begun to be used as standard sirens. They are absolute distance indicators, as they do not need to rely on any calibrators to infer the luminosity distance (Soares-Santos et al., 2019). To construct a redshift-distance relationship and measure H_0, the redshift is also needed, and for that, the host galaxy needs to be identified. For some GW events, the host galaxy can be identified from the electromagnetic counterpart emission (Abbott et al., 2017, 2021). However, there are sources such as binary black hole (BBH) merges without EM emission, which are called dark sirens, and even in those cases cosmological parameters can be inferred.

---

[24] It is the distance that accounts how the luminosity of a source is reduced due to the photon travel in an expanding Universe.

1.4 Current challenges

Although we have strong observational evidence for the ΛCDM model, there are still some open questions and tensions suggesting that the model is not complete. In this section, we summarize some of them (see Di Valentino et al. (2025) for a detailed review). We made a distinction between fundamental questions/challenges and tensions/anomalies.

Open questions

- What is the nature of dark matter? Although the SM of particle physics explains the fundamental particles and interactions successfully, it does not include dark matter, which accounts for almost 84% of the matter content in the Universe.
- What kind of fluid is dark energy? While understanding DM is already a challenge, explaining a fluid with a negative pressure is even harder. One of the main goals of the modern cosmology is to determine the nature and evolution of the DE, which constitutes the 69% of the total energy density of the Universe.
- Inflation: B-mode quest. The level of flatness and homogeneity cannot be explained without inflation. However, it has yet to be confirmed, and the precise details of how it occurred remain unknown. Upcoming CMB experiments, such as LiteBIRD (Hazumi et al., 2020) or The Simons Observatory (SO) (Ade et al., 2019), have been specifically designed to search for the polarisation B-mode (see Section 2.2.4), an imprint in the CMB of the GW generated during inflation.

Tensions

- **Lithium problem:** The amount of $^7\mathrm{Li/H}$ predicted by the BBN is a factor of 3.6 larger than the observed one (Cooke, 2024).
- **$\mathbf{H_0}$ tension:** The value of the Hubble constant inferred from early- and late-Universe does not agree. Early-time measurements come mainly from the CMB anisotropies, BBN abundance, and BAO, which actually do not directly constrain H_0, but rather it is derived from the model. The latest values inferred from the *Planck* mission, including lensing and BAO, is $H_0 = 67.81 \pm 0.38\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. Late-time measurements are obtained through different calibrations of the distance ladder. The SH0ES team uses parallax to calibrate Cepheids and Cepheids to calibrate SNIa. They infer a value around $H_0 = 73\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ with an approximate uncertainty of $1\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Riess et al., 2022). This is around 5σ tension. Additional late-time measurements come from the TRGB led by The Chicago Carnegie Hubble Program (CCHP). Their latest results seem to be compatible with the values inferred from CMB, $H_0 = 69\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Freedman et al., 2024). These results indicate that significant work remains to fully

understand the discrepancies among late-time cosmological measurements. In case of the tension between late- and early-time Universe being confirmed, resolving the discrepancy may require new physics beyond the standard ΛCDM model.

- **$\mathbf{S_8}$ tension:** This refers to a discrepancy in the S_8 parameter between the early- and late-time measurements. S_8 is commonly defined as $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$, where σ_8 is the root mean square of the amplitude of matter perturbations smoothed over $8\,h^{-1}$ Mpc. A larger value of σ_8 means that the matter is more strongly clustered on this scale. The low redshift probes, which include cosmic shear (Secco et al., 2022; Li et al., 2023) and galaxy clustering (Philcox & Ivanov, 2022), measured a lower S_8 value than expected from the evolution of the fluctuations observed in the CMB (Tristram et al., 2024; Madhavacheril et al., 2024). Recent analysis of the final data release from the Kilo-Degree Survey (KiDS) (Stölzner et al., 2025) shows measurements consistent with *Planck* values.

- **CMB anomalies:** Although the CMB data shows an excellent agreement with the prediction of the ΛCDM model, a set of unexpected features have been found in the temperature maps at the level of 2-3σ. The study of some of these anomalous features in the temperature and polarisation maps is the main topic of this thesis. See Section 3 for a review on this topic.

2

Cosmic Microwave Background

"The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction."

Sidney Coleman

Over the past 60 years, the CMB radiation has been a fundamental pillar of the cosmological standard model, offering a detailed picture of the Universe when it was only $380,000$ years old. It has been observed using balloons, as well as space- and ground-based telescopes, which ultimately led to the detection of anisotropies, tiny fluctuations in the intensity field reflecting perturbations in the density at the time of recombination. The statistical analysis of these fluctuations has become a cornerstone in modern cosmology, revealing key properties of the Universe, and a powerful cosmological probe as they are accurately described within the linear regime. The most recent space mission, *Planck*, in synergy with ground-base telescopes, has provided an accurate picture of these anisotropies over a broad range of angular scales. In recent years, the attention has shifted towards the precise observation of the CMB polarisation anisotropies, particularly the accurate measurement of the B-mode, a smoking-gun of cosmic inflation. This chapter provides a brief review of these topics, while the next one offers an overview of CMB anomalies, which constitute the main topic of the work I carried out during my PhD.

2.1 History

After the discovery of the CMB radiation in 1965 by A. Penzias and R. Wilson (Penzias & Wilson, 1965), the first goal was to measure its spectral energy distribution and compare it to a blackbody spectrum as predicted by the Big Bang model. The Far-Infrared Absolute Spectrophotometer (FIRAS) instrument aboard the COsmic Background Explorer (COBE) (Mather et al., 1994) satellite, launched by NASA in 1989, was the first mission to do so. FIRAS measured the CMB spectrum with high precision, finding excellent agreement with a blackbody spectrum at 2.726 ± 0.010 K.This temperature corresponds to the redshifting of the photons released at recombination around $z \approx 1100$ when the Universe was at a temperature close to 3000 K. Additionally, COBE/FIRAS was the first mission putting upper limits on the CMB spectral distortions, which are tiny deviations from a perfect blackbody spectrum induced by processes that inject energy in the CMB photons.

Although the CMB is remarkably homogeneous, it is not perfectly smooth. It presents small anisotropies, temperature variations, which provide a snapshot of the density perturbations in the photon-baryon plasma at the time of recombination. The strongest anisotropy is the dipole, which is actually mainly produced by the Doppler effect due to the motion of the Solar System with respect to the CMB rest frame. COBE/FIRAS measured a dipole with an amplitude of $3372 \pm 14\ \mu$K, which is compatible with the current value provided by *Planck* of $3366.6 \pm 2.7\mu$K (Planck Collaboration et al., 2020f). Apart from the dipole, there are fainter anisotropies, at the level of 10^{-5} K, of different angular sizes. These angular scales are conveniently described in terms of spherical harmonics, with smaller angular features corresponding to higher multipole moments ℓ. In particular, the shape of the angular power spectrum of these anisotropies, discussed in detail in Section 2.2.1, encodes information about the primordial fluctuations and the photon-baryon plasma physics before recombination, and therefore, provides valuable information about the composition and evolution of the Universe. The Differential Microwave Radiometer (DMR) instrument aboard COBE measured, for the first time, temperature variations on large angular scales (Wright et al., 1996).

The early 2000s were characterised by numerous ground-based and balloon-borne experiments that measured temperature anisotropies at smaller angular scales and across different frequency bands. BOOMERanG (Masi, 2002) was the first ballon-borne experiment dedicated to observing the CMB with high angular resolution. It provided the first clear detection of the first acoustic peak in the angular power spectrum and a hint of the second and third peaks. BOOMERanG was the first experiment discovering that the geometry of the Universe is close to flat. These results were confirmed by the MAXIMA balloon (Lee et al., 2001), and the DASI (Halverson et al., 2002) and the Very Small Array (VSA) (Scott et al., 2003) ground-based telescopes, which detected the second and third acoustic peaks. In particular, DASI was the first experiment detecting the CMB polarisation, which measured the amplitude of E-mode at the level of 4.9σ (Leitch et al., 2002), result that was improved later to 6.3σ (Leitch et al., 2005). The Cosmic Background Imager (CBI) (Sievers et al., 2003) and the Arcminute Cosmology Bolometer Array Receiver (ACBAR) (Kuo et al., 2004) ground-based telescopes provided complementary information by measuring higher-order peaks and the Silk damping tail (Silk, 1968) over a more extended multipole range. In particular, CBI revealed the the E-mode angular power spectrum shows peaks and valleys shifted half a cycle with respect to those observed in the temperature spectrum (Readhead et al., 2004).

After COBE, the next generation space satellite dedicated to the CMB observation was the Wilkinson Microwave Anisotropy Probe (WMAP) (Hinshaw et al., 2013), a NASA spacecraft operating from 2001 to 2010. It observed the CMB in 5 frequency bands spanning in the $23-94$ GHz range, providing a full-sky measurement of temperature anisotropies. WMAP played a key role establishing the ΛCDM model. It detected an anticorrelation in the TE cross-power spectrum on large angular scales, around $\ell \approx 150$, which is a signature of adiabatic superhorizon fluctuations at the time of recombination (Peiris et al., 2003), and a prediction of the inflationary models.

After COBE and WMAP, the third, and currently the latest, generation of space satellite for CMB was the *Planck* mission (Planck Collaboration et al., 2020a), a European Space Agency (ESA) spacecraft operating from 2009 to 2013. It carried two instruments: the Low Frequency Instrument (LFI), observing at 30, 44, and 70 GHz, and the High Frequency Instrument (HFI), observing at 100, 143, 217, 353, 545, and 857 GHz, with a combined resolution of 5 arcmin. Thanks to this, *Planck* provided cosmic variance limited (see Section 2.2.1) measurements of the temperature power spectrum up to $\ell \approx 1600$, and a highly precise determination of the cosmological parameters. There have been four official data releases: *Planck* 2013 (PR1), 2015 (PR2), 2018 (PR3), and the most recent one PR4/NPIPE [1] (Planck Collaboration et al., 2020f). Figure 2.1 shows the *Planck* 2018 `SMICA` temperature map.

Regarding ground-based telescopes, there are several ongoing and future experiments with the aim of measuring with high precision the CMB polarisation, particularly the B-mode, on degree and sub-degree angular scales. These are the BICEP/Keck Array (The BICEP/Keck Collaboration et al., 2024), the Simons Observatory (SO) (Ade et al., 2019), the South Pole Telescope[2] (SPT) (Zebrowski et al., 2025), Atacama Cosmology Telescope (ACT) (Louis et al., 2025), Polarbear (Adachi et al., 2022), and CMB-S4[3] (Abazajian et al., 2016). In synergy with these experiments, LiteBIRD (LiteBIRD Collaboration et al., 2023) will be the next generation satellite mission to measure the CMB polarisation on large angular scales. Figure 2.2 shows the current state of the art, highlighting the synergy between space missions and ground-based experiments in reconstructing both the large and small scales of

---

[1] In PR4, the two *Planck* instruments have been processed jointly which has effectively reduced the large-scale systematics on polarisation data.

[2] The first goal for the SPT was a survey to search for clusters of galaxies using the Sunyaev-Zel'dovich (SZ) effect.

[3] Recently, the U.S. Department of Energy (DOE) and the National Science Foundation (NSF) have jointly decided that they can no longer support the CMB-S4 Project.

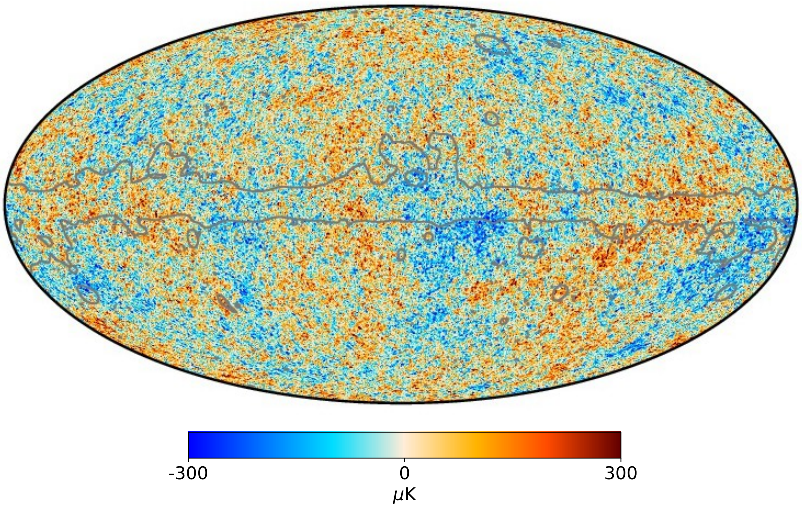

FIGURE 2.1: *Planck* 2018 SMICA temperature map. Grey line shows the *Planck* mask, which is the region where the residuals from foreground emission are expected to be non negligible. This region is inpainted for illustrative purposes. Figure adapted from Planck Collaboration et al. (2020a).

the anisotropies.

2.2 Theory of CMB anisotropies

The CMB is the unique cosmological observable providing direct information about the Universe when it was just $380,000$ years old. Studying the correlations of its anisotropies and, more generally, their statistical properties, offers a powerful tool to describe the early Universe and constrain the standard cosmological model. For instance, the angular correlations above degree scale probe directly the primordial quantum conditions due to their superhorizon nature at the time of recombination. On the other hand, sub-degree angular correlations probe the propagation of acoustic waves in the primordial photon-baryon plasma, and the damping tail caused by the photon diffusion. These are the “primordial anisotropies", which are generated at the last scattering surface during decoupling. Apart from them, there are also the so-called “secondary anisotropies", which refer to signals generated after decoupling due to physical processes affecting the photons on their way towards us. Weak gravitational lensing, caused by the deflection of the photons due to the matter distribution, the Sunyaev-Zel’dovich effect, which produces distortions in the CMB spectra, and the integrated Sachs-Wolfe (ISW) effect, which arises from the time variation of gravitational potentials along line of sight, are some examples.

The following sections provide a brief review on the theory of CMB anisotropies. Section 2.2.1 introduces the statistics used to describe the temperature anisotropies. The origin and evolution of the anisotropies is presented in Section 2.2.2, including a brief description of the secondary anisotropies. Polarisation anisotropies are introduced in Section 2.2.3, and

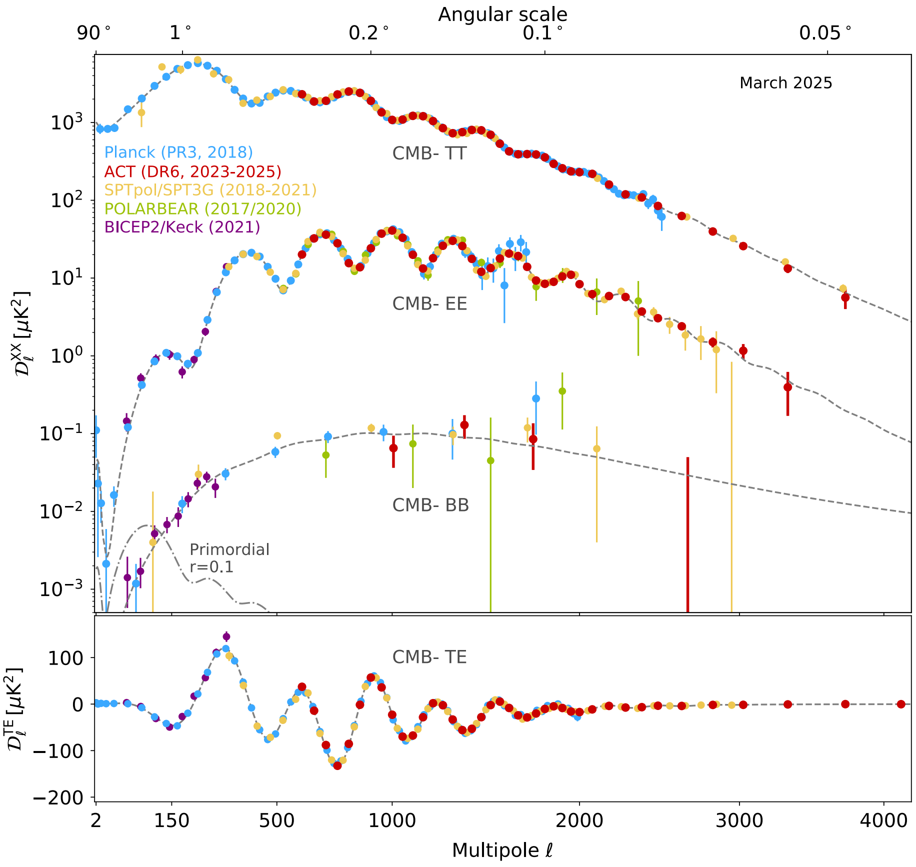

FIGURE 2.2: Compilation of CMB angular power spectrum measurements with different experiments. It includes temperature (TT), polarisation E-and B-modes, and TE cross power spectra. The dashed grey line corresponds to the best-fit $\Lambda\mathrm{CDM}$ model to the combination of ACT with the *Planck* large-scales data. Figure adapted from Louis et al. (2025).

finally, Section 2.2.4 summarizes the main challenges of the CMB.

2.2.1 Statistics of CMB temperature anisotropies

The CMB shows small perturbations of the order of 10^{-5} K above the background temperature, $\bar{T}_{\rm CMB} = 2.7255 \pm 0.0006$ K (Fixsen, 2009). Therefore, the CMB temperature measured by an observer in a direction $\hat{\mathbf{n}} = (\theta, \phi)$ in the sky is equal to

$$T(\hat{\mathbf{n}}) = \bar{T}_{\rm CMB}[1 + \Theta(\hat{\mathbf{n}})], \tag{2.1}$$

where $\Theta(\hat{\mathbf{n}})$ is the fractional temperature fluctuation. As already mentioned, the largest anisotropy in the CMB is a dipolar feature that comes from the motion of the solar system with respect to the CMB rest frame,

$$\delta T(\hat{\mathbf{n}})/T = \hat{\mathbf{n}} \cdot \mathbf{v} = v \cos\theta. \tag{2.2}$$

For $\delta T = 3366\,\mu$K (Planck Collaboration et al., 2020f) the inferred velocity is $v \approx 370$ km/s. Once the mean temperature and the dipole are subtracted, the anisotropy field can be statistically studied.

Since the anisotropies originate from primordial quantum fluctuations of a Gaussian field (see Section 2.2.2), they are fully characterised by their two-point correlation function,

$$C(\theta) = \langle \Theta(\hat{\mathbf{n}}) \Theta(\hat{\mathbf{n}}') \rangle \,, \tag{2.3}$$

where $\cos\theta = \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'$, and the average is over an ensemble of Universes.

An equivalent and more common description of the statistical properties of the temperature anisotropies is provided in harmonic space. Given that the anisotropies are defined on the celestial sphere, it is convenient to expand the temperature field in a spherical harmonic basis,

$$\Theta(\hat{\mathbf{n}}) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \qquad a_{\ell m} = \int \Theta(\hat{\mathbf{n}}) Y^*_{\ell m}(\hat{\mathbf{n}}) d\Omega, \tag{2.4}$$

where the coefficients $a_{\ell m}$ are the multipole moments. The sum starts at $\ell = 2$, the quadrupole, because $\ell = 0, 1$ corresponds to the monopole ($\bar{T}_{\rm CMB}$) and the dipole, respectively. All the cosmological information is encoded in the distribution of these coefficients. They follow a Gaussian distribution with zero mean and variance given by the angular power spectrum, C_ℓ, which is the equivalent in the Fourier space of the two-point correlation function in real space,

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{\ell \ell'} \delta_{m m'} C_\ell^{TT}, \qquad C(\theta) = \sum_\ell \frac{2\ell + 1}{4\pi} C_\ell^{TT} P_\ell(\cos\theta), \tag{2.5}$$

where the average is done over an ensemble of Universes, and P_ℓ are the Legendre polynomials. The Kronecker deltas reflect two key properties: statistical isotropy, implying that the power spectrum is independent of the magnetic quantum number (m); and the statistical independence of different multipole moments. Essentially, the C_ℓ's quantify the power of temperature fluctuations in the CMB on different angular scales. A higher C_ℓ for a particular ℓ means that there are larger temperature variations, more pronounced hot and cold spots, at that angular size. Usually, it is common to work with D_ℓ's instead of C_ℓ's as shown in Figure 2.2,

$$D_\ell = \frac{\ell(\ell+1)}{2\pi} C_\ell. \tag{2.6}$$

So far, the two-point correlation function and the angular power spectrum have been defined as an average over an ensemble of Universes. However, we live in one Universe and we can only observe that realisation. The solution is the statistical isotropy. Given that all the $a_{\ell m}$'s for a given ℓ have the same variance, for a fixed multipole we have $2\ell+1$ moments to estimate the corresponding C_ℓ,

$$\hat{C}_\ell^{TT} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \tag{2.7}$$

This estimator is unbiased, but it has a non-zero sampling variance which is commonly known as cosmic variance,

$$\frac{\Delta C_\ell}{C_\ell} = \sqrt{\frac{1}{\ell+0.5}}. \tag{2.8}$$

As expected, the uncertainty is larger at low multipoles because there are less independent moments to build up the estimator.

Finally, it is worth mentioning that the effect of the solar system motion with respect to the CMB rest frame also induces second-order effects that lead to distortions at both low and high multipoles. These include a frequency-dependent dipolar-modulation of the CMB anisotropies and an aberration effect on small angular-scales (Challinor & van Leeuwen, 2002).

2.2.2 From quantum fluctuations to CMB

This section provides a brief introduction to the initial curvature perturbations generated during inflation, and presents the key equation that relates these primordial perturbations to the observed power spectrum of CMB anisotropies. The objective is to offer a qualitative explanation for the origin of the (scalar) CMB anisotropies without entering into the full details, which would require the complete derivation and solution of the perturbed Boltzmann-Einstein equations. Typically, this task is performed using specific software such as `camb`[4](Lewis et al., 2000) or `CLASS`[5](Blas et al., 2011) Boltzmann solvers. All of this is done at linear order, consistent with the small amplitude of the anisotropies observed in the CMB. For a more detailed description, the reader is referred to Baumann's textbook Baumann (2022) or Ma & Bertschinger (1995); Lesgourgues (2013).

Primordial perturbations

Inflation theory not only provides a solution to the horizon and flatness problems but also offers a natural mechanism for the generation of primordial scalar and tensor perturbations. The scalar perturbations, in particular, serve as the initial seeds for the formation of structures in the late Universe.

In linear perturbation theory, scalar perturbations in the energy-momentum tensor are coupled to scalar metric perturbations through the linearized Einstein equations. These together with the continuity and Euler equations determine the whole evolution. However, the perturbation variables are sensitive to the choice of coordinates in both space and time. This freedom, known as the gauge freedom, means that, depending on the coordinate choice, the energy density perturbations can be removed and transferred into metric perturbations or vice versa. This issue can be addressed either by fixing a specific gauge and consistently tracking all perturbations, or by working with gauge-invariant quantities. A particularly useful example is the curvature perturbations, ζ and $\mathcal{R}$, which for adiabatic perturbations take the following form in Newtonian gauge,

$$\zeta = \Phi - \frac{\delta\rho}{3(\bar{\rho}+\bar{P})}, \qquad \mathcal{R} = \Phi - \mathcal{H}v, \tag{2.9}$$

where Φ is the Newtonian gravitational potential[6] and $\mathcal{H} = aH$ is the conformal Hubble parameter. Both curvature perturbations become equal on superhorizon scales, i.e., on wavelengths larger than the Hubble radius ($k << aH$), and more importantly, they are conserved for adiabatic perturbations. This makes $\mathcal{R}$ a perfect variable to connect the inflationary era to the post-inflationary Universe.

As discussed earlier, inflation is driven by the inflaton field, $\phi(t)$. In addition to its classical behavior, it also exhibits small quantum fluctuations above the classical background arising from the uncertainty principle,

$$\phi(t,\vec{x}) = \bar{\phi}(t) + \delta\phi(t,\vec{x}). \tag{2.10}$$

Expressing the curvature perturbation $\mathcal{R}$ as a function of the inflaton perturbation, we

---

[4] https://camb.readthedocs.io/en/latest/

[5] http://class-code.net/

[6] It is also one of the Bardeen potentials, which is one of the gauge-invariant variables for the metric perturbations.

have

$$\mathcal{R} = \Phi - \frac{\mathcal{H}}{\dot{\phi}}\delta\phi. \tag{2.11}$$

By considering the flat gauge ($\Phi = 0$), it is possible to relate the curvature perturbation to the inflaton one. Since these inflaton fluctuations can be interpreted as local time shifts in the evolution of the background, in this picture primordial fluctuations arise from spatial variations in the time at which inflation ends. Therefore, different regions of space undergo slightly different amounts of inflation, leading to small variations in their energy densities.

After inflation, inflaton field decays to the SM particles during reheating. In the single-field inflation model, the initial condition for the perturbations of all the cosmological species is determined from the same curvature perturbation $\mathcal{R}$. This is one of the main predictions of the simplest inflationary model, the primordial fluctuations are purely adiabatic, a prediction that has been tested with the *Planck* data (Planck Collaboration et al., 2020e). This condition can be expressed, for two species a and b, as

$$\frac{\delta_a}{1+\omega_a} = \frac{\delta_b}{1+\omega_b}. \tag{2.12}$$

where $\delta_i \equiv \delta\rho_i/\rho_i$ is the density contrast of component i, and $\omega_i \equiv P_i/\rho_i$ is its equation of state parameter, with i = a, b.

We now turn to the statistical properties of curvature perturbations. The evolution of the inflaton fluctuations can be obtained by perturbing at linear order the metric and the inflaton field in the Klein-Gordon equation. This gives rise to the well-known Mukhanov-Sasaki equation (Baumann, 2022),

$$f'' + \left(k^2 - \frac{z''}{z}\right) f = f'' + \omega(\tau, k)^2 f = 0, \tag{2.13}$$

where the Mukhanov variable is defined as $f \equiv a\delta\phi$ and $z \equiv \frac{a\bar{\phi}'}{\mathcal{H}}$, with $\bar{\phi}'$ denoting the derivative of the homogeneous field with respect the conformal time. This is the equation of a harmonic oscillator with a time dependent frequency. This equation can be evaluated in two regimes. At early times, in the limit where $k^2 >> z''/z$, all the scales are well inside the horizon and the fluctuations follow the standard harmonic oscillator with a fixed frequency $\omega = k$. As comoving Hubble radius shrinks, modes exit the horizon. In the limit where $k^2 << z''/z$, the solutions for the Mukhanov-Sasaki equation are a growing mode $f \propto z$ and a decaying mode $f \propto z^{-2}$. In particular, the growing mode corresponds to the frozen curvature perturbations.

The power spectra of the curvature perturbation $\mathcal{R}$ is given by

$$\Delta^2_{\mathcal{R}} = \left(\frac{H}{\dot{\bar{\phi}}}\right)^2 \Delta^2_{\delta\phi}, \tag{2.14}$$

where $\Delta_{\delta\phi}$ denotes the spectrum of the Bunch–Davies vacuum quantum fluctuations of the inflaton field. Since the origin of the primordial perturbations are quantum vacuum fluctuations, it is necessary to quantize the Mukhanov variable following the canonical quantization from quantum mechanics. In this picture, Mukhanov-Sasaki equation describes a quantum oscillator. By imposing the Bunch-Davies vacuum as initial condition, the spectrum of inflaton fluctuations is found to be

$$\Delta^2_{\delta\phi} \approx \left(\frac{H(t)}{2\pi}\right)^2\Bigg|_{k=aH(t)}. \tag{2.15}$$

For each Fourier mode, the spectrum is evaluated at the time when it crosses the horizon. In the limit case where H does not evolve in time, the expected spectrum should be scale-invariant, but this would mean eternal inflation. Since H is evolving, this leads to a slight scale dependence. In particular, under the slow-roll approximation, the power spectrum is given by

$$\Delta^2_{\mathcal{R}}(k) = \frac{1}{8\pi^2\epsilon}\frac{H^2}{M^2_{\rm Pl}}\Bigg|_{k=aH}, \tag{2.16}$$

where $M_{\rm Pl}$ is the reduced Planck mass. This expression provides the amplitude of quantum-generated curvature perturbations, evaluated at the moment of horizon crossing. The primordial power spectrum is usually modelled as

$$\Delta^2_{\mathcal{R}}(k) = A_s\left(\frac{k}{k_*}\right)^{n_s-1}, \qquad A_s = \frac{1}{8\pi^2\epsilon_*}\frac{H^2_*}{M^2_{\rm Pl}} \tag{2.17}$$

where $\mathcal{A}_s$ and n_s are the cosmological parameters described in 1.2.1, and H_* and ϵ_* are evaluated at the time when the reference scale k_* crosses the horizon. The near scale-invariance of this spectrum reflects the quasi-de Sitter expansion during inflation, and its slight tilt encodes information about the slow-roll parameters

$$n_s - 1 \equiv \frac{d\ln\Delta^2_{\mathcal{R}}(k)}{d\ln k} \approx -2\epsilon - \kappa, \qquad \kappa = \frac{\dot{\epsilon}}{H\epsilon}. \tag{2.18}$$

Current constraints exclude the exact scale-invariant scenario by $\approx 8.8\sigma$ (Tristram et al., 2024). Since the slow-roll parameters vary in time, n_s could also exhibit time evolution, which is parametrized by the spectral running index defined in 1.25. Therefore, n_s is evaluated at the time when the reference scale k_* crosses the horizon.

To summarize, the quantum fluctuations of the inflaton field ($\delta\phi$) are stretched to macroscopic scales by inflation, generating the cosmological curvature perturbations ($\mathcal{R}$). This process is a consequence of the rapid accelerated expansion of space during which the comoving Hubble radius shrinks, allowing perturbation modes to exit the horizon, freeze, and become classical fluctuations. Once outside the Hubble radius, these modes remain frozen until they re-enter during the post-inflationary expansion.

Inflation also predicted a spectrum of primordial gravitational waves that arise from the freezing and amplification of the quantum fluctuations in the tensor perturbations of the spacetime metric. It can be probed that the action for the tensor perturbations is just two copies of the action of a scalar field, the same for each of the two polarisation modes ($h_+, h_\times$). By solving the corresponding Mukhanov-Sasaki equation, where now the Mukhanov variables are $f_{+,\times} = \frac{aM_{\rm Pl}}{\sqrt{2}} h_{+,\times}$, it can be found that the power spectrum of tensor fluctuations is

$$\begin{aligned}\Delta_h^2(k) &= \langle h_{ij} h^{ij} \rangle = 2\Delta_{h_+}^2(k) + 2\Delta_{h_\times}^2(k) = \\ &= 2\left(\frac{2}{M_{\rm Pl}}\right)^2 \Delta_{\delta\phi}^2(k)\bigg|_{k=aH} = \frac{2}{\pi^2}\left(\frac{H}{M_{\rm Pl}}\right)^2\bigg|_{k=aH}.\end{aligned} \tag{2.19}$$

This power spectrum is also modelled as a power law,

$$\Delta_h^2(k) = \mathcal{A}_t \left(\frac{k}{k_*}\right)^{n_t}, \qquad \mathcal{A}_t \equiv \frac{2}{\pi^2}\frac{H_*^2}{M_{\rm Pl}^2}, \qquad n_t \equiv -2\epsilon_*, \qquad r \equiv \frac{\mathcal{A}_t}{\mathcal{A}_s}, \tag{2.20}$$

where we define the tensor-to-scalar ratio as the ratio between both amplitudes. An important aspect is that the amplitude is a direct measure of the expansion rate during inflation and the energy scale of inflation.

From primordial perturbations to now

As shown in the previous section, the power spectra of the primordial curvature perturbation is featureless and well described by a simple power law. However, the CMB angular power spectrum shows a non-trivial structure characterised by a series of peaks and valleys. Actually, it is the transfer function $\mathcal{T}_\ell(k)$ that relates $\Delta_{\mathcal{R}}^2(k)$ to C_ℓ by capturing the evolution of perturbations in the photon-baryon plasma, the free streaming of photons after decoupling and their projection onto the sky,

$$C_\ell = 4\pi \int \mathrm{d}\ln k \; \mathcal{T}_\ell^2(k) \Delta_{\mathcal{R}}^2(k). \tag{2.21}$$

While the primordial spectrum encodes information about $\mathcal{A}_s$ and n_s, the transfer function contains information about the baryon and cold dark matter content, as well as the geometry of the Universe.

After decoupling, photons free-stream and maintain their distribution function, a Bose-Einstein spectrum perturbed by small temperature fluctuations, $\Theta(\hat{\mathbf{n}})$. The Boltzmann equation links the evolution of this temperature fluctuations to the evolution of the photon energy in an inhomogeneous spacetime, which is governed by the geodesic equation,

$$\frac{d\Theta}{d\eta} = \frac{d\,\ln(aE)}{d\eta} = -\frac{d\Psi}{d\eta} + \Phi' + \Psi'. \tag{2.22}$$

By integrating this equation from recombination up to the present time, the relation between the observed temperature fluctuations in the $\hat{\mathbf{n}}$ direction and the fluctuations in the last-scattering surface becomes

$$\frac{\delta T}{\bar{T}}(\hat{\mathbf{n}}) = \left(\frac{1}{4}\delta_\gamma + \Psi\right)_* - (\hat{\mathbf{n}} \cdot \mathbf{v}_b)_* + \int_{\eta_*}^{\eta_0} d\eta(\Phi' + \Psi'), \tag{2.23}$$

where * indicates that those terms are evaluated at the last-scattering surface. The three contributions in the right-hand side are the Sachs-Wolfe term, the Doppler term, and the integrated Sachs-Wolfe term, respectively. So far, instantaneous recombination has been assumed. This means that before decoupling photons are tightly coupled to baryons, so $\mathbf{v_b} \sim \mathbf{v}_\gamma$, and instantly they become free. A more careful calculation requires one to take into account the thickness of the last-scattering surface by including the visibility function that quantifies the probability of a photon being last scattered between η and $\eta + \delta\eta$. Let us briefly discuss each of the three effects.

- **Sachs-Wolfe (SW) effect (Sachs & Wolfe, 1967):** It combines the temperature anisotropy associated to the intrinsic photon fluctuations ($\delta_\gamma/4$) with the gravitational redshift (Ψ) due to the potential fluctuations at the last-scattering surface. According to the first one, regions with a larger (lower) photon density are slightly hotter (colder). The second effect indicates that an overdensity, which corresponds to $\Psi < 0$, leads to a temperature decrement as the photons lose energy climbing out of a potential well.

- **Doppler effect:** This accounts for the temperature anisotropies associated with the fluctuations in the bulk velocity of electrons at the recombination time. The local peculiar velocity of the electrons scattering photons in the last-scattering surface induces a Doppler shift in the photon energy, resulting in temperature fluctuations proportional to $\mathbf{p}_\gamma \cdot \mathbf{v}_e$. Here, $\mathbf{p}_\gamma$ denotes the momentum of the scattered photon and $\mathbf{v}_e$ the velocity of the electron, with can be substituted by $\mathbf{v}_b$.

- **Integrated Sachs-Wolfe (ISW) effect:** The last term accounts for the extra gravitational redshift due to the evolution of the metric potentials along the line-of-sight. A photon gains energy when it falls in a gravitational potential well and loses an equivalent amount when climbing it out. However, if the potentials evolve while the photon is crossing the well, the net effect will not be zero. The gravitational potentials evolve during the radiation- and dark-energy-dominated eras, remaining approximately constant during the matter-dominated era, and changing gradually during the transitions from radiation to matter domination and from matter to Λ domination. At early times, during the matter-dominated era, a residual radiation component contributed to the total energy density, which results in a non-zero early ISW effect. At late times, dark energy dominates the energy density, producing a late ISW effect.

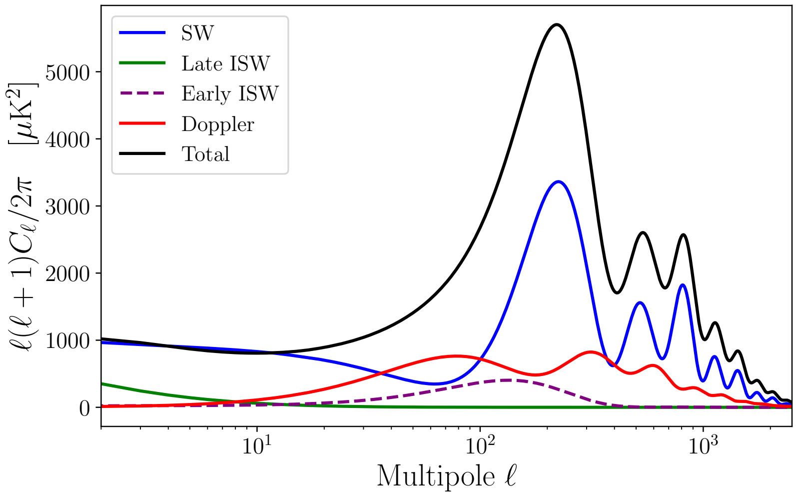

FIGURE 2.3: Contribution of the Sachs-Wolfe, Doppler, and early and late time integrated Sachs-Wolfe effects to the CMB temperature power spectrum using the following cosmological parameters: $H_0 = 67.81$, $\Omega_c h^2 = 0.1186$, $\Omega_b h^2 = 0.02229$, $A_s = 2.1073 \times 10^{-9}$, $n_s = 0.969$, $\tau = 0.06$. Note that the sum of individual terms does not add up to the total spectrum (black line) since the cross terms also contribute, positively or negatively. Figure made with A. Lewis' `camb.symbolic` package[7].

The contribution of each effect to the CMB power spectrum is shown in Figure 2.3. The free streaming of photons after decoupling relates the fluctuations on the last-scattering surface to the observed temperature variations in the sky (eq. 2.23). Determining the correlations between the temperature in different directions of the sky requires tracing the evolution of the photon perturbations in the primordial plasma from the initial conditions to the last-scattering surface. This is best studied in Fourier space where each Fourier mode evolves independently in the linear regime. The contributions from all modes are then summed and projected onto the sky. This entire process is encoded in $\mathcal{T}_\ell(k)$. Let us discuss briefly how these effects and the physics of the photon-baryon plasma affect different scales in the angular power spectrum. Large angular scales ($\ell < 100$) are created by superhorizon fluctuations, i.e. modes that were outside the Hubble horizon at the recombination time. The SW effect is the dominant contribution to the temperature anisotropies on these angular scales, with a small contribution from the late ISW. For the adiabatic initial conditions, since decoupling occurs during matter era, the superhorizon limit implies $\delta_\gamma \approx -\frac{8}{3}\Phi$. Consequently, the CMB temperature fluctuations on large angular scales are,

$$\frac{\delta T}{\bar{T}}(\hat{\mathbf{n}}) \approx \frac{1}{5}\mathcal{R}_*. \tag{2.24}$$

This means that these scales did not evolve before decoupling and probe directly the initial conditions. Additionally, on these scales the gravitational redshift is more relevant than the intrinsic temperature, so hot (cold) spots actually correspond to underdensity (overdensity)

---

[7] https://camb.readthedocs.io/en/latest/ScalEqs.html

regions.

Multipoles above $\ell \approx 100$ are sourced by sub-Hubble horizon modes whose non-trivial evolution must be taken into account. Before recombination, the Hubble horizon grows with time, allowing Fourier modes of the primordial curvature perturbations to cross and enter the horizon. Once a mode enters, its evolution is governed by the dynamics of the tightly coupled photon-baryon plasma. Gravitational potential wells, sourced by dark matter, attract baryons and induce compression. At the same time, since photons are coupled to the baryons, they are dragged into these wells. The resulting increase in radiation pressure counteracts gravity, generating oscillations in the plasma, the sound waves propagating in the photon-baryon fluid. These dynamics can be described by the equation for the photon density contrast that takes the form of a damped driven harmonic oscillator. At recombination, modes with different wavelenghts are captured at different moments in their evolution. Modes that have reached maximum compression (rarefaction) at that time are imprinted as odd (even) peaks in the CMB angular power spectrum. Additionally, the Doppler effect contributes most significantly when the velocity is maximal, which typically occurs at the zero-crossing of the density contrast. This is the reason why the Doppler effect appears out of phase with respect to the SW effect in Figure 2.3.

Above $\ell \sim 1100$ the CMB fluctuations are exponentially damped. There are two reasons for this: the Silk damping (Silk, 1968) and the Landau damping. The first one is a consequence of the photon diffusion, which is important for modes deep inside the Hubble radius. This effect arises by going beyond the leading order in the tight-coupling approximation, which introduces a friction term that includes thermal conduction and effective photon viscosity. The second effect arises by considering that recombination is not instantaneous, which has an impact for short-wavelength fluctuations.

So far, we have considered the temperature fluctuations generated by primordial scalar perturbations. However, tensor perturbations can also produce temperature fluctuations because of the presence of primordial gravitational waves that affect the temperature anisotropies at the last-scattering surface. Taking into account that the amplitude of gravitational waves decays inside the horizon, this signal is mainly expected on large scales. In particular, the strongest constraints on the tensor-to-scalar ratio from *Planck* data alone come from large angular scales.

Excluding the ISW effect, we have described the primary anisotropies of the CMB, i.e. those generated at the last-scattering surface. However, as CMB photons travel towards us, they are further affected by various processes that give rise to secondary anisotropies. In the following, we describe the most relevant ones:

- **Reionisation:** After recombination, CMB photons propagate freely as there are no free electrons in the medium to be scattered with. This transparency persists through most of the cosmic history, until $z \sim 6$, when the ultraviolet light from the first stars and galaxies reionises the intergalactic medium. Using the definition of optical depth in Equation 1.23, which assumes that the scattering from decoupling to reionisation is negligible, the probability that a photon reaches us without being scattered is $e^{-\tau}$, while $(1 - e^{-\tau})$ is the probability of being scattered. Since the electrons are not relativistic, Thomson scattering redistributes photons' direction but does not modify their energy. As a result, the scattered photons only contribute to the monopole and lose their directional anisotropic imprint. The observed temperature today in a given direction is

$$T_0(\hat{\mathbf{n}}) = \bar{T}_0[1 + \Theta(\hat{\mathbf{n}})]e^{-\tau} + \bar{T}_0(1 - e^{-\tau}) = \bar{T}_0[1 + \Theta(\hat{\mathbf{n}})e^{-\tau}]. \tag{2.25}$$

This expression shows that the anisotropies are suppressed by a factor of $e^{-\tau}$, and therefore the power spectrum on scales smaller than the horizon at recombination ($\ell > 10$) is reduced by a factor of $e^{-2\tau}$. Since the optical depth simply rescales the amplitude of the power spectrum, it is highly correlated with the amplitude of primordial scalar perturbations, A_s. Breaking the degeneracy requires the measurement of the CMB polarisation.

- **Sunyaev-Zeldovich effect:** Among the processes that generate secondary anisotropies are the thermal (tSZ) and kinetic (kSZ) Sunyaev-Zeldovich effects (Sunyaev & Zeldovich, 1970), as well as gravitational lensing (Lewis & Challinor, 2006). Thermal SZ arises from the interaction of the CMB photons with the electrons in the hot gas inside galaxy clusters. As CMB photons go through the hot gas, inverse Compton scattering with the free electrons boosts their energy, resulting in a spectral distortion of the CMB blackbody spectrum. Accordingly, the tSZ signal appears in the CMB temperature maps as a cold (hot) spot at frequencies below (above) 217 GHz centred around the position of galaxy clusters. In contrast, the kSZ effect is produced by the bulk motion of the free electron gas. It mainly produces is a shift in the blackbody temperature that is proportional to the line-of-sight peculiar velocity of the scattering electrons ($v_{||}$),

$$\theta_{\mathrm{kSZ}}(\hat{\mathbf{n}}) = \sigma_T \int n_e v_{||} dl. \tag{2.26}$$

  In particular, a cluster moving away from the observer produces a negative shift, $\theta_{\mathrm{kSZ}}(\hat{\mathbf{n}}) < 0$, while an approaching cluster results in $\theta_{\mathrm{kSZ}}(\hat{\mathbf{n}}) > 0$. Both effects may have relativistic corrections when the electron temperature or velocities become large enough.

- **Gravitational lensing:** As CMB photons travel through the Universe, their trajectories can be deflected due to the large-scale structure. This effect, known as gravitational lensing, is predicted by the general relativity theory, which states that light trajectories are curved when passing near massive objects. Given that the deflection angles are small, this effect is typically studied in the weak lensing approximation. Essentially, lensing produces that a photon coming from the direction $\hat{\mathbf{n}}$ is deviated to the position $\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})$, where ϕ is the lensing potential, i.e. the line-of-sight projection of the gravitational potentials,

$$\phi(\hat{\mathbf{n}}) = -2\int_0^{\chi_*} \mathrm{d}\chi \frac{\mathrm{S_K}(\chi_* - \chi)}{\mathrm{S_K}(\chi_*)}\Psi(\chi\hat{\mathbf{n}}; \eta_0 - \chi). \tag{2.27}$$

In this equation, S_K is the transverse comoving distance defined in Equation 1.4, and χ_* and χ are the radial comoving distances to the last-scattering surface and to the position of the gravitational potential Ψ, respectively. The deflection angles cause distortions on the CMB power spectrum and add some non-Gaussianity. In particular, it smooths out the acoustic peaks on small angular scales. Additionally, it rotates the polarisation of the CMB photons in such a way that some amount of power from the E-mode is leaked into the B-mode (see Section 2.2.3). As will be shown later in Section 2.2.4, this acts as a foreground that complicates the detection of the primordial B-mode. The power spectrum of the lensing potential, $C_\ell^{\phi\phi}$, is also a cosmological probe by itself (Madhavacheril et al., 2024).

Impact of the cosmological parameters on the CMB temperature power spectrum

The CMB power spectrum has a rich structure, whose shape depends on the cosmological parameters. In this section, we show how these parameters affect the position and amplitude of the acoustic peaks, which allows us to constrain them by using the measured *Planck* power spectrum. The main problem of having such a multidimensional parameter space is the existence of degeneracies, which means that the effect of one parameter in the power spectrum can be mimicked by varying others in a certain way. These degeneracies require using external data or polarisation measurements to break them. Let us now consider the effect of some of the cosmological parameters in the temperature power spectrum.

We start with the parameter that governs the geometry of the Universe, Ω_k. By modifying this, we effectively modify the angular diameter distance to the last-scattering surface. The positions of the acoustic peaks are closely related to the angular size of the sound horizon at the recombination epoch[8], θ_*. This quantity, as shown in Equation 1.22, is defined as the ratio between the comoving size of the sound horizon, $r_s(z_*)$, and the angular diameter distance, $D_A(z_*)$. By fixing H_0 and the physical densities $\Omega_b h^2$ and $\Omega_c h^2$, we fix r_s. In this case, varying Ω_k, we modify θ_* through $D_A(z_*)$, and therefore, the positions of the acoustic peaks are shifted in multipole space. In particular, in an open (close) Universe, the physical scales are projected onto a smaller (larger) angular scales, shifting the peaks to larger (lower) multipoles. Note that in this case, Ω_Λ is adjusted accordingly to satisfy the Friedmann equa-

---

[8] The Fourier modes that most contribute to the acoustic peaks are the ones that are in the maximum and minimum of the oscillation.

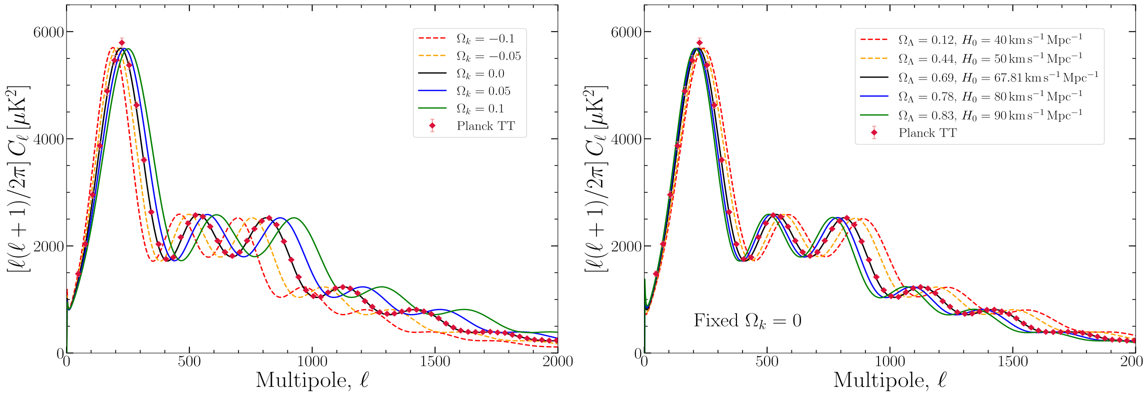

FIGURE 2.4: Impact of spatial curvature Ω_k (*left panel*) and dark energy density Ω_Λ (*right panel*) on the temperature angular power spectrum. The red dots represent the binned power spectrum as measured by the *Planck* mission, while the black curve corresponds to the best-fit ΛCDM model. The other cosmological parameters are fixed to the best-fit model. Note that in the *right panel*, H_0 cannot be fixed independently when the physical matter densities and Ω_k are fixed. Plots generated using `camb` (Lewis et al., 2000; Lewis & Challinor, 2011) software[9].

tion (see Eq. 1.18). The shift in the peak locations can also be mimicked in a flat Universe by varying the dark energy content. As curvature does, dark energy primarily affects the CMB by altering the distance to the last-scattering surface. They are partially degenerate as both influence the late-time Universe. This is shown in Figure 2.4, which also includes the binned *Planck* measurements.

We now focus on A_s, n_s, and τ. Changing A_s means rescaling all C_ℓs by an overall factor. The same effect can be obtained by modifying the optical depth to reionisation. As shown in Equation 2.25, the effect of τ is a suppression of the anisotropies by a factor of $e^{-\tau}$. The large angular scales are not affected by τ as they have entered the horizon recently. However, these scales are dominated by the cosmic variance and cannot break the degeneracy by themselves. Additionally, shifting $n_s \to n_s + \alpha$ modifies the tilt of the primordial power spectrum (see Eq. 2.17). Figure 2.5 shows the impact of these three parameters in the C_ℓs.

Finally, we discuss the effect of the baryon and CDM densities. Both parameters mainly affect the pre-recombination physics by modifying the size of the sound horizon, $r_s(z_*)$. Consequently, they also produce a shift in the locations of the acoustic peaks through a modification of θ_*. Additionally, they have an impact on the expansion history and the angular diameter distance, which are more sensitive to the dark matter content due to its greater contribution to the total matter density. Both densities also affect the diffusion scale, and thus, the damping tail.

The baryon content is mainly constrained from the relative amplitude between the odd and even peaks. Essentially, baryons add inertia to the photon-baryon fluid, which enhances the compression (odd peaks) relative to the rarefactions (even peaks). Physically, this occurs because they contribute to the gravitational pull, but not to the radiation pressure.

---

[9] See `https://camb.readthedocs.io/en/latest/` for documentation.

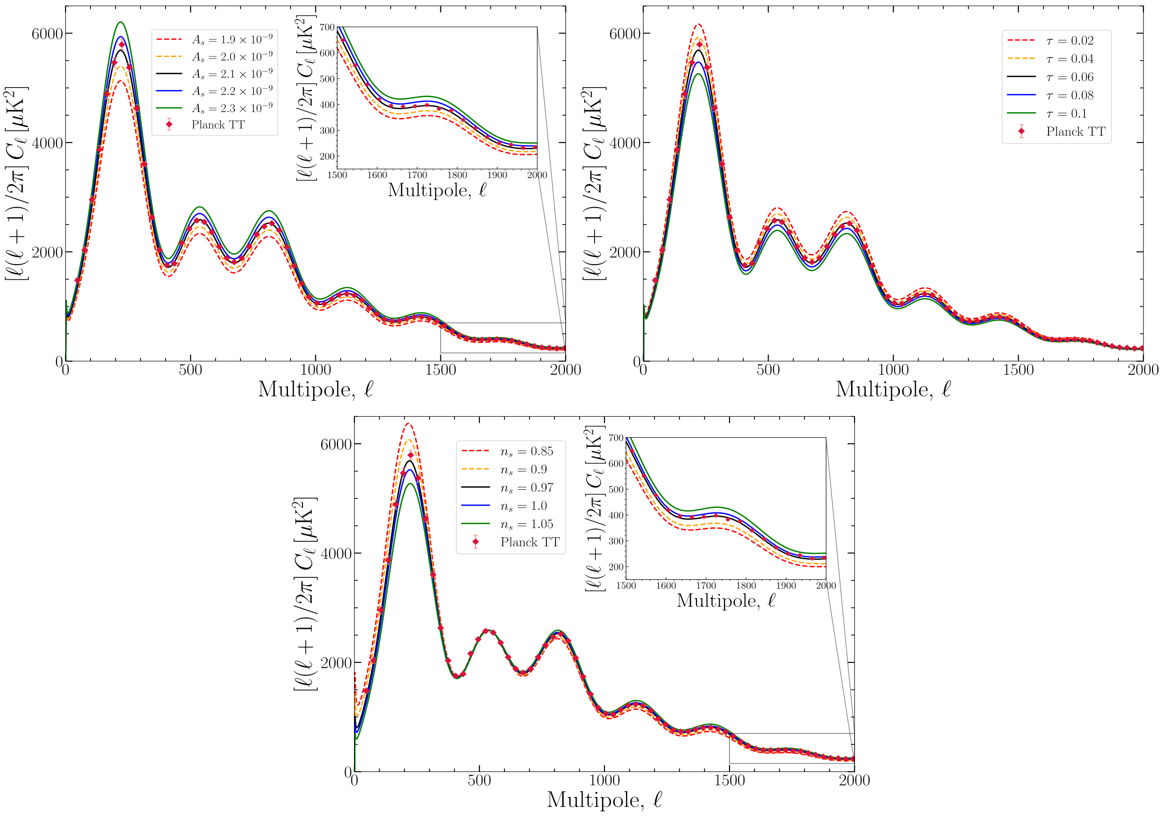

FIGURE 2.5: Impact of the amplitude (*top left panel*) and spectral index (*bottom panel*) of the primordial density fluctuations, and optical depth to reionisation (*top right panel*) on the temperature angular power spectrum. The red dots represent the binned power spectrum as measured by the *Planck* mission, while the black curve corresponds to the best-fit ΛCDM model. The other cosmological parameters are fixed to the best-fit model. Plots generated using `camb` (Lewis et al., 2000; Lewis & Challinor, 2011) software.

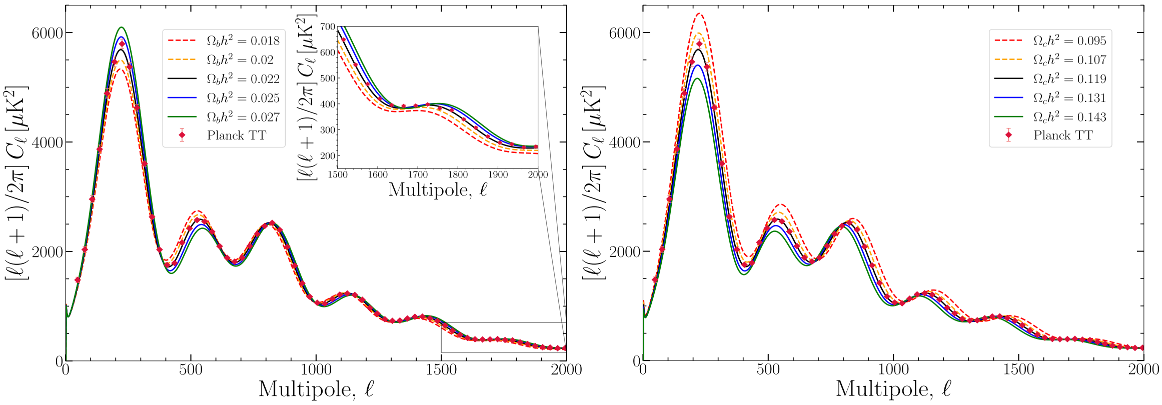

FIGURE 2.6: Impact of the physical baryonic (*left panel*) and cold dark matter densities (*right panel*) on the temperature angular power spectrum. The red dots represent the binned power spectrum as measured by *Planck* mission, while the black curve corresponds to the best-fit ΛCDM model. The other cosmological parameters are fixed to the best-fit model. Plots generated using `camb` (Lewis et al., 2000; Lewis & Challinor, 2011) software.

CDM, in turn, governs the gravitational wells into which baryons fall, and consequently, the driving force for the acoustic oscillations. In particular, reducing CDM content delays the matter-radiation equality, causing more modes to enter the horizon during the radiation era. During this epoch, the gravitational potential decays, resulting in a time-dependent driving force that enhances the amplitude of the oscillations. The shift in the equality time also affects the early ISW, boosting the height of the first acoustic peak. All of these effects are shown in Figure 2.6.

2.2.3 CMB polarisation

The CMB polarisation is an almost independent source of information that can be used, among other things, to better constrain the cosmological model, test CMB anomalies presented in the temperature data, and probe cosmic inflation through the polarisation B-mode. In this section, we briefly introduce the mechanism that polarised the CMB radiation, followed by a discussion of the statistical information it provides and the challenges that we face for the detection of primordial B-mode.

CMB polarisation is generated when an electron, scattering photons through Thomson scattering, experiences a quadrupolar anisotropy in the incoming radiation field (Hu & White, 1997; Cabella & Kamionkowski, 2004; Baumann, 2022).

The differential cross section of the Thomson scattering is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{3\sigma_T}{8\pi}(\hat{\epsilon}\cdot\hat{\epsilon}')^2, \tag{2.28}$$

where $\hat{\epsilon}$ and $\hat{\epsilon}'$ are the incident and scattered polarisation vectors, respectively. Under certain conditions, Thomson scattering can induce linear polarisation in the scattered radiation. If the radiation is isotropic, then the outgoing radiation will remain unpolarized as

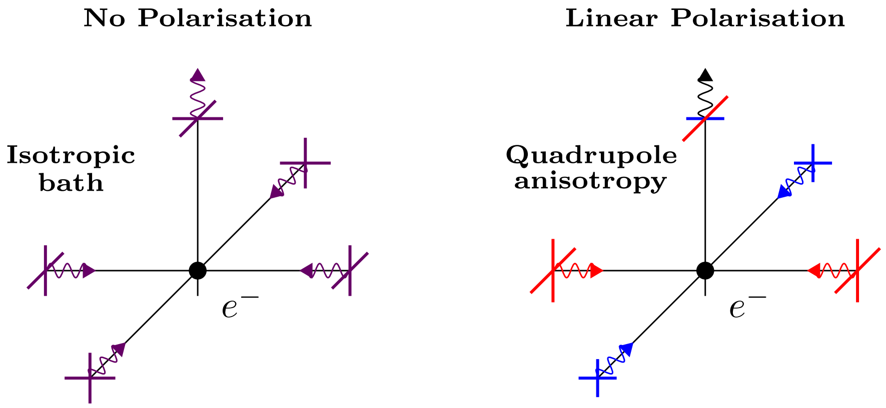

FIGURE 2.7: Schematic illustration of how CMB polarisation is generated through Thomson scattering. *Left panel:* If the incoming radiation is isotropic, i.e., it has equal intensity from all directions, the scattered radiation remains unpolarised. *Right panel:* In the presence of a quadrupolar anisotropy in the radiation field, the scattered radiation will be partially linearly polarised.

orthogonal polarisation stages, coming from directions separated by 90°, will balance each other. In particular, in the presence of a quadrupolar anisotropy of the radiation field, these components do not cancel perfectly, resulting in a net linear polarisation. This is shown in Figure 2.7, where the radiation coming from the left is larger than that coming from the top. From a mathematical point of view, given the angular dependence of the differential cross section of the Thomson scattering, the orthogonality of the spherical harmonics guarantees that only the quadrupole moment can generate polarisation (Baumann, 2022).

Before recombination, during the tight coupling regime, quadrupole anisotropies are suppressed by the frequent scattering of the photons. Just before recombination, the mean free path of photons increases, and the electrons start seeing a quadrupole moment in the radiation field. Consequently, polarisation is only generated near recombination, through Thomson scattering of a radiation field with a quadrupole moment at the last-scattering surface. Given the sharp drop in the free electron density, there is only a narrow window in which Thomson scattering can effectively produce polarisation. As a result, the CMB polarisation fraction is approximately 10%.

Statistics of CMB polarisation

The statistical properties of the CMB polarisation field can be described in a very similar way as for temperature. The polarisation of electromagnetic radiation propagating in the z direction is described by the Stokes parameters

$$\begin{aligned} I &\equiv |E_x|^2 + |E_y|^2, \\ Q &\equiv |E_x|^2 - |E_y|^2, \\ U &\equiv 2|E_x||E_y| \cos\varphi, \\ V &\equiv 2|E_x||E_y| \sin\varphi. \end{aligned} \tag{2.29}$$

The first Stokes parameter, I, measures the intensity[10], while Q, U, and V determine the polarisation state. For unpolarised light, $Q = U = V = 0$, whereas fully polarised radiation satisfies the relation $I^2 = Q^2 + U^2 + V^2$. In particular, the Stokes parameter V describes circular polarisation, while Q and U characterise the linear polarisation. Specifically, Q represents the difference between the intensity along the x and y axes, and U quantifies the difference around two axes rotated by 45 degrees. In the CMB case, due to the nature of the Thomson scattering, the radiation is linearly polarised, with a polarisation fraction of around 10%,

$$\frac{P}{I} = \frac{\sqrt{Q^2 + U^2}}{I} < 10\%, \tag{2.30}$$

and has no V component (see Eimer et al. (2024) for the current constraint on C_ℓ^{VV}). For each point in the sphere, the linear polarisation is described by the Stokes parameters Q and U, which provide the amplitude $P^2 = Q^2 + U^2$ and orientation $\phi = \frac{1}{2}\arctan(U/Q)$. The main problem is that these quantities are not independent of the choice of coordinates. Under a rotation of angle ψ in the tangent plane at the direction $\hat{\mathbf{n}}$, Q and U transform as

$$\begin{aligned} Q' &= Q\cos 2\psi + U \sin 2\psi, \\ U' &= -Q \sin 2\psi + U \cos 2\psi, \end{aligned} \tag{2.31}$$

or in a compact form,

$$Q' \pm iU' = e^{\mp 2i\psi}(Q \pm iU). \tag{2.32}$$

This shows that the polarisation field behaves as a spin-2 quantity, and therefore, its decomposition in the spherical harmonic space must be made in terms of the spin-weighted spherical harmonics,

$$(Q \pm iU)(\hat{\mathbf{n}}) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} {}_{\pm 2}a_{\ell m \pm 2}Y_{\ell m}(\hat{\mathbf{n}}), \tag{2.33}$$

---

[10] Intensity fluctuations can be converted to temperature fluctuations of a blackbody using the Planck's law. It is more practical to describe the amplitude of the radiation in terms of temperature since, in that way, CMB is signal is frequency independent.

where ${}_{\pm 2}a_{\ell m}$ coefficients are calculated like

$$ {}_{\pm 2}a_{\ell m} = \int (Q \pm iU)(\hat{\mathbf{n}})_{\pm 2}Y^{*}_{\ell m}\mathrm{d}\Omega. \tag{2.34}$$

Note that the summation starts in $\ell = 2$ as polarisation has neither monopole nor dipole.

Instead of describing CMB polarisation in terms of Q and U, it is more convenient to decompose it in terms of E- and B-mode, which are rotationally invariant quantities. The E/B decomposition is similar to the decomposition of a vector field in terms of a curl-free gradient part and a divergence-free curl part, which also explains the names of E-mode and B-mode in an analogy with the electrostatic field decomposition in electric $\vec{E}$ and magnetic $\vec{B}$ fields. The harmonic coefficients in Eq. 2.34 can be linearly combined to get the coefficients for the E/B-modes (Zaldarriaga & Seljak, 1997),

$$\begin{aligned} e_{\ell m} &= -({}_{2}a_{\ell m} + {}_{-2}a_{\ell m})/2, \\ b_{\ell m} &= -({}_{2}a_{\ell m} - {}_{-2}a_{\ell m})/2i. \end{aligned} \tag{2.35}$$

These two quantities behave differently under a parity transformation. While E-mode is parity-even ($E(-\hat{\mathbf{n}}) = E(\hat{\mathbf{n}})$) and remains unchanged, B-mode is parity-odd ($B(-\hat{\mathbf{n}}) = -B(\hat{\mathbf{n}})$) and changes sign. These harmonic coefficients can be combined with the spin-0 spherical harmonics to get the scalar E-mode and pseudo-scalar B-mode maps,

$$\begin{aligned} E(\hat{\mathbf{n}}) &= \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} e_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \\ B(\hat{\mathbf{n}}) &= \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} b_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}). \end{aligned} \tag{2.36}$$

As for temperature, the statistical properties of the CMB radiation can be fully described, under the Gaussian assumption, in terms of the variances of the spherical harmonic coefficients,

$$\begin{aligned} \langle e_{\ell m} e^{*}_{\ell' m'} \rangle &= \delta_{\ell\ell'} \delta_{mm'} C_{\ell}^{EE}, \\ \langle b_{\ell m} b^{*}_{\ell' m'} \rangle &= \delta_{\ell\ell'} \delta_{mm'} C_{\ell}^{BB}, \\ \langle a_{\ell m} e^{*}_{\ell' m'} \rangle &= \delta_{\ell\ell'} \delta_{mm'} C_{\ell}^{TE}, \end{aligned} \tag{2.37}$$

where C_{ℓ}^{EE} and C_{ℓ}^{BB} are the angular power spectra of E-mode and B-mode, respectively, and C_{ℓ}^{TE} the cross spectrum between temperature and E-mode. The other two possible cross spectra, C_{ℓ}^{TB} and C_{ℓ}^{EB}, vanish because B-modes have opposite parity to T and E, and the standard cosmological model does not include any parity-violating processes. Therefore, only four power spectra are needed to fully characterise the statistical properties of the CMB.

Scalar perturbations do not produce B-modes, but these are produced by primordial gravitational waves and by the conversion from E-to-B due to gravitational lensing. Both will be briefly discussed in Section 2.2.4. In absent of tensor modes, the E-mode power spectrum and its cross spectrum with the temperature anisotropies are

$$\begin{aligned} C_\ell^{EE} &\propto \int_0^\infty \mathrm{d}\ln\mathrm{k}|\mathcal{E}_\ell(\mathrm{k})|^2\Delta^2_{\mathcal{R}}(\mathrm{k}), \\ C_\ell^{TE} &\propto \int_0^\infty \mathrm{d}\ln\mathrm{k}|\mathcal{T}_\ell(\mathrm{k})\mathcal{E}^*_\ell(\mathrm{k})|\Delta^2_{\mathcal{R}}(\mathrm{k}), \end{aligned} \tag{2.38}$$

where $\mathcal{T}_\ell(k)$ and $\mathcal{E}_\ell(k)$ are the intensity and polarisation transfer functions. In particular, the polarisation transfer function depends on the quadrupole moment of the photon field at the last-scattering surface.

Breaking the degeneracies

In the previous section, we mentioned that the degeneracies between parameters can be broken using the CMB polarisation data. Here, we present an example for the $\mathcal{A}_s$-τ degeneracy. Figure 2.8 shows the impact of these parameters on the E-mode power spectrum. Although the effect seems quite similar at multipoles above $\ell > 10$, there is a clear distinction on large scales. The first feature we observe is that the acoustic peaks in the E-mode spectrum are out of phase with those in the temperature spectrum. This is because the E-mode is sourced by a gradient in the photon-baryon fluid velocity. Apart from that, we also observe a bump at low multipoles, which is sourced by the re-scattering of quadrupole anisotropy by the free-electrons during reionisation. This feature depends directly on the optical depth to reionisation, making the polarisation E-mode large scales a powerful probe for breaking the degeneracy with the amplitude of the scalar perturbations.

2.2.4 CMB challenges and B-mode quest

The measurement of the CMB radiation, and in particular its polarisation, is an enormous challenge due to the astrophysical foregrounds and instrumental systematic effects. In addition to the CMB, other astrophysical processes, located between us and the last-scattering surface, emit in the microwave frequency range. In particular, the galactic emission is the strongest contaminant, but there is also an extragalactic contribution. The standard approach to deal with these foregrounds is to clean the data using component separation algorithms (Planck Collaboration et al., 2020b). These methods exploit the frequency dependence of foregrounds to extract the CMB signal. For this to be possible, a broad frequency coverage is needed. The Figure 2.9 shows the Spectral Energy Distribution (SED) for the different astrophysical components in brightness temperature units, which also include the frequency bands of the *Planck* mission. In *Planck*, foreground cleaned CMB maps were mainly used for consistency checks and Gaussianity and statistical isotropy analy-

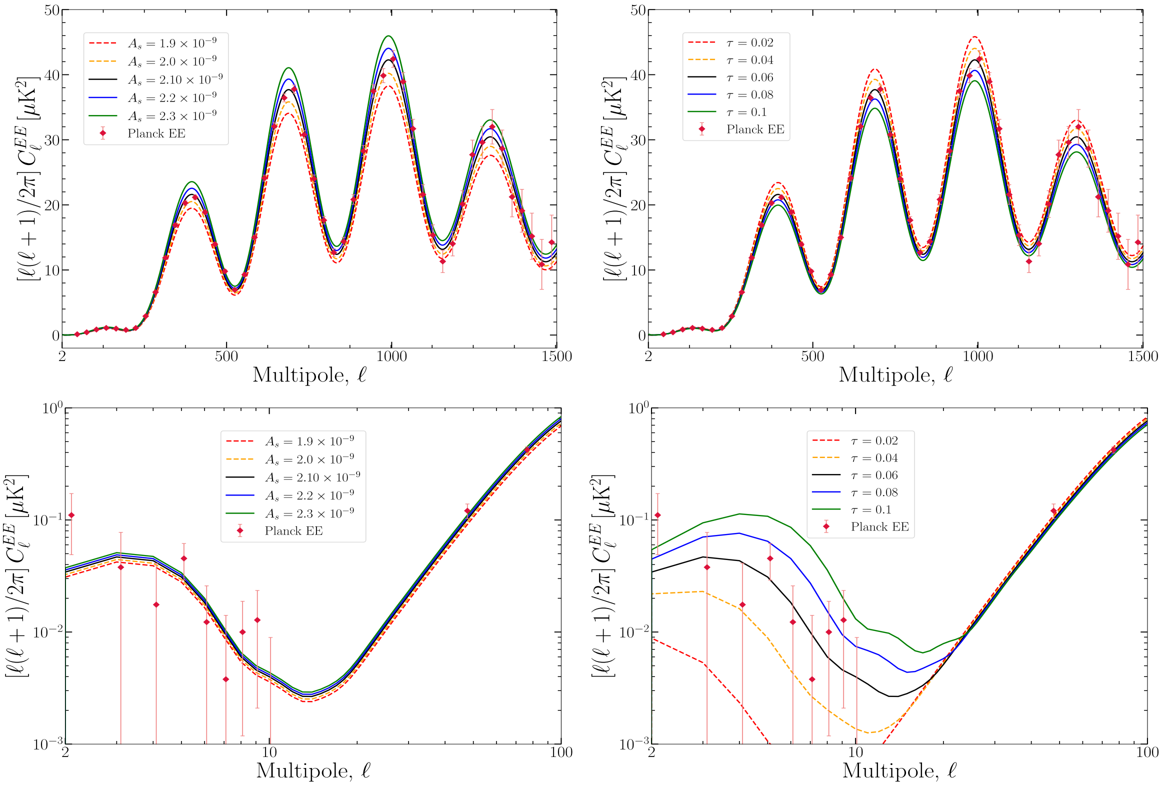

FIGURE 2.8: Breaking $A_s - \tau$ degeneracy with polarisation E-mode large scales. *Top panels:* Impact of the amplitude of primordial density fluctuations (*left panel*) and optical depth to reionisation (*right panel*) on the polarisation E-mode angular power spectrum. *Bottom panels:* Same but restricted to the large angular scales. The red dots represent the binned power spectrum as measured by the *Planck* mission, while the black curve corresponds to the best-fit $\Lambda\mathrm{CDM}$ model. Plots generated using `camb` (Lewis et al., 2000; Lewis & Challinor, 2011) software.

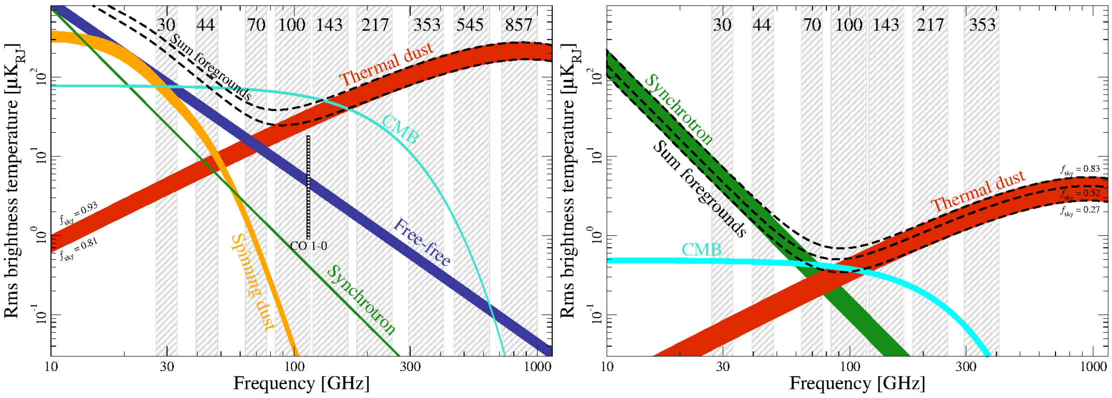

FIGURE 2.9: Spectral Energy Distribution (SED) of the main astrophysical components in temperature (*left panel*) and polarisation (*right panel*). The vertical grey bands indicate the frequency channels of the *Planck* mission. Figure taken from Planck Collaboration et al. (2020a).

ses (Planck Collaboration et al., 2020d). Some of these methods also derived foreground products, which were later used to characterise and model these components. In general, cosmological parameters were extracted from power spectrum based likelihoods, where frequency power spectra [11] were used (Planck Collaboration et al., 2020c; Tristram et al., 2024). In this case, the frequency maps were partially masked, and foregrounds and instrumental effects were modelled by nuisance parameters, which were marginalised over.

In intensity, the main foregrounds are Galactic diffuse components: synchrotron emission, free-free, thermal dust emission, and anomalous microwave emission (AME). Additionally, the cosmic infrared background (CIB), and extragalactic point sources also contribute. In polarisation only synchrotron and thermal dust seem to be polarised, although the polarisation of AME is still an open debate (Génova-Santos et al., 2017; González-González et al., 2025). Some point sources also exhibit some degree of polarisation. Thermal dust refers to the light emitted by interstellar dust grains when they are heated by UV and optical photons, while synchrotron emission corresponds to the radiation generated by relativistic cosmic ray electrons spiralling around the Galactic magnetic field lines.

Most of the component separation algorithms operate on multi-frequency maps, and they can be classified into parametric and non-parametric methods. Non-parametric methods do not assume any physical model of the sky components, while parametric methods assume the components SED's to fit for. In *Planck* analysis, four component separation methods were considered: one parametric, `Commander` (Eriksen et al., 2008), and three non-parametric, `SMICA` (Cardoso et al., 2008), `NILC` (Basak & Delabrouille, 2012), and `Sevem` (Fernández-Cobos et al., 2012).

In particular, primordial B-mode detection will be experimentally challenging as we

---

[11] Typically, the channels where CMB signal is less affected by foregrounds, i.e. 100-, 143-, and 217-GHz channels.

are searching for a very faint signal whose amplitude is still unknown. This signal is highly contaminated by polarised Galactic foregrounds, which dominate over all frequencies and sky regions. Therefore, cleaning the sky through the component separation algorithms will be crucial, especially if we want to achieve the sensitivity in r expected for the next generation experiments such as LiteBIRD, $r_{95\%} \lesssim 0.001$. This requires, first, a better understanding and modelling of the foregrounds, and second, optimal component separation methodologies. A non-proper treatment of foregrounds could bias our results and lead to a false claim of primordial gravitational waves detection, as it happened with BICEP2 collaboration (BICEP2 Collaboration et al., 2014; BICEP2/Keck Collaboration et al., 2015). In particular, RadioForegroundsPlus project [12] aims to combine data sets of different experiments to provide the best possible characterisation of the physical properties of polarized emission in the microwave frequency range.

Planned and future experiments will achieve such high-sensitivity that it will require unprecedented systematics control through an accurate modelling of the instrument. Some of the systematics are:

- **Polarisation angle:** Uncertainties in the polarisation angle of detectors bias the tensor-to-scalar ratio r by mixing the E- and B-mode signals. In particular, LiteBIRD will need to calibrate the polarisation angles to the arcminute level in the most sensitive frequency band (Vielva et al., 2022) to achieve the proposed sensitivity on r.
- **Beam characterisation:** An imperfect modelling of the beam is a potential source of systematic effects. In particular, the uncertainty in the beam far side-lobe is expected to be one of the main systematics in future CMB observations (Leloup et al., 2024).
- **Half-wave plate (HWP):** Measuring polarisation signal requires new sophisticated technologies that include the use of fast-rotating HWP. This approach modulates the linear polarisation to 4 times the frequency of the HWP, allowing a quasi-instantaneous measurement of Stokes parameters by each detector (Patanchon et al., 2024). This helps reducing the intensity to polarisation leakage and the $1/f$ noise. However, HWP can introduce new systematics. Imperfections in the HWP can bias the tensor-to-scalar ratio (Patanchon et al., 2024; Monelli et al., 2024).

As already stated in the previous section, primordial scalar perturbations can not produce B-mode polarisation signal. This signal is only generated from tensor perturbations, which also produce imprints in the temperature and E-mode power spectra as shown in Figure 2.10. Primordial gravitational waves induce time-varying distortions in spacetime that affect the propagation of photons. As they propagate, the two polarisation states, h_+ and $h_\times$, stretch and compress the space along orthogonal directions, with a relative orientation of $45°$, as Figure 2.11 illustrates. This distortion induces a quadrupole anisotropy in the radiation field, leading to linear polarisation.

---

[12] `https://research.iac.es/proyecto/radioforegroundsplus/`

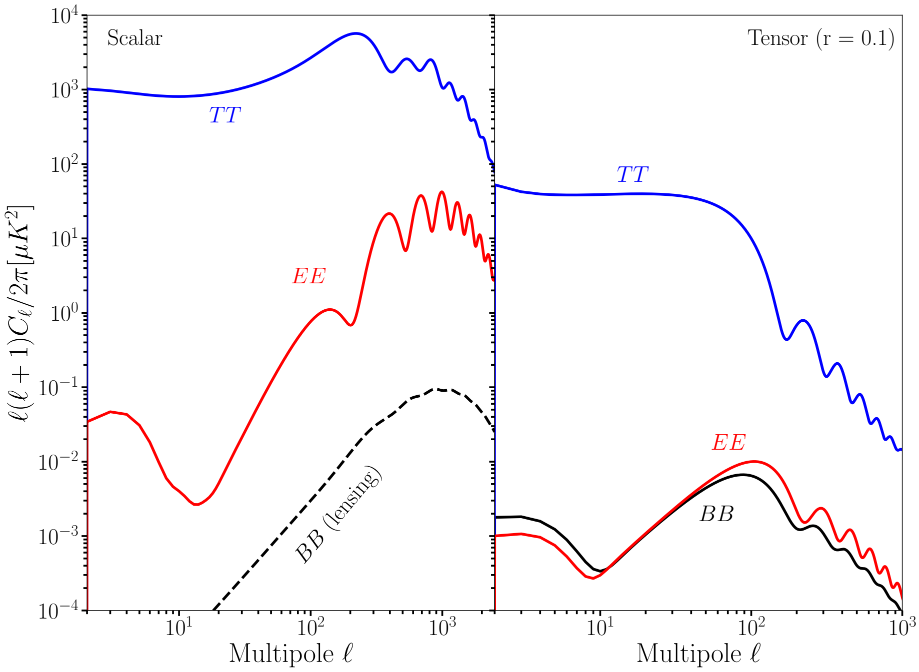

FIGURE 2.10: Temperature and polarisation angular power spectrum induced by scalar (*left panel*) and tensor (*right panel*) perturbations. A tensor-to-scalar ratio of 0.1 has been considered. In the case of scalar perturbations, BB spectrum corresponds to lensing induced B-modes. Plot generated using `CLASS` (Blas et al., 2011) software.

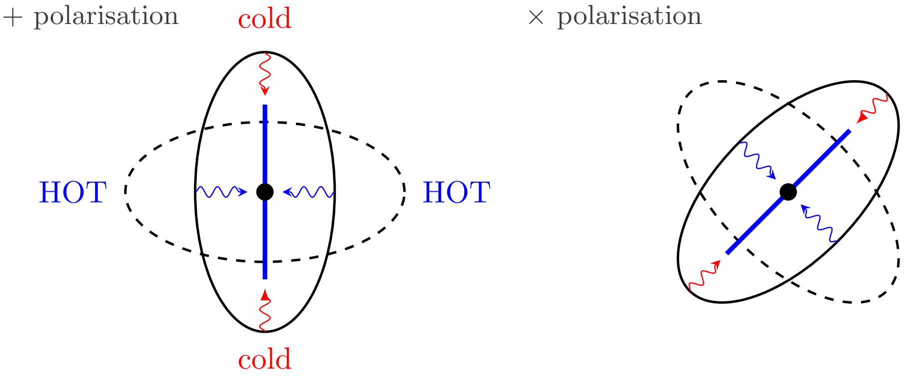

FIGURE 2.11: Illustration of CMB polarisation induced by gravitational waves. As the wave propagates, it stretches and compresses the space along orthogonal directions, generating a temperature quadrupole in the photon distribution. The two polarisation states of the gravitational wave, $+$ and $\times$, produce linear polarisation patters rotated by $45°$ with respect to each other. Figure adapted from Baumann (2022).

The B-mode power spectrum is characterised by two peaks, the reionisation and recombination bumps. The first peak is located at $\ell < 10$, and is caused by the re-scattering of the CMB photons at reionisation. It depends on both the tensor-to-scalar ratio and the optical depth to reionisation. The second peak depends on r and is located at $\ell \approx 100$. Above this multipole the signal decays. While the reionisation bump is only reachable from space, the recombination peak is the main target of ground-based experiments. However, this peak is hidden by the lensing B-mode. Weak lensing leaks power from E- to B-mode (Lewis & Challinor, 2006) in such a way that it produces a lensed B-mode power spectrum which effectively acts like $\sim 5\,\mu\mathrm{K}\cdot\mathrm{arcmin}$ white noise up to $\ell \approx 1000$ (see Fig. 2.10). Delensing techniques are the standard way to mitigate this effect. Lensed B-mode has been already detected by several experiments as shown in Figure 2.2.

3

CMB anomalies

> *"Every question leads to new answers, new discoveries, and new smarter questions."*
>
> Bill Nye

The cosmological principle, which states that the Universe is statistically isotropic and homogeneous on large scales, is a fundamental pillar of the ΛCDM model. This model relies on the assumption that the geometry of the Universe on large scales is well described by the FLRW metric, and it is strongly supported by the simplest inflationary models and CMB observations. In particular, data from the WMAP and *Planck* satellites have provided a very accurate confirmation of the base ΛCDM model. However, these observations have also revealed deviations from statistical isotropy in the temperature maps. These features, often referred to as CMB anomalies, have been found at modest levels of statistical significance (i.e., 2–$3\,\sigma$), and have the potential to challenge fundamental assumptions. Their importance lies in their large-scale nature, which is a common property for most of them. Large angular scales offer a window into the initial conditions of the Universe, and therefore, CMB anomalies have the potential to probe new physics in the very early Universe, particularly during the epoch of inflation. This chapter aims to provide a brief review on the CMB anomalies, which constitute the main topic of this thesis.

The main purpose of this chapter is to summarise the current state of the large-scale CMB anomalies, the features suggesting a potential deviation from the statistical isotropy and Gaussianity assumptions. It is important to note that these two properties are what allow all cosmological information to be encoded in the CMB power spectra (see Eq. 2.5) (Planck Collaboration et al., 2016c). The first anomaly dates back to the COBE DMR data, where they found a surprisingly low value for the quadrupole moment and a lack of correlation for angular scales above 60 degrees (Hinshaw et al., 1996a,b). The most important anomalies are: a low-ℓ power deficit and a low variance ; a lack of correlation on large angular scales in the angular 2-point correlation function; alignment between the quadrupole and octopole moments; a hemispherical power asymmetry; a parity asymmetry; and an anomalous "Cold Spot" on angular scales of approximately 10°. Most of them emerged during the analysis of WMAP first-year data (Bennett et al., 2003; Eriksen et al., 2004b; Schwarz et al., 2004; Vielva et al., 2004; Hansen et al., 2009), and were later confirmed by the *Planck* satellite (Ade et al., 2014, 2016; Planck Collaboration et al., 2020d).

There are three plausible explanations for the origin of these mild tensions. The most exciting one is a potential cosmological origin, which implies new physics beyond the standard model. Some attempts have been made to explain the anomalies. For instance, a cutoff in the primordial power spectrum, $\mathcal{P}(k)$, could account for the lack of correlation and the low variance, while a modulation of the spectrum could produce the observed asymmetry (Contaldi et al., 2003; Planck Collaboration et al., 2020e). Although it is relatively straightforward to construct models in which a fast-roll phase preceding slow-roll inflation naturally explains the lack of power at large angular scales, it is considerably more difficult to develop inflationary models that predict the phenomenological modulations (Planck Collaboration et al., 2016c, 2020e). Standard slow-roll models in the single-field inflation context simply cannot produce such asymmetry (Erickcek et al., 2008).

More exotic models have therefore been explored. Examples include multi-field inflation. The inclusion of an isocurvature superhorizon perturbation in the curvaton field (Erickcek et al., 2008, 2009) can produce the observed scale-dependent asymmetry without violating the small scales constraints coming from high redshift quasars (Hirata, 2009). Other phenomenological modulations are those studied in Zibin & Contreras (2017), Contreras et al. (2017), Contreras et al. (2018), and Dai et al. (2013). However, these mechanisms have their own weaknesses and none of them is favoured by *Planck* measurements (Planck Collaboration et al., 2020e).

A second explanation is that these anomalies could result from foreground or systematic effects. However, this is highly unlikely (Schwarz et al., 2016), as they have been observed by two independent space missions with differing scanning strategies, systematic uncertainties, and frequency coverage. Moreover, the frequency independence of the observed anomalies and the fact that the anomalies show up at similar statistical significance in the four foreground cleaned CMB maps produced by *Planck* are strong arguments against a foreground related origin. Recent work (Jung et al., 2024) has attempted to quantify the impact of the Sunyaev-Zeldovich (SZ) signal from the local Universe on large-scale CMB anomalies, concluding that the local tSZ and kSZ effects can not account for the observed deviations from isotropy. In parallel, recently a hint of a potential new foreground detection (Luparello et al., 2023; Hansen et al., 2023; Cruz et al., 2025) has been proposed as a possible explanation for large-scale CMB anomalies. By stacking the CMB observations, an indication of a temperature decrement has been identified around local galaxies.

A third possibility is that the anomalies are either statistical flukes, or that the derived significance levels are not properly computed in absent of a well-motivated theoretical model. The main discussion here is that the significance of large-scale anomalies may be subject to a posteriori corrections (Bennett et al., 2011; Bennett et al., 2013), the well-known “look-elsewhere" effect. Another interesting debate is their statistical independence. A recent work suggests that their joint probability is above 5σ (Jones et al., 2023), and claims that the Universe is not isotropic. Given that intensity measurements have reached the cosmic variance limit, the largely independent information provided by large-scale polarisation is needed to clarify the origin of these anomalies. However, it is well established that the large angular scales in the *Planck* E-mode measurements are limited by systematic effects. Therefore, future polarisation observations, such as those from LiteBIRD, will provide valuable insights into this topic.

3.1 Current state of art

In this section, we review the main CMB anomalies, excluding the hemispherical power asymmetry (HPA), which will be discussed in more detail in Section 3.2, given its particular relevance to this work. The statistical significance of these anomalies lies around $2-3$ σ, depending on the estimator used. All of them are well documented (see, for example, (Schwarz et al., 2016; K. Aluri et al., 2023) and references therein).

Low variance

Previous analyses of WMAP (Monteserín et al., 2008; Cruz et al., 2011; Gruppuso et al., 2013) and *Planck* (Ade et al., 2014, 2016; Planck Collaboration et al., 2020d) data have reported an anomalous low variance in the temperature maps compared to the simulations based on the ΛCDM model.

In Planck Collaboration et al. (2020d) the unit variance estimator, described in Cruz et al. (2011), is applied to the *Planck* 2018 component-separated maps to measure the variance, skewness, and kurtosis. The method works in the following way[1],

---

[1] Skewness and kurtosis require appropriate normalization.

- The normalized CMB temperature is given by,

  $$u_i^T = \frac{T_i}{\sqrt{\sigma_0^2 + \sigma_{i,N}^2}}, \tag{3.1}$$

  where T_i is the observed temperature at pixel i, σ_0^2 is the variance of the CMB signal, which is the same for all the pixels, and $\sigma_{i,N}^2$ is the variance of the noise for that pixel, which is computed from the Monte Carlo (MC) simulations.

- The variance of the CMB, σ_0^2, is then estimated by finding the value that minimizes the deviation of the normalized map's variance from unity,

  $$\hat{\sigma_0^2} = s_0^2 : \min ||\mathrm{var}(\mathrm{u}(\mathrm{s}_0^2)) - 1||. \tag{3.2}$$

  This estimator was found to be more optimal than the estimator that minimizes the standard Kolmogorov-Smirnov distance (Cruz et al., 2011).

Once the variance is estimated, the lower-tail probability is computed by comparing the value obtained from the data with the distribution obtained from simulations. The p-value[2] was computed for different resolutions, finding consistent results among the four component-separated CMB maps. The results show a decreasing lower-tail probability with decreasing resolution, highlighting the presence of lack of power on large angular scales. Furthermore, the p-value was slightly larger, around 1% against the 0.5% reported in previous analysis. They attributed this difference to the fact that the PR3 common mask keeps a larger sky fraction compared to the PR2 common mask. In Gruppuso et al. (2013) and Ade et al. (2014), it was already reported that the lower-tail probability depends on the mask, with the variance appearing more anomalously low when a mask covering larger fraction of the sky was applied. The skewness and kurtosis were compatible with a Gaussian field. Similar results were obtained for the `Sevem` frequency-cleaned CMB maps (70, 100, 143, and 217 GHz), which suggest that foreground contamination is not the reason for the observed low variance.

In Cruz et al. (2011), an hemispherical dependence of the low variance was also found, which was later confirmed in Ade et al. (2014). By analysing the variance separately in each hemisphere, it was found that the northern ecliptic hemisphere exhibited an anomalously low variance, while the southern hemisphere remains consistent with the MC simulations. Similar results were obtained using the Galactic hemispheres. They also investigated the potential link with the quadrupole-octopole alignment, concluding that the data were fully

---

[2] In general, the p-value is defined as the fraction of simulations with a value of a given estimator equal to or more extreme (either higher or lower depending on the context) than that observed in the data. In this section, the terms p-value and lower-tail probability are used interchangeably. Note, however, that the p-value is a broader concept, and in some cases it corresponds to the upper-tail probability, even if this is not explicitly stated.

compatible with simulations after quadrupole and octopole moments were subtracted. This indicates a potential common origin for both features.

In Planck Collaboration et al. (2020d) the low variance in polarisation data was also tested. They follow a different approach due to the low signal-to-noise ratio, and they checked the variance and the cross-variance between *odd-even* (OE) and *half-mission* (HM) splits[3] for the component-separated polarisation CMB maps, for the `Sevem` frequency-cleaned maps, and also for the cross-variance between pairs of frequencies. In general, they found a reasonable consistency with MC simulations, but they also claimed that the low signal-to-noise ratio of *Planck* polarisation data and the uncertainties in the noise characterization limited the detection of potential anomalies.

A more recent work (Billi et al., 2024) has studied this anomaly using the latest PR4 data release, and a new class of optimised estimators in harmonic space that are able to test the anomaly considering only temperature, only polarisation, or jointly. They confirmed the presence of lack of power applying the new estimator to the PR3 (PR4) data set, with a p-value $\leq 0.33\%(1.76\%)$. The results using only polarisation E-mode data are fully compatible with ΛCDM, although they found significant differences between both data sets due to different levels of systematics. The joint estimator provides similar results to the ones obtained with only temperature data.

Lack of correlation at large scales

The non-Gaussianity of the temperature and polarisation data was also studied in Planck Collaboration et al. (2020d) using the N-point correlation functions, which is defined as the average product of the temperature or polarisation field at N directions,

$$C_N(\theta_1, ..., \theta_{2N-3}) = \langle X(\hat{\mathbf{n}}_1)...X(\hat{\mathbf{n}}_N)\rangle \, . \tag{3.3}$$

Under the statistical isotropy assumption, this function depends only on the relative angles between directions (θ_1,...,θ_{2N-3}). In particular, the 2-point function and some configurations[4] for the 3-point function were considered in Planck Collaboration et al. (2020d). Deviations from zero in the 3-point function are used to test non-Gaussianity. Actually, MC simulations already showed some level of deviation which indicates that the modelled systematics can induce some non-Gaussianity. Using a χ^2 statistic to quantify the agreement between data and simulations, they found no statistical evidence of non-Gaussianity. However, the results for the TT correlation function (see Fig. 8 in Planck Collaboration et al. (2020d)) confirmed the existence of a lack of correlation for large angular separations, for

---

[3] In the *Planck* data analysis, *odd-even splits* refer to dividing the data according to alternating stable pointing periods (or rings), while *half-mission splits* separate the observations into two halves corresponding to the first and second parts of the mission. These splits are commonly used to test for internal consistency and to identify potential residual systematics.

[4] The 2-point function only depends on a single parameter, the angular distance between the two pixels, while the 3-point correlation function depends also in the relative position between three pixels.

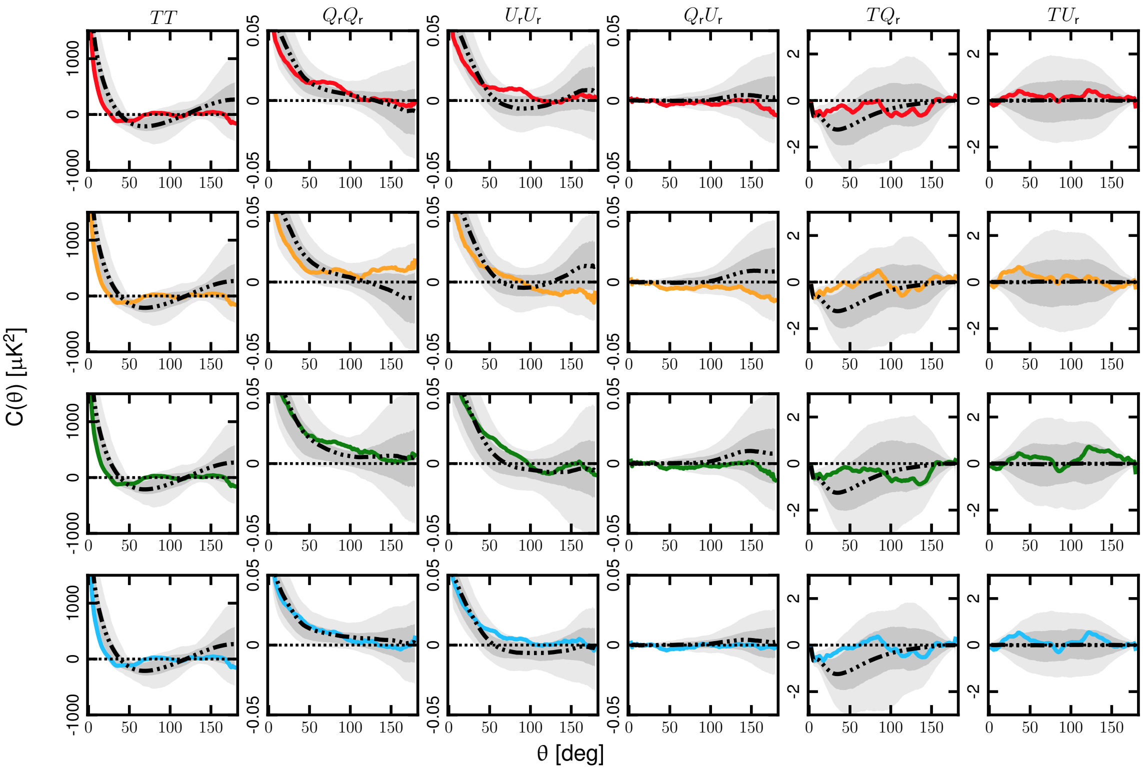

FIGURE 3.1: Two-point correlation functions from *Planck* 2018 temperature and polarisation data. Solid coloured lines show the results for the four component-separation methods, `Commander`, `NILC`, `Sevem`, and `SMICA`, ordered from top to bottom. The black dot-dashed line corresponds to the mean correlation function from the FFP10 simulations, while shaded grey regions indicate the 68% and 95% confidence intervals, also obtained from simulations. Figure taken from Planck Collaboration et al. (2020d).

angles above $\sim 60°$, as previously noted in Hinshaw et al. (1996b), Bennett et al. (2003), and Copi et al. (2015b). This is shown in the first column of Figure 3.1.

To quantify this lack of correlation on large angular scales typically the $S_{1/2}$ statistic is used, first proposed in Spergel et al. (2003) and extended to consider TE correlation in Copi et al. (2013). The most general statistic is defined by

$$S^{XY}(\theta_1, \theta_2) = \int_{\cos\theta_2}^{\cos\theta_1} [\hat{C}_2^{XY}(\theta)]^2 \mathrm{d}(\cos\theta), \tag{3.4}$$

where $X, Y \equiv \{T, Q, U\}$, and $\hat{C}_2^{XY}(\theta)$ denotes the estimated 2-point correlation function, which can be calculated either in real space (Planck Collaboration et al., 2020d) or in harmonic domain using the (pseudo-)power spectrum, as in Eq. 2.5 (Copi et al., 2013; Gruppuso, 2014; Copi et al., 2015b). For temperature, the $S_{1/2}$ statistic was defined by a posterior choice of $\theta_1 = 60°$ and $\theta_2 = 180°$. In particular, depending on the analysis choices, the reported statistical significance of this statistic has been between 0.1% and 0.5% in both WMAP and *Planck* data (Copi et al., 2007, 2009; Gruppuso, 2014; Copi et al., 2015b; Ade et al., 2016; Planck Collaboration et al., 2020d). In some cases even below 0.01% (Gruppuso, 2014) using nine-year WMAP data. As reported in Gruppuso (2014), similar to the case of the low variance anomaly, a more conservative masking procedure was found to increase the significance of the lack of power.

The four component separation methods used in Planck Collaboration et al. (2020d) showed p-values of $< 0.1\%$ for temperature analysis. In order to avoid the "look-elsewhere" effect, the *Planck* team defined a global p-value, which was found to be of the order of $\sim 1\%$. This p-value increased to $\sim 13\% - 14\%$ when the quadrupole was subtracted from the data prior to computing the statistic, highlighting the connection between the low quadrupole amplitude and the lack of power. Taking into account that the shape of the 2-point correlation function for $\theta > 60°$ is dominated by the large angular scales or the lowest multipoles in the harmonic domain, it seems natural to link both. This was already pointed out in Gruppuso (2014). The p-values obtained in Planck Collaboration et al. (2020d) using polarisation data were not statistically significant, and the inconsistency in p-values shown between different component-separated maps rather indicated the presence of different levels of residuals. Additionally, the inferred S^{TQ} value, which was shown in Copi et al. (2013) to be an appropriate quantity to test for the fluke hypothesis[5], did not allow one to rule it out.

In Ade et al. (2016), the *Planck* team also used an alternative χ^2 statistic to test the anomaly, obtaining a slightly smaller p-values (see Table 13 in Ade et al. (2016)) of around 2%. This approach was adopted to address the criticism that the $S_{1/2}$ statistic does not account for correlations between bins in the 2-point correlation function.

Quadrupole and octopole alignment

The quadrupole and octopole (Q-O) alignment was first explored in de Oliveira-Costa et al. (2004). They found that the cosmic octopole is anomalously planar, which means that most of the octopole's power is concentrated in the $a_{\ell m}$ coefficients with $|m| = 3$, or, in other words, that the hot and cold spots of the octopole lie predominantly near a single plane, in this case, close to the Galactic plane. Furthermore, they found that the preferred axes, $\hat{\mathbf{n}}_2$ and $\hat{\mathbf{n}}_3$, for the quadrupole and octopole are anomalously aligned. These axes are the axis $\hat{\mathbf{n}}$ that maximised the angular momentum dispersion,

$$\sum_m m^2 |a_{\ell m}(\hat{\mathbf{n}})|^2, \tag{3.5}$$

where $a_{\ell m}(\hat{\mathbf{n}})$ denotes the spherical harmonic coefficients computed in a rotated frame where the z-axis is aligned with the $\hat{\mathbf{n}}$ direction. Using the dot product between both axes, $|\hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_3|$, as a statistic to quantify the alignment, they found a strong alignment with a p-

---

[5] Under the fluke hypothesis, the observed lack of correlation in the temperature data is not a hint of new physics, but rather a rare statistical fluctuation within the ΛCDM model.

value[6] of 1.6%. This is highly unexpected under the statistical isotropy assumption, as the multipole moments are uncorrelated. Consequently, $\hat{\mathbf{n}}_2$ and $\hat{\mathbf{n}}_3$ are expected to be random realisations drawn from a uniform distribution in which all directions are equally probable.

Other works have studied this anomaly using alternative statistics for the alignment (Schwarz et al., 2004; Copi et al., 2007, 2015a). A useful tool is the decomposition of the CMB sky in terms of multipole vectors as described in (Copi et al., 2004),

$$\sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m} \approx A^{(\ell)} \prod_{i=1}^{\ell} (\hat{\mathbf{v}}^{(\ell,i)} \cdot \hat{e}), \tag{3.6}$$

where $\hat{\mathbf{v}}^{(\ell,i)}$ is the i^{th} multipole vector of the ℓ^{th} multipole. In this picture, the quadrupole and octopole are described in terms of two and three multipole vectors, respectively. The usual way of comparing the multipole vectors to test the Q-O alignment is as follows:

- For each ℓ compute the $\ell(\ell-1)/2$ cross products, $\mathbf{w}^{(\ell,i,j)} \equiv \pm(\hat{\mathbf{v}}^{(\ell,i)} \times \hat{\mathbf{v}}^{(\ell,j)})$. These are also called area vectors. In particular, the quadrupole and octopole have 1 and 3 area vectors, respectively.

- The dot products between the quadrupole area vector and the three octopole area vectors are computed.

$$\begin{aligned} A_1 &\equiv |\mathbf{w}^{(2,1,2)} \cdot \mathbf{w}^{(3,1,2)}|, \\ A_2 &\equiv |\mathbf{w}^{(2,1,2)} \cdot \mathbf{w}^{(3,2,3)}|, \\ A_3 &\equiv |\mathbf{w}^{(2,1,2)} \cdot \mathbf{w}^{(3,3,1)}|. \end{aligned} \tag{3.7}$$

- An S statistic is defined as the sum of the dot products, $\mathrm{S} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{A}_{\mathrm{i}}$.

- The S statistics can be modified to quantify the alignment of the area vectors with a fixed direction in the sky, $\hat{\mathbf{e}}$,

$$\mathrm{S} \equiv \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} |\mathbf{w}_{\mathrm{i}} \cdot \hat{\mathbf{e}}|, \tag{3.8}$$

where the sum is over a set of area vectors.

Using the S statistic, a p-value of 0.13 for the Q-O alignment was found in the first-year WMAP maps. Surprisingly, both multipoles were also found to lie within the ecliptic plane and close to the dipole axis, with p-values below 1% and 0.4%, respectively.

The unusual Q-O alignment persisted in the other WMAP releases and also in *Planck* data. In Copi et al. (2007), a p-value of 0.4% was found for the alignment after removing the kinetic quadrupole (KQ) component, which corresponds to the Doppler contribution to

---

[6] In this context, the p-value is referred to the upper tail probability, i.e., only 1.6% of simulations exhibited a dot product above the one observed in the data.

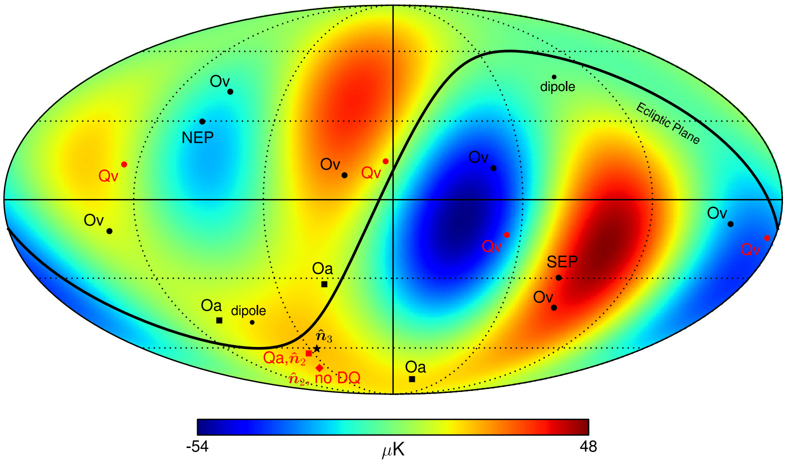

FIGURE 3.2: Combined quadrupole and octopole pattern from the *Planck* 2013 `SMICA` map, after correcting for the Doppler induced quadrupole. Multipole vectors for the quadrupole (octopole) are indicated by red (black) circles, labelled as Qv (Ov). The directions of the area vectors, $\mathbf{w}^{(\ell,i,j)}$, are shown by red (black) squares for the quadrupole (octopole). In the quadrupole case, this coincides with the direction that maximises the angular momentum dispersion. The black star indicates the direction that maximises the angular momentum dispersion of the octopole. Additionally, the red diamond displays the direction that maximises the angular momentum dispersion of the quadrupole before applying the KQ correction (DQ in the figure). Figure taken from Copi et al. (2015a).

the quadrupole. In particular, the alignment was found to becomes even more anomalous after KQ correction, something that was confirmed later by *Planck* (Ade et al., 2014). They also confirmed the alignment with the dipole axis, with a p-value of around 0.3% and an unexpected alignment with the Galactic poles at the level of $\sim 1\%$, while the significance of the correlation with the ecliptic plane decreased to 4%. In Ade et al. (2014), the *Planck* team used the angular momentum dispersion statistic to test the Q-O alignment in the first data release, obtaining a p-value of $\sim 1\%$ for all the component-separated maps after KQ correction (see Table 18 in Ade et al. (2014)). Similarly, in Copi et al. (2015a), they applied both statistics, the angular momentum dispersion and the S statistic, finding p-values consistently below 0.6% for the WMAP's ILC[7] 7yr, 9yr, and *Planck*'s first release maps. In particular, the p-values for the alignments with the ecliptic plane, the Galactic poles, and the dipole axis were found to be around 3%, 1%, and 0.2%, respectively. The p-value for these alignments increased above 6% when considering ΛCDM realisations with the Q-O alignment at least as close as in the data. Figure 3.2 shows the combined quadrupole-octopole `SMICA` map for *Planck* first release data, including the mutipole and area vectors, and the maximum angular momentum dispersion directions.

Computing the multipole vectors from a cut-sky introduces significant uncertainties that affect the statistics. Therefore, the alignment analyses are usually performed in the

---

[7] Internal Linear Combination (ILC) is a foreground-cleaning method that combines linearly the CMB frequency maps with a set of weights that are chosen to minimize the foreground contamination while preserving the CMB signal.

component-separated full-sky maps (Copi et al., 2007), which makes our ability to remove foregrounds the main limitation to measure the alignment. Foreground contamination as the origin of the alignment does not seem to be a satisfactory explanation as the anomaly persists in the data even after applying different component-separation methods. More importantly, in Bennett et al. (2011), it was shown that the alignment is not dominated by one or two discrete regions.

A recent work (Patel et al., 2025) tested the low-ℓ alignment using the five-year WMAP data and the *Planck* PR1, PR2, and PR3 data releases. As an alignment statistic, they used the dot product between the principal eigenvectors (PEV) of the power tensor. Essentially, the PEV is the direction that maximises the angular momentum dispersion. The analysis confirmed the presence of the alignment across all data sets. Furthermore, Jung et al. (2024) reported a p-value below 1% using the latest PR4 `Sevem` and `Commander` maps.

Parity asymmetry

In this section, we focus on the point-parity or even-odd asymmetry. Motivated by the potential impact of the galactic foregrounds on the observed large-scales anomalies, an estimator for the point-parity symmetry, based on the even and odd multipoles of the power spectrum, was introduced in Land & Magueijo (2005). It was applied to the WMAP data to search for foreground residuals, which were found to favour even multipoles. However, contrary to expectations, the analysis revealed a preference for odd parity, although not at a statistically significance level.

The CMB temperature field can be decomposed in terms of its positive and negative parity components,

$$T(\hat{\mathbf{n}}) = T^{+}(\hat{\mathbf{n}}) + T^{-}(\hat{\mathbf{n}}), \tag{3.9}$$

where

$$T^{\pm}(\hat{\mathbf{n}}) = \frac{T(\hat{\mathbf{n}}) \pm T(-\hat{\mathbf{n}})}{2}. \tag{3.10}$$

The positive and negative parity components are constructed only from even or odd multipoles, respectively. Similarly for polarisation field. On large angular scales the Sachs-Wolfe plateau ensures that both parity modes have a comparable amplitude. Therefore, testing for a relative amplitude difference between them enables the detection of a possible even or odd point-parity preference.

An odd-parity preference was detected in WMAP data (Kim & Naselsky, 2010b,a; Gruppuso et al., 2011), with p-values below 0.5%, depending on the analysis choices, such as whether the power spectrum was estimated from full-sky or cut-sky maps, or on the adopted maximum multipole ($\ell_{\max}$).

The point-parity anomaly is typically quantified using the following estimator,

$$R^{TT}(\ell_{\max}) = \frac{C_+^{TT}(\ell_{\max})}{C_-^{TT}(\ell_{\min})}, \tag{3.11}$$

where $C_\pm^{TT}$ are given by

$$C_\pm^{TT}(\ell_{\max}) = \frac{1}{\ell_{\rm tot}^\pm} \sum_{\ell=2,\ell_{\max}}^{+,-} \frac{\ell(\ell+1)}{2\pi} C_\ell^{TT}, \tag{3.12}$$

with $\ell_{\rm tot}^\pm$ is the total amount of even and odd multipoles considered in the sum. The power spectrum can be computed straightforwardly in the full-sky case. For the cut-sky case, it can be done by using an optimal Quadratic Maximum Likelihood (QML) estimator, as proposed in Gruppuso et al. (2011), or a pseudo-C_ℓ estimator (Hivon et al., 2002), as applied in Kim & Naselsky (2010a). In particular, Kim & Naselsky (2010a) reported a p-value of 0.3% for the WMAP seven-year map. This value decreased up to 0.2% using the whole sky. These results were later confirmed by *Planck* analyses (Ade et al., 2014, 2016; Planck Collaboration et al., 2020d), which also showed that the significance depends on the maximum multipole considered (see Figure 25 in Planck Collaboration et al. (2020d)). In particular, the significance peaks for $\ell_{\max} \approx 20-30$. In Planck Collaboration et al. (2020d), the p-value was found to be $\sim 1\%$ in that range using the *Planck* 2018 common mask. This result was consistent among the four component-separated CMB maps. An additionally considered case, which used the power spectrum obtained from the `Commander` code, showed a slightly smaller p-value with a minimum probability of 0.2% for $\ell_{\max} = 24$. Correcting for the "look elsewhere" effect, this value increased to 1.6%.

For polarisation analysis, the following alternative estimator was considered to avoid problems in the denominator

$$D^X(\ell_{\max}) = C_+^X(\ell_{\max}) - C_-^X(\ell_{\max}), \tag{3.13}$$

where X corresponds to TE and EE. No anomalous lower-tail probability was found, being the low signal-to-noise of *Planck* polarisation data the main limitation.

The Cold Spot and other large-scale peaks

The Cold Spot (CS) is an unusual cold region of around 10° in the southern hemisphere, at Galactic coordinates $(\ell, b) = (209^\circ, -57^\circ)$, surrounded by a hot ring. The CS has been extensively analysed using a variety of statistical estimators designed to trace non-Gaussian features (see Vielva (2010) for an overview). It was first detected in Vielva et al. (2004) by performing a spherical wavelet analysis on the first-year WMAP data. The methodology was the following one:

- They filtered the map with the spherical Mexican hat wavelet (SMHW) using different scales. As pointed out in Vielva et al. (2004), wavelets were found to be more efficient for detecting certain non-Gaussian features, and most importantly, they made possible to identify the angular scales in which the non-Gaussian signature was more prominent. The SMHW is described in detail in Martínez-González et al. (2002), where it was demonstrated to be more efficient than other wavelets for detecting non-Gaussian features.

- For each scale, three statistics were defined: the variance $\sigma^2(R)$, the skewness $S(R)$, and the kurtosis (R),

  $$\begin{aligned} \sigma^2(R) &= \frac{1}{N_R}\sum_{i=1}^{N_R} w_i(R)^2, \\ S(R) &= \frac{1}{N_R}\sum_{i=1}^{N_R} w_i(R)^3/\sigma(R)^3, \\ K(R) &= \frac{1}{N_R}\sum_{i=1}^{N_R} w_i(R)^4/\sigma(R)^4 - 3, \end{aligned} \tag{3.14}$$

  where N_R is the number of wavelet coefficients $w_i(R)$ that depends on the mask.

- Using 10^4 simulations, the acceptance intervals for each statistic were computed. An excess of kurtosis was detected for $R = 4^\circ.17$ and $R = 5^\circ$ scales, with a right-tail probability of $\approx 0.4\%$, which means that only 40 simulations showed a kurtosis value as large as the one observed in the WMAP data. In particular, a dedicated SMHW analysis in each hemisphere showed that the excess was located in the southern hemisphere, with a p-value of $\approx 0.1\% - 0.2\%$. No foregrounds, systematics, or uncertainties in the fiducial model were found to be responsible for that non-Gaussian detection.

- Analysing the SMHW coefficients, a very cold spot was identified as the possible source of non-Gaussianity, with a value of -4.57σ for $R = 4^\circ.17$ scale at $(\ell, b) = (209^\circ, -57^\circ)$. Furthermore, it did not show a frequency dependence.

These results were also confirmed in Cruz et al. (2005). They applied six estimators on seven regions, confirming the presence of a big cold spot in the south-west region. In particular, they found that the probability of observing a spot similar to the CS is $\approx 0.2\%$. A similar p-value was obtained applying an area estimator to the SMHW coefficients. After masking the CS region, this probability increased by a factor of about 100, and the kurtosis decreased, making the data much more compatible with the Gaussian hypothesis.

An alternative analysis based on the Higher Criticism (HC) statistic was performed in Cayón et al. (2005). This statistic was found to be optimal for detecting non-Gaussian features, and also valuable for identifying their origin by determining the portion of the data

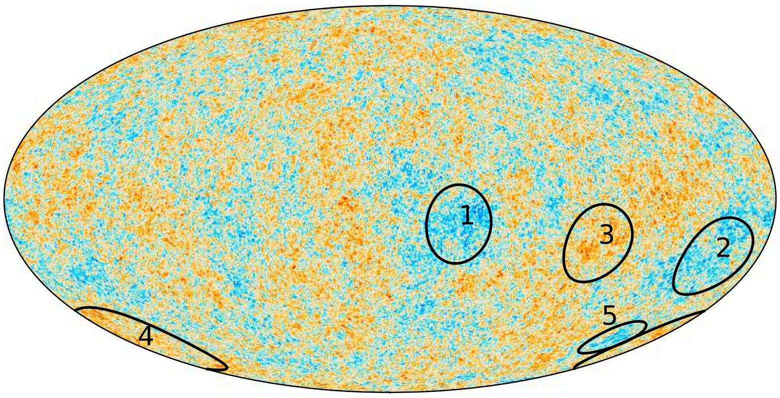

FIGURE 3.3: Inpainted *Planck* 2018 `Commander` map. The black circles indicate the locations of the large-scale peaks, labelled form 1 to 5 (Cold Spot). Figure taken from Planck Collaboration et al. (2020d).

that contributes most to the deviation. Using 5000 simulations, a p-value of 0.54% was found in the WMAP data set. Furthermore, a ring centred in $(\ell, b) \approx (209^\circ, -57^\circ)$ containing 490 pixels was localised as the origin of such deviation. These pixels are part of the CS pointed out in previous works.

A morphological study of the CS revealed that it has an almost circular shape (Cruz et al., 2006). To account for the "look-elsewhere" effect, a new test was performed in Cruz et al. (2007b) to eliminate a posteriori choices. Using this approach, the p-value obtained from the skewness and kurtosis estimators was 1.85%.

The CS was later confirmed by the *Planck* mission (Ade et al., 2014, 2016; Planck Collaboration et al., 2020d). In Ade et al. (2014) an excess of kurtosis was detected on wavelet scales of around 5° with p-values of $\approx 1\%$. Furthermore, applying the area estimator, an anomalous cold area was detected with a minimum p-value of 0.1%. This result was robust across different component-separated maps and masks. Most recently, the most prominent peaks on the sky have been studied using the *Planck* data. In Marcos-Caballero et al. (2017a) a multiscale analysis of the CMB temperature field was done through its derivatives up to second order. In particular, by looking to extreme deviations in the derivatives field, the CS was detected as an extreme value on the curvature parameter, which is given by the Laplacian of the temperature field. As mentioned in the paper, studying the curvature at different scales (by filtering the temperature field with Gaussian function of varying widths), is equivalent to performing a multiscale analysis using the SMHW. This is because the SMHW is the Laplacian of a Gaussian, and therefore, the anomalous curvature corresponds to the same underlying signature represented by the anomalous wavelet coefficients observed in Vielva et al. (2004) and Cayón et al. (2005). Furthermore, in the same work a statistical connection between the CS and the low variance anomaly was found. The curvature appeared more anomalous when the covariances obtained from the data were used, while the statistical significance decreased using the covariances from the best-fit model. Surprisingly, all extreme values of the derivatives field were found in the southern hemisphere, particularly on the western side, indicating a potential connection with the observed power asymmetry (see the next section).

In Marcos-Caballero et al. (2017b), a multipolar profile analysis was performed for the most prominent peaks in the temperature field, including the CS. These are shown in Figure 3.3, where CS is labelled by 5. The study found that the CS anomaly is mainly caused by its large curvature in the centre, which exhibits a 4σ deviation (see Figure 2 in Marcos-Caballero et al. (2017b)). In particular, when both the peak height and curvature were conditioned to the value observed at the centre, the angular profile of the CS was found to be compatible with the predictions of the ΛCDM model, even reproducing the surrounding hot ring. However, when only the peak height was conditioned, the profile presented a 4.7σ deviation. In Planck Collaboration et al. (2020d), the same analysis was performed on the polarisation data. However, no peaks could be detected because of the low signal-to-noise ratio, highlighting that more sensitive polarisation data would be required.

In summary, the origin of the CS remains unclear and is still an active research area under investigation. Its frequency independence makes an origin in systematics or foregrounds highly unlikely, and also rules out a SZ origin. Some authors have tried to explain the CS through the void hypothesis, suggesting that the presence of a supervoid could produce the observed CS through the ISW and the Rees-Sciama effects (Inoue & Silk, 2006; Finelli et al., 2016). However, in Marcos-Caballero et al. (2016), it was shown that a very rare void would be required as the ISW effect within the standard model cannot reproduce the anomalous CS profile. Other physical mechanisms, including cosmic textures (Cruz et al., 2007a, 2008), have also been proposed, but none have provided a fully satisfactory explanation.

3.2 Hemispherical power asymmetry

In this section, an overview of the Hemispherical Power Asymmetry (HPA) is provided. The first evidence of a hemispherical asymmetry was observed in the analysis of the first-year WMAP data (Eriksen et al., 2004a; Hansen et al., 2004) by studying the angular power spectrum on opposing hemispheres. In particular, two complementary analyses were performed. First, the sky was divided into 164 slightly overlapping discs of 9.5° radius. In each disc the power spectrum was locally estimated in bins of width 3 and compared to the spectra derived from an ensemble of 6144 simulated maps. This analysis revealed that the amplitudes of fluctuations in discs located in the northern (southern) Galactic hemisphere were, in general, lower (higher) in the WMAP data than in simulations. The analysis was then extended by estimating the power spectrum in the northern and southern hemispheres defined by 82 different directions. For each of the 164 hemispheres, the total power was es-

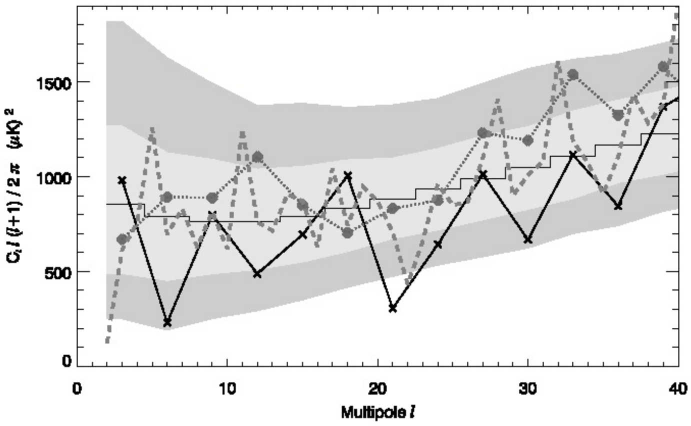

FIGURE 3.4: Temperature angular power spectrum from WMAP. Black crosses represent the binned power spectrum computed from the northern hemisphere, while grey dots correspond to the southern hemisphere. The hemispheres are defined with respect to the axis that maximises the asymmetry. The thin solid black line denotes the best-fit WMAP power spectrum, and grey bands indicate the 1σ and 2σ confidence regions derived from simulations. The dashed line shows the power spectrum computed from the masked WMAP data. Figure taken from Eriksen et al. (2004a).

timated as $C_{\mathrm{hemisphere}} = \sum_b C_b$, and the asymmetry quantified by the ratio between the two hemispheres for each of the 82 orientations. This approach made it possible to identify the axis that maximised the asymmetry without introducing bias from a priori choice of hemispheres. The results showed that, over the multipole range $\ell = 2 - 40$, the ratio observed in the WMAP data exceeded that of at least 99.7% of the simulations, corresponding to a more than 3σ detection of hemispherical asymmetry. Additionally, it was found for some multipole cuts that the axis maximising the asymmetry lied close to the north ecliptic pole. As shown in Figure 3.2, the quadrupole and octopole already contribute to such asymmetry. Figure 3.4 shows the first evidence of the asymmetry, as shown in Eriksen et al. (2004a).

A simple statistically anisotropic model can be constructed by modulating an isotropic map in the following way

$$\delta T_{\mathrm{sky}}(\hat{\mathbf{n}}) = (1 + M(\hat{\mathbf{n}}))\delta T_{\mathrm{iso}}(\hat{\mathbf{n}}), \tag{3.15}$$

where $\delta T_{\mathrm{iso}}(\hat{\mathbf{n}})$ is the underlaying statistically isotropic field, and $M(\hat{\mathbf{n}})$ is the modulating field. The simplest case is a dipolar modulation (Gordon et al., 2005; Gordon, 2007),

$$M(\hat{\mathbf{n}}) = A\hat{\mathbf{d}} \cdot \hat{\mathbf{n}}, \tag{3.16}$$

where A and $\hat{\mathbf{d}}$ are the amplitude and direction of the modulation, respectively. Nevertheless, a higher-order modulation may also be considered, such as a quadrupolar modulation. Alternatively, a modulation in k-space (Zibin & Contreras, 2017) has also been proposed in an attempt to reconcile the observations with the cosmological model, which in practice could be more related to the potential physics behind this anomaly. The dipolar modulation parametrisation has been extensively studied in both WMAP (Eriksen et al., 2007; Hansen et al., 2009; Hoftuft et al., 2009; Akrami et al., 2014) and *Planck* data (Akrami et al., 2014; Ade et al., 2014, 2016; Marcos-Caballero & Martínez-González, 2019; Planck Collaboration et al., 2020d; Gimeno-Amo et al., 2023), and provides a framework within which Bayesian analysis can be performed. In fact, in Eriksen et al. (2007), the hemispherical power asymmetry was studied in the three-year WMAP data by applying a Bayesian analysis. The WMAP data suggested a dipolar distribution of the power with a best-fit modulation amplitude of around 12% in the direction $(\ell, b) = (225^\circ, -27^\circ)$. The probability of finding such modulation in the corresponding simulations was only $\sim 1\%$, and a Bayesian log evidence[8] difference, $\Delta \log E = \log E_1 - \log E_0$, of around $1.5 - 1.8$. These results were robust against sky cuts and frequency channels, suggesting that the observed asymmetry was not affected by foregrounds or masking effects. The same analysis was performed with the five-year WMAP data (see Hoftuft et al. 2009) finding a best-fit modulation amplitude of $A = 0.072 \pm 0.022$ in the direction $(\ell, b) = (224^\circ, -22^\circ) \pm 24^\circ$. In this analysis, the Bayesian log-evidence difference increased to 2.6 showing a strong evidence for the dipolar modulation model.

In the first studies, dipolar modulation was found to mainly affect large angular scales ($\ell \leq 64$). In Hansen et al. (2009), the analysis presented in Eriksen et al. (2004a) was extended to determine whether the asymmetry persists at higher multipoles. A strong correlation was found among the directions of the fitted dipoles characterising the power distribution across six independent multipole ranges between $\ell = 2$ and $\ell = 600$. These directions appeared to be anomalously aligned, indicating that the HPA extends across a broader range of angular scales. Furthermore, the dipole amplitude was found to vanish on small angular scales, for multipoles above $\ell \sim 600$, indicating that the modulation is scale-dependent, $A \rightarrow A_\ell$. These results were confirmed in Axelsson et al. (2013) using the nine-year WMAP data, where variations in cosmological parameters across the sky were also explored. In a more recent work (Marcos-Caballero & Martínez-González, 2019), a scale-dependent modulation model was fitted to the data, finding that the Bayesian evidence supports it against a scale-invariant model. Additionally, the Q-O alignment was found to be more likely in the dipolar modulation scenario than in the standard model, suggesting a potential connection between both phenomena.

The *Planck* team also investigated the HPA using a variety of statistical estimators (Ade et al., 2014, 2016; Planck Collaboration et al., 2020d), assuming that the asymmetry arises from a dipolar modulation. These estimators include: fitting a dipole to the variance map; performing a Bayesian fit of the dipolar modulation model to the temperature map; mea-

---

[8] The Bayesian evidence is given by $E \equiv \int P(\mathrm{d}|\theta, \mathrm{H})\mathrm{P}(\theta|\mathrm{H})\mathrm{d}\theta$, where $P(\mathrm{d}|\theta, H)$ is the likelihood and $P(\theta|H)$ the priors. Essentially, it is the average of the likelihood over the prior volume. By computing this quantity for two hypotheses, H_1 and H_2, the difference can be used to test if one of the models is preferred by the data over the other.

suring the coupling between adjacent multipoles, ℓ and $\ell + 1$, as induced by the anisotropic model; and testing for potential alignments in the directions of power distribution across different multipole intervals. Each of these approaches is described in more detail in the following list.

- **Variance asymmetry:** The Local-Variance method was first introduced in Akrami et al. (2014), where it was applied to the nine-year WMAP and *Planck* 2013 temperature data to study the HPA. It was found that none of the 1000 isotropic simulations exhibited a variance asymmetry as large as the one observed in the `SMICA` temperature map. For WMAP, the statistical significance was slightly smaller, but with a direction fully consistent with that derived from *Planck* data. Additionally, it was found that only the dipole component exhibited an anomalous amplitude in the power spectrum of the variance map.

  The methodology includes two free parameters, the input resolution of the CMB map and the disc size considered for the analysis. In Ade et al. (2016) it was found that the significance levels dropped when larger discs were considered due to the cosmic variance introduced by the largest-scales. By filtering out the low multipoles they showed that the variance asymmetry was well detected in the *Planck* 2015 data (p-value $\leq 0.1\%$) for all discs sizes with directions closely aligned.

  The local-variance method was first applied to the polarisation E-mode in Aluri & Shafieloo (2017), finding a power asymmetry with a p-value around 3% and a preferred direction closely aligned with the CMB dipole.

  The *Planck* 2018 data were analysed in Planck Collaboration et al. (2020d), finding results consistent with previous studies and among the four component-separated maps. Using 1000 simulations to define the probabilities, the p-value was found to be below 0.1%, i.e., none of the simulations exhibited a local-variance dipole amplitude as large as that observed in the data. Taking advantage of the improved control of polarisation systematics, the analysis was extended to the E-mode maps. The resulting p-value was found to vary significantly across component-separated maps, with `Commander` and `Sevem` providing values as low as 0.7% and 0.4%, respectively, while `NILC` and `SMICA` exhibited values above 5%. These differences suggested the presence of residual systematic effects in the component-separated maps. Surprisingly, the direction of the local-variance dipole was found to be closely aligned with that of the temperature analysis. This is shown in Figure 3.5. The probability of finding such alignment in simulations was below $1\% - 2\%$ for three component-separated maps.

  Full details of the methodology are provided in Chapter 4, where we present the results for the variance asymmetry analysis performed on PR4 data set.

- **Dipole modulation:** The basic model adopted in the dipole modulation approach is

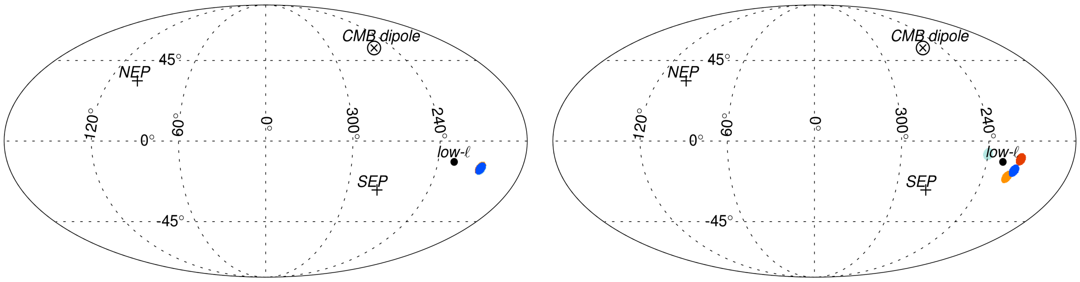

FIGURE 3.5: Dipole directions estimated using the local-variance method in temperature (*left panel*) and polarisation(*right panel*) from *Planck* 2018 data set. Different colours correspond to different component separation methods: `Commander` (red), `NILC` (orange), `Sevem` (green), and `SMICA` (blue). For reference, the black dot, labelled as low-ℓ, indicates the direction obtained from QML estimator. Figure adapted from Planck Collaboration et al. (2020d).

similar to those presented in Eriksen et al. (2007) and Hoftuft et al. (2009). The sky is described as follows,

$$\mathbf{d}(\hat{\mathbf{n}}) = B(1 + M(\hat{\mathbf{n}}))\mathbf{d}_{\rm iso}(\hat{\mathbf{n}}) + \mathbf{n}(\hat{\mathbf{n}}), \tag{3.17}$$

where $M(\hat{\mathbf{n}})$ is the dipolar modulation described in Eq. 3.16, and B and $\mathbf{n}$ refer to the convolution by the instrumental beam and the instrumental noise, respectively. In Ade et al. (2016), the *Planck* 2013 temperature data were directly fitted to this model using a multivariate Gaussian likelihood,

$$\mathcal{L}(A, \hat{\mathbf{d}}, q, n) \propto \frac{e^{-\frac{1}{2}\mathbf{d}^t \mathrm{C}^{-1}\mathbf{d}}}{\sqrt{|\mathrm{C}|}}, \tag{3.18}$$

where C is the covariance matrix, given by

$$\mathrm{C} = \mathrm{BMS_{iso}M^T B^T} + \mathrm{N}, \tag{3.19}$$

with $\mathrm{M_{ij}} = (\mathrm{A}\hat{\mathbf{d}} \cdot \hat{\mathbf{n}}_{\rm i})\delta_{\rm ij}$. Here, $\mathrm{S_{iso}}$ and N denote the covariance matrices of the underlying isotropic signal and noise, respectively. In particular, the covariance matrix for the isotropic signal, $S_{\rm iso}$ was analytically computed from the power spectrum, modelled as

$$C_\ell(q, n) = q\left(\frac{\ell}{30}\right)^n C_{\ell,\rm fid}, \tag{3.20}$$

where $C_{\ell,\rm fid}$ is the best-fit power spectrum. These two parameters, q and n, were included to accommodate the low variance of low-ℓ. As already mentioned in Section 3.1, the significance of the CS was larger by considering the anomalous low variance.

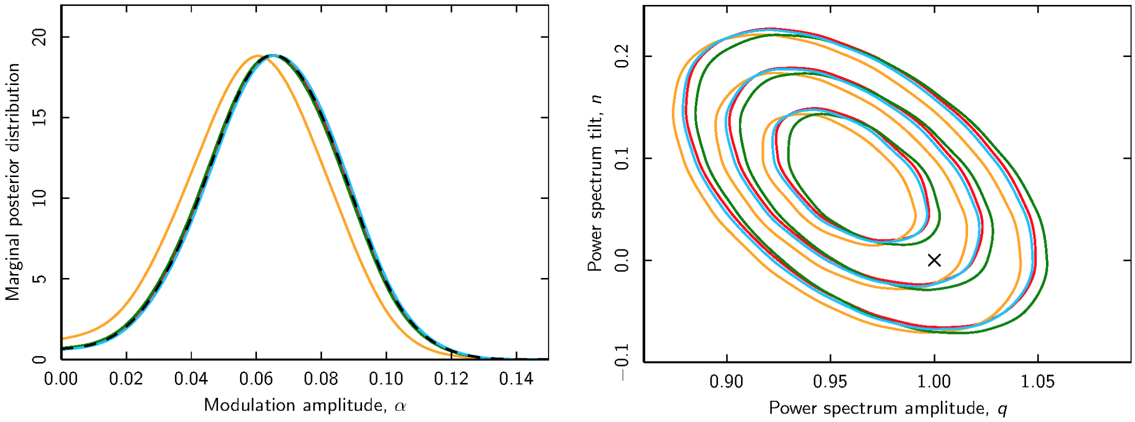

FIGURE 3.6: Marginal posteriors on the dipolar modulation amplitude (*left panel*) and (q, n) parameters modelling the power spectrum in Eq. 3.20 (*right panel*), derived from *Planck* 2015 temperature data. Constraints for `Commander` (red), `NILC` (orange), `Sevem` (green), and `SMICA` (blue) are provided. Figure adapted from Ade et al. (2016).

The analysis was carried out at low resolution, using a `HEALPix`[9] (Górski et al., 2005) pixel resolution of $N_{\rm side} = 32$, and smoothing with a Gaussian function with a full width at half maximum (FWHM) ranging from 5° to 10°. The analysis also included a set of foreground templates in the covariance matrix to marginalise over. The significance was found to vary with the smoothing scale. In fact, the largest significance was found for 5°, with all the the component-separated maps exhibiting significances $\sim 3\sigma - 3.5\sigma$, and amplitude and direction values of $\sim 7\%$ and $(\ell, b) = (226^\circ, -16^\circ)$, respectively. Furthermore, these results were consistent with those derived from WMAP.

The analysis was updated with the *Planck* 2015 data set (Ade et al., 2016), confirming the results of the previous data release, i.e., best-fit dipole modulation amplitude of $6 - 7\%$ and a deficit of around $3 - 5\%$ of power in low-ℓ compared to the best-fit ΛCDM model. Figure 3.6 shows the marginalised constraints for dipolar modulation and power spectrum parameters taken from Ade et al. (2016).

- **Bipolar spherical harmonics (BipoSH) formalism:** As previously mentioned, the power spectrum only contains all the statistical information under the statistical isotropy and Gaussianity assumptions. In this scenario, the 2-point correlation function only depends on the distance, and can be estimated by averaging all the temperature products, $\delta T(\hat{\mathbf{n}}_1)\delta T(\hat{\mathbf{n}}_2)$, between all pairs of pixels with the same angular separation, θ. In harmonic space, this translates to null off-diagonal terms of the covariances $\langle a_{\ell m}, a_{\ell', m'} \rangle$. In Hajian & Souradeep (2003), a κ_ℓ statistic, based on the Bipolar

---

[9] Hierarchical Equal Area isoLatitude Pixelation (`HEALPix`) is a scheme for partitioning the sphere into equal-area pixels, commonly used in CMB analyses. The resolution is specified by the parameter $N_{\rm side}$, where the total number of pixels on the sphere is given by $N_{\rm pix} = 12N_{\rm side}^2$. Larger values of $N_{\rm side}$ correspond to higher angular resolution.

Spherical Harmonic (BipoSH) formalism, was introduced to quantify the breakdown of statistical isotropy.

If statistical isotropy does not hold, the 2-point correlation function becomes direction-dependent and poorly determined, since only one product is available for each pair of pixels for a single sky realisation. To extract rotationally meaningful information about potential statistical isotropy violation, the 2-point correlation function is expanded in the BipoSH basis,

$$C(\hat{\mathbf{n}}_1, \hat{\mathbf{n}}_2) = \sum_{\ell,\ell',L,M} A_{\ell\ell'}^{LM} \{Y_\ell(\hat{\mathbf{n}}_1) \otimes Y_{\ell'}(\hat{\mathbf{n}}_2)\}. \tag{3.21}$$

The $A_{\ell\ell'}^{LM}$ are the BipoSH coefficients,

$$A_{\ell\ell'}^{LM} = \sum_{mm'} \langle a_{\ell m} a^*_{\ell' m'} \rangle \, (-1)^{m'} \mathcal{C}_{\ell m \ell' -m'}^{LM}, \tag{3.22}$$

where $\mathcal{C}_{\ell m \ell' -m'}^{LM}$ are the Clebsch-Gordan coefficients.

The BipoSH provides a complete framework for characterising any deviation from statistical isotropy. This decomposition allows simultaneously to identify the nature of the anisotropy described by the bipolar multipole L, and determine if it is concentrated in a specific multipole range. In particular, for a dipolar anisotropy the information is encoded in L = 1, while quadrupolar anisotropy corresponds to L = 2, octopolar to L = 3, and so on. The standard angular power spectrum corresponds to $A_{\ell\ell}^{00}$ coefficients since the correlations between different multipoles vanish.

The BipoSH coefficients for the model described in Eq. 3.16 are given by the following expression (Ade et al., 2016),

$$\begin{aligned} A_{\ell\ell+1}^{1M} &= \bar{A}_{\ell\ell+1}^{1M} + m_{1M} G_{\ell\ell+1}^{1}, \\ G_{\ell\ell+1}^{1} &= \frac{C_\ell + C_{\ell+1}}{\sqrt{4\pi}} \sqrt{\frac{(2\ell+1)(2\ell+3)}{3}} \mathcal{C}_{\ell 0 (\ell+1) 0}^{10}, \end{aligned} \tag{3.23}$$

where m_{1M} are the spherical harmonic coefficients of the modulation field, and $\bar{A}_{\ell\ell+1}^{1M}$ the BipoSH coefficients of the underlying unmodulated CMB field[10]. The m_{1M} coefficients can be obtained using a minimum variance estimator as done in Ade et al. (2014), and in fact, it can be done over specific angular scales. Furthermore, the amplitude of the modulation field is related to the power in the modulation field, $m_1 = (|m_{11}|^2 + |m_{10}|^2 + |m_{1-1}|^2)/3$, by $A = \frac{3}{2}\sqrt{m_1/\pi}$.

This formalism was used in Ade et al. (2014) using the *Planck* 2013 data. A dipole modulation signal (L = 1) was detected between 3σ and 4σ level (see left panel in Fig-

---

[10] This term includes the cosmic variance, and the bias introduced by the mask, which can be quantified through simulations.

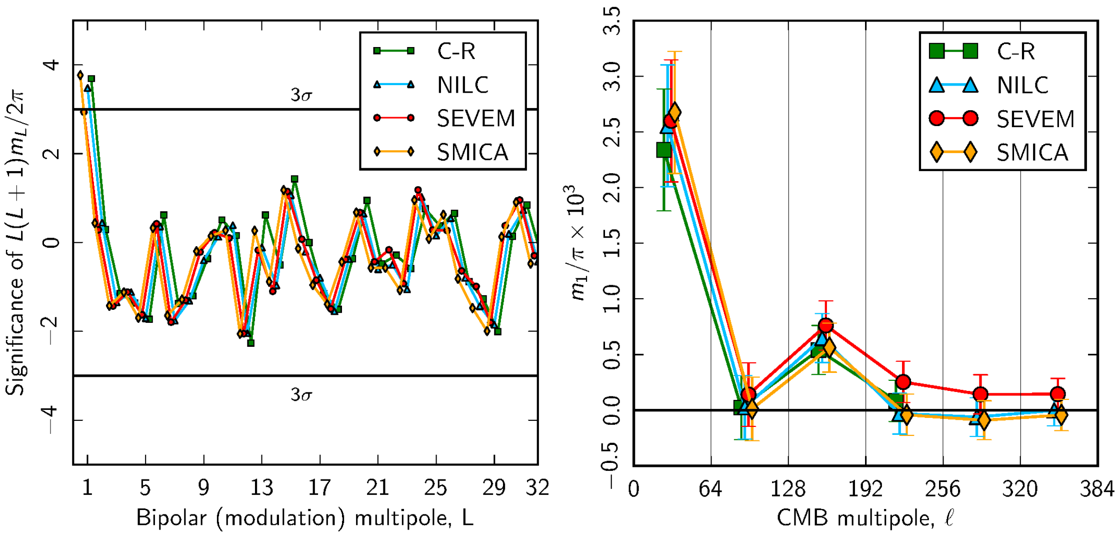

FIGURE 3.7: *Left panel:* Significance of the m_L for different bipolar multipoles estimated from *Planck* 2013 data. Dipolar modulation (L = 1) is detected with a significance between 3 and 4σ. *Right panel:* Dipole modulation power across different multipole bins. Figure adapted from Ade et al. (2014).

ure 3.7) on large angular scales ($\ell \leq 64$) with an amplitude ($A \sim 0.07$) and direction compatible with those inferred from the direct fitting. As shown in the right panel of Figure 3.7, no significance amplitude was detected for higher multipoles. Additionally, the left panel also displays the significance of the modulation power for other bipolar multipoles. All values are within 2σ, and in particular, no quadrupolar modulation was observed.

These results were confirmed in Ade et al. (2016) using the *Planck* 2015 maps. For the `SMICA` map the estimated modulation amplitude and direction were found to be $A = 0.069 \pm 0.022$ and $(\ell, b) = (228°, -18°) \pm 30°$, respectively. Similar values were obtained for the other component-separated CMB maps. The analysis was also performed using the `Sevem` cleaned frequency maps, finding no evidence for frequency dependence. Even if no statistical evidence was found for the amplitude on higher multipoles, the directions were seen to be clustered together.

- **Quadratic Maximum Likelihood (QML) analysis:** An analysis closely related to the BipoSH formalism was performed in Ade et al. (2016). In this case, the couplings between ℓ and $\ell \pm 1$ modes in the CMB covariance matrix, induced by a scale-invariant dipolar modulation, were investigated using a QML approach (see Section 7.2 and Appendix C of Planck Collaboration et al. (2020d) for further details). In the multipole range $\ell = 2 - 67$, a p-value of $0.9\% - 1.0\%$ was reported, which increased to 10% after accounting for the "look-elsewhere" effect.

  This method also allows for the estimation of both the amplitude and the direction of the modulation. The values obtained from the QML analysis were found to be fully

Table 3.1: Dipolar modulation amplitude and direction from *Planck* 2018 `SMICA` map. Values taken from Planck Collaboration et al. (2020d). The errors are calculated from the FFP10 simulations [11], including the systematic effects and statistical uncertainties.

Data	Amplitude	Direction (ℓ, b) [deg]
`SMICA` (TT)	$0.070^{+0.032}_{-0.015}$	$(221, -22) \pm 31$
`SMICA` (TT, TE, EE)	$0.068^{+0.032}_{-0.015}$	$(221, -19) \pm 31$

consistent with those derived from the BipoSH and Bayesian approaches. For instance, using the `SMICA` map, the inferred modulation amplitude was $A = 0.062^{+0.026}_{-0.013}$, with a corresponding direction of $(\ell, b) = (213^\circ, -26^\circ) \pm 28^\circ$, where uncertainties were estimated from statistically isotropic CMB realisations.

The analysis was extended to polarisation in Planck Collaboration et al. (2020d). The results obtained from the temperature data were found to be consistent with previous analysis, while the addition of polarisation data alongside temperature had a minimal impact on the inferred dipolar modulation amplitude and direction, as shown in Table 3.1. No evidence for dipolar modulation was found when analysing the polarisation data alone. Finally, the amplitude inferred from temperature in the multipole range $\ell = 2 - 220$ was found to be around a factor of 3 smaller than that obtained in the $\ell = 2 - 64$ range.

- **Angular clustering of the power distribution:** The first indications of angular clustering were found in Hansen et al. (2009) by computing the local power spectrum in discs binned in 6 independent blocks of multipoles, ranging from $\ell = 2$ to $\ell = 600$, and fitting for a dipolar asymmetry in the resulting maps. The six dipoles pointed to a similar direction.

  The *Planck* team adopted an alternative model-independent estimator to quantify angular clustering (Ade et al., 2014, 2016; Planck Collaboration et al., 2020d). A detailed description of the methodology is provided in Chapter 6. In summary, the method fits dipoles to maps that encode the spatial variation of power within specific multipole intervals, and then, quantifies the degree of alignment among the resulting dipole directions. The level of alignment in the data is compared to that obtained from simulations in order to assess its statistical significance. In fact, this analysis does not focus on the amplitude of the dipoles, which may be consistent with the expected distribution of values for isotropic CMB realisations, but rather on whether there is a preferred direction. The presence of such alignment constitutes a clear signature of power asymmetry. In the standard cosmological model, the multipoles are uncorrelated, and consequently, the power distribution associated with different multipole

---

[11] A set of end-to-end simulations produced by the *Planck* team incorporating realistic models of the instrument, noise properties, foreground emission, scanning strategy, and CMB signal.

intervals are expected to be statistically independent, with their preferred directions randomly oriented across the sky. Note that the significance of the clustering should not be affected by noise properties and systematic effects, as long as those are correctly included in the simulations.

Applying this analysis to the *Planck* 2013 data revealed a clear alignment between dipole directions up to $\ell_{\max} = 600$, with none of the simulations exhibiting a mean alignment between all dipoles as that observed in the data. Remarkably, all directions were found to lie close to the HPA axis reported in previous works. These results were shown to be robust across different component-separated maps. The analysis was also performed on the `Sevem` 143 GHz cleaned frequency map, corrected by the Doppler modulation effect[12], showing that the statistical significance of the asymmetry decreases above $\ell_{\max} \sim 600$. The robustness of these results was further supported by applying three global statistics, all of them finding strong evidence for asymmetry.

The analysis was also extended to the *Planck* 2015 data in Ade et al. (2016), where the Rayleigh statistic was introduced to quantify the alignment of dipole directions. A detailed description of this statistic is provided in Chapter 6. In this case, the p-values were found to remain low even for $\ell_{\max} > 750$. After applying different analysis choices, playing with the $\ell_{\min}$ and $\ell_{\max}$, it was finally concluded that the measured clustering was a consequence of the existing correlations between directions of both high and low multipoles, with the effect extending up to $\ell_{\max} = 1500$. Remarkably, for $\ell_{\max} < 100$, the p-values did not indicate a low-ℓ asymmetry, in contrast to the results obtained with the previously described methods. In fact, this discrepancy may be related to the high variance of the Rayleigh statistic in this regime, which is a consequence of the limited number of available bins.

In Planck Collaboration et al. (2020d), the analysis was extended to polarisation by analysing the *Planck* 2018 data. The temperature results confirmed those obtained in the *Planck* 2015 analysis, with p-values consistently below 1% for multipoles up to $\ell_{\max} \approx 1200$. For the OE and HM data splits, the p-values started to be above 1% for $\ell_{\max} < 1000$, which may be attributed to the inclusion of the unobserved pixels mask. For the polarisation E-mode signal, a hint of alignment was observed at multipoles around $\ell = 150 - 250$, with p-values reaching levels below 1%. A similar trend was observed by analysing the correlations between directions obtained from TT and EE analysis, while the TE was consistent with simulations.

In Chapter 6, we present the results of the angular clustering analysis using the PR4 temperature and polarisation maps. We extend the analysis to search for dipolar features in the cosmological parameters, mainly motivated by the observed anomalous anisotropic distribution of the power in temperature data, as well as by recent studies

---

[12] Unlike aberration effect, Doppler modulation was not included in the simulations.

(Fosalba & Gaztañaga, 2021; Yeung & Chu, 2022), where a directional dependence of the cosmological parameters is claimed.

3.2.1 Inpainting CMB data

As discussed in Section 2.2.4, one of the main issues in characterising the CMB anisotropies is how to deal with extragalactic and Galactic foreground emissions such as synchrotron, free-free, or thermal dust. These are astrophysical emissions between the last scattering surface and us. The CMB has been measured in several frequency channels to exploit their different frequency response and to allow component-separation algorithms to subtract an important part of them.

However, certain regions, such as the Galactic plane and the locations of bright extragalactic point sources, remain strongly contaminated and are not suitable for statistical analyses, even after applying component-separation techniques. In these regions, foreground emissions are too intense and complex to be effectively cleaned. The standard approach is to mask such regions using the confidence masks provided by each component-separation method. In practice, this leads to a loss of information and other technical difficulties. For example, it prevents accurate characterisation of large angular scales, which is crucial for the study of CMB anomalies.

In the case of polarisation, masking presents even more complications, not just at harmonic level but also in real space. In particular, masking Q and U Stokes parameters introduces an undesired mixing between E- and B-modes (Tegmark & de Oliveira-Costa, 2001; Lewis, 2003), commonly referred to as E-to-B leakage. This is critical for the detection of the primordial B-mode signal since the E-mode signal is significantly larger. In real space, the lack of full-sky Q and U maps prevents an accurate reconstruction of E- and B-modes as a result of the non-local nature of the transformations. In fact, this challenge motivates the work described in Chapter 5. Reconstructed maps are of interest for studies of isotropy and statistics, including the anomalies present at large angular scales. It is worth noting that for power spectrum estimation in masked skies, other standard techniques are available. For example, in Alonso et al. (2019) a very fast algorithm is described (`NaMaster`), based on the pseudo-$\mathcal{C}_\ell$ formalism, to estimate the angular power and cross-power spectra. This implementation is applicable to both spin-0 and spin-2 fields and is able to correct for the coupling between multipoles induced by the mask. Nevertheless, it is known to be suboptimal on large angular scales. An alternative approach is the QML method (Tegmark, 1997; Tegmark & de Oliveira-Costa, 2001; Bilbao-Ahedo et al., 2021), which is optimal for low multipoles but computationally demanding. Given these constraints, two strategies can be adopted for the statistical analysis of CMB anomalies: (1) adapt the estimators to partial sky maps, which in general are more complicated and numerically expensive, or (2) fill the maps using the so-called inpainting algorithms.

Inpainting techniques aim to reconstruct a full-sky CMB map statistically coherent with the observed data. Inpainting is not a novel concept, it was already used in image processing to fill in missing pixels and restore blurred photographs (Masnou & Morel, 1998). This idea was extended to the Cosmology field and applied to fill the CMB sky (Abrial et al., 2008; Inoue et al., 2008; Perotto et al., 2010; Bucher & Louis, 2012; Kim et al., 2012). The simplest approach is a diffuse inpainting, where the algorithm iteratively fills the masked pixels by averaging the neighbouring unmasked pixels. More sophisticated methodologies have also been developed, such as the purified inpainting employed in (Planck Collaboration et al., 2020d), which does not assume any underlying cosmological model.

In Chapter 5, we consider an alternative approach based on Gaussian constrained realisations (Hoffman & Ribak, 1991; Bucher & Louis, 2012; Kim et al., 2012; Benoit-Lévy et al., 2013; Marcos-Caballero & Martínez-González, 2019). Assuming that the field is Gaussian, we use the best-fit power spectrum to estimate the pixel correlations and fill the missing pixels drawing samples from the corresponding conditional probability distribution. This is explained in more detail in Chapter 5. Although an exact implementation is not feasible for high-resolution maps, where correlations for millions of pixels are needed, this method can be useful for low-resolution maps, up to $N_{\text{side}} = 64$ and $\ell_{\max} = 192$. In fact, this range covers the full range of the reionisation and the recombination peaks of the B-mode. In this context, such an approach could be a powerful tool for future experiments, such as LiteBIRD. The use of Neural Networks has also been proposed as an alternative way to inpaint the CMB (Puglisi & Bai, 2020).

An implementation of this inpainting technique has been employed in Chapter 4 to study the HPA in the E-mode maps.

Part II

Main results of this thesis

The second part of this thesis presents my contributions to the cosmology field. Except for the abstracts, which have been adapted to align with the structure of the thesis, Chapters 4, 5, and 6 reproduce the content of the following publications Gimeno-Amo et al. (2023, 2024, 2025):

- P.1: **Gimeno-Amo, C., Barreiro, R.B, Martínez-González, E., & Marcos-Caballero, A. (2023).** Hemispherical power asymmetry in intensity and polarization for *Planck* PR4 data. *Journal of Cosmology and Astroparticle Physics*, JCAP12(2023)029.
- P.2: **Gimeno-Amo, C., Martínez-González, E., & Barreiro, R.B, (2024).** CMB-PAInT: An inpainting tool for the cosmic microwave background. *Journal of Cosmology and Astroparticle Physics*, JCAP09(2024)038.
- P.3: **Gimeno-Amo, C., Hansen, F.K., Martínez-González, E., Barreiro, R.B, & Banday, A.J. (2025).** Exploring Statistical Isotropy in *Planck* Data Release 4: Angular Clustering and Cosmological Parameter Variations Across the Sky. *arXiv e-prints* arXiv:2504.05597.

During the thesis I have performed three studies in the context of the CMB anomalies and the statistical isotropy of the universe. In particular, I have studied the HPA by applying the local-variance estimator to the PR4 temperature and polarisation E-mode maps. For this, it was necessary to implement an optimal inpainting approach. I did that in my second publication by developing a `Python` code, `CMB-PAInT`, based on the Gaussian constrained realisation method. Finally, I have also studied the HPA from the angular clustering estimator perspective, where I also include an analysis on the isotropy of the cosmological parameters by studying their spatial distribution across the sky.

Apart from these works, as an active member of the LiteBIRD collaboration, I have contributed to several studies which are out of the scope of this thesis. These contributions include:

- Development of an optimal likelihood approach for the future detection of the polarisation B-modes.
- Study of the impact of HWP-related systematic effects on the calibration of polarisation angles.
- Forecasting activities:
  - Bayesian analysis of cosmological parameters using a MCMC sampler with the reconstructed lensing-only spectrum. The results of this activity are reported in Ruiz-Granda et al. (2025).

- An analysis on LiteBIRD's capability to test CMB anomalies using polarisation E-modes. This work is directly related to the topic covered in this thesis. The aim of this work is to investigate the ability of cosmic variance limited E-mode measurements to constrain the fluke hypothesis, i.e., the possibility that the CMB temperature anomalies are statistical fluctuations. We use several estimators described in Chapter 3, and in particular, the `CMB-PAInT` code is applied. The results of this activity are reported in Banday et al. (2025).

4

Hemispherical power asymmetry in intensity and polarisation for *Planck* PR4 data

The purpose of this chapter is to study and characterise the hemispherical power asymmetry on the new *Planck* Data Release 4 by performing an asymmetry analysis on both intensity and polarisation maps. We apply the local-variance estimator introduced in Chapter 3. As a first step, we reanalyse the PR3 data inpainted using a method based on Gaussian constrained realisations, which is described in detail in Chapter 5. The results are consistent with those previously reported by the *Planck* collaboration. For PR4, using the `Sevem` component-separated maps, we find that the p-value is slightly reduced to 2.8%, i.e., only 2.8% of the E2E simulations exhibit a local-variance dipole amplitude in the E-mode as large as that observed in the data, while the dipole direction remains close to the that derived in PR3. Furthermore, we identify a hint of a potential T-E alignment between the asymmetry axes at the level of approximately 5%. Our results indicate that more sensitive all-sky CMB polarisation data, such as those expected from the future LiteBIRD experiment, will be necessary to reach a more robust conclusions on the possible existence of deviations from statistical isotropy in the form of a hemispherical power asymmetry.

4.1 Data and Simulations

In the present analysis we have used the *Planck* data[1] from the latest two releases, the *Planck* 2018 full-mission release (PR3) (Planck Collaboration et al., 2020a), and the fourth release (PR4), which has been processed by the NPIPE pipeline (Planck Collaboration et al., 2020f). Both pipelines use the LFI and HFI raw and uncalibrated data to generate frequency maps in the HEALPix (Górski et al., 2005) format, 9 in the case of temperature and 7 for Q and

---

[1] *Planck* space telescope, operated by the ESA, had two instruments, the Low Frequency Instrument (LFI) and the High Frequency Instrument (HFI), with the goal of measuring the total intensity and polarisation of CMB photons in a wide frequency range from 30 to 857 GHz (30 to 353 for polarisation). This range was covered by 9 different frequencies, with the three lowest frequencies measured by the LFI (30, 44 and 70 GHz) and the rest by the HFI (100, 143, 217, 353, 545 and 857 GHz).

U Stokes parameters as the two highest frequency channels were not sensitive to polarisation. These frequency maps are propagated through the component separation methods (Planck Collaboration et al., 2020b), `Sevem` (Fernández-Cobos et al., 2012), `SMICA` (Cardoso et al., 2008), `NILC` (Basak & Delabrouille, 2012), and `Commander` (Eriksen et al., 2008), which generate each a clean CMB map. For PR3 we have used the data from the four separation algorithms. For PR4, clean CMB maps have been provided only by `Sevem` and `Commander`. We use both maps for the analysis in intensity and only `Sevem` products for polarisation.[2] All the data and simulations are available at NERSC[3] at full resolution[4] ($\mathrm{N_{side}} = 2048$), and have been downgraded (to $\mathrm{N_{side}} = 64$) in harmonic space, according to

$$a_{\ell m}^{out} = \frac{b_\ell^{out} p_\ell^{out}}{b_\ell^{in} p_\ell^{in}} a_{\ell m}^{in}, \tag{4.1}$$

where the superscripts in and out refer to the HEALPix resolution of the input and output maps, respectively, the $a_{\ell m}$'s are the harmonic coefficients of the considered maps, the p_ℓ's give the corresponding pixel window functions, and the b_ℓ's are the Gaussian smoothing functions with a FWHM of $5'$ for the original cleaned CMB maps and of $160'$ for the downgraded maps. The output $a_{\ell m}$'s are then converted to a pixel map considering a maximum multipole $\ell_{\max} = 3\mathrm{N_{side}} - 1$. For this procedure, we have used the Healpy package[5] (Zonca et al., 2019), which is basically a Python wrapper for HEALPix. `Sevem` maps for both *Planck* releases are shown in Figure 4.1. Differences in polarisation between both releases are apparent, mainly due to the reduction of large-scale systematics in the PR4 data. Conversely, intensity maps are very consistent.

All the details of the PR4 processing and the differences with respect to PR3 are given in Planck Collaboration et al. (2020f). Most notably, while previous *Planck* pipelines process LFI and HFI in an independent way, PR4 was designed to process them simultaneously which produces a significant improvement in the calibration of both instruments. In addition, PR4 reduces the noise by including the data from the repointing manoeuvres, which is around 8% of the total mission. These differences, together with other added features and modifications, have the net effect of reducing especially the polarisation systematics at large scales as well as the level of noise. However, the signal at the smallest multipoles is also reduced due to the resulting transfer function (see Section 4.2.4 and Figure 4.6).

In order to test the null hypothesis, we have used as the reference point the FFP10 PR3 and NPIPE Monte Carlo (MC) simulations, provided by the *Planck* collaboration along with the data maps. This set of simulated maps includes realisations of the CMB signal, the

---

[2] Although `Commander` polarisation products are also available for PR4, we have not included them in this work, since the associated simulations were not well suited for our analysis (see Section 4.3.3).

[3] National Energy Research Scientific Computing Center (NERSC), `https://www.nersc.gov/`, is a primary scientific computing facility operated by Lawrence Berkeley National Laboratory, located in California. It provides high-performance computing and storage facilities where *Planck* latest data and simulations can be found.

[4] Data is also available at *Planck* Legacy Archive (PLA), `https:/pla.esac.esa.int/`.

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

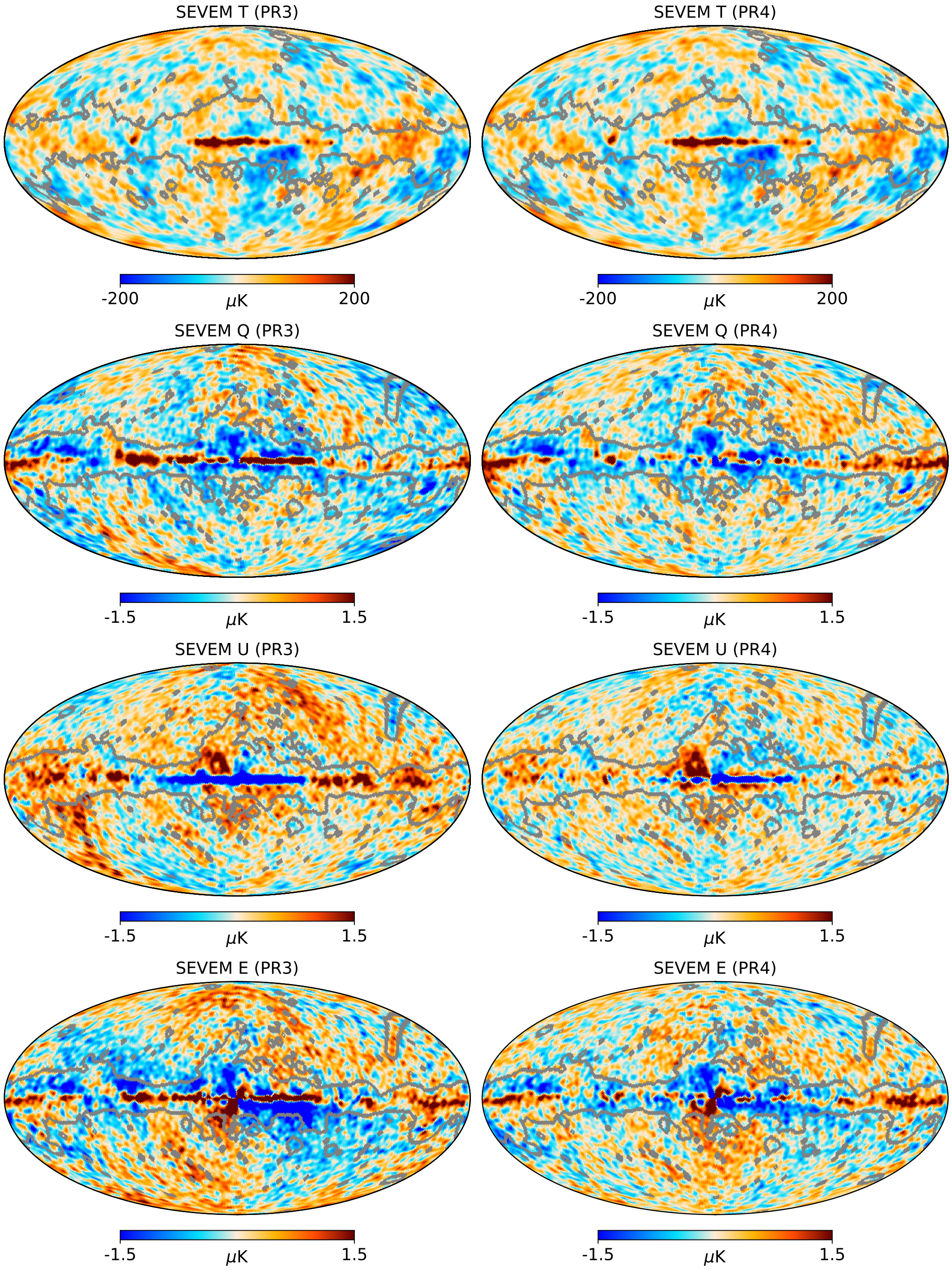

FIGURE 4.1: `Sevem` component-separated CMB maps at 160′ resolution. The first column shows T, Q, U, and E-mode maps for the *Planck* 2018 pipeline (PR3), while the second column shows the same maps for NPIPE (PR4). E-mode maps are derived from the full-sky Stokes Q and U maps. The grey lines indicate either the temperature or the polarisation *Planck* component separation confidence mask (Planck Collaboration et al., 2020b).

instrumental noise and the systematics, which try to capture the characteristics of the full data processing such as the scanning strategy, the detector responses, or the calibration errors introduced during data reduction. For both PR3 and PR4, a number of realisations of the CMB sky are generated according to a fiducial CMB power spectrum based on the cosmology described in Table 6 of Planck Collaboration et al. (2020f). These sky realisations include lensing, Rayleigh scattering, and Doppler boosting. Additionaly, they are convolved by the frequency-specific beam. In contrast, noise simulations have been obtained following a more complex approach (see Section 4 of Planck Collaboration et al. (2020d) and Section 5 of Planck Collaboration et al. (2020f)), processing them using the same algorithm as for the real data via an *end-to-end* (E2E) pipeline, including instrumental noise and systematics. For PR3, after propagating E2E simulations through component-separation pipelines, the final product is a set of 999 CMB and 300 noise simulations for each of the four methods. On the other hand, for the PR4 pipeline, 600 CMB and noise realisations are available for `Sevem`, while 100 simulations are provided for intensity in the case of `Commander`.

Regarding masks, we use the PR3 intensity and polarisation component separation confidence masks downgraded to $N_{\rm side} = 64$, which retains the 71.3% and 72.4% of the sky, respectively. Following Planck Collaboration et al. (2020d), to generate lower resolution binary masks, we first smoothed them according to Eq. 4.1 and then set a threshold of 0.95. All the pixels with a value below the threshold are set to 0, while the remaining ones are set to 1. Masks at $N_{\rm side} = 2048$ are available in the *Planck* Legacy Archive[6] (PLA). The intensity mask is used without modification for the local variance analysis in the temperature maps. On the other hand, the polarisation mask is used as the starting point for the inpainting of the Q and U maps and to generate a set of extended customised masks for the analysis of the E- mode map, as explained in Section 4.2.2.

4.2 Methodology

In this section, we first present an overview of our analysis and then explain in detail the different steps of the methodology in the corresponding subsections. In the current work we perform an exhaustive analysis of the hemispherical power asymmetry anomaly for intensity and polarisation following the methodology of Planck Collaboration et al. (2020d), where the local-variance estimator is used to find dipolar-like features. The estimator is introduced in Subsection 4.2.3, and essentially the output is the amplitude and direction of the local-variance dipole, which can be use as a tracer of the modulation.

To apply this analysis to the polarisation data, we first need to construct an E-mode map. This is a non-local quantity, and ideally we need full-sky maps of the Q and U Stokes parameters to have an accurate reconstruction of the E-mode map. However, since CMB observations are strongly contaminated by residual foregrounds in certain regions of the

---

[6] https://pla.esac.esa.int/

sky, we cannot consider in the analysis those pixels excluded by the polarisation confidence mask. To solve this problem, a useful approach is to carry out inpainting in the contaminated regions, a process that assigns new values to the excluded pixels such that they are more consistent with the rest of the CMB map. In particular, we have implemented a method based on *Gaussian Constrained Realisations* (GCR) which is used to inpaint the regions masked by the polarisation confidence mask in the Q and U data maps. In Subsection 4.2.1 the inpainting method is described in more detail. We note that this technique requires the knowledge of the covariance matrix of all the components present in the map. For the CMB signal, this can be obtained analytically given a cosmological model, whereas for the noise and systematics component, we compute it from the E2E noise simulations. However, to avoid overfitting (due to the lack of convergence of the estimated matrix because of the relatively small number of simulations), the simulations considered in the analysis must be independent from those used to compute the covariance matrix. Therefore, we are forced to split the simulations in two halves, the first set used to estimate the covariance matrix and the second one to obtain the distribution of the amplitudes of the local-variance dipole and the corresponding p-value.[7] Different partitions of the simulations are also considered in order to test the robustness of the results. In particular, Subsection 4.2.2 describes how the p-value is estimated taking into account the different data splits.

Another important step is to check the validity of the E-mode map constructed from the inpainted Q and U data maps. In particular, using the E2E simulations, we generate an extended customised mask by imposing a threshold in the pixel error of the reconstruction of the E-mode map. This mask is then used for the analysis of the polarisation data (see Subsection 4.2.2 for details on how this mask is constructed).

Finally, Section 4.2.4 presents some tests to validate our methodology. We check with the E2E simulations that the estimated amplitude and direction of the local-variance dipole are not biased as a consequence of the inpainting process. We also give a comparison between the inpainting scenario and the one where E-modes are derived from a simple masking of the Q and U Stokes parameters. Furthermore, assuming a dipolar modulation model (Eq. 3.16), we obtain some interesting results such as the sensitivity of the method taking into account the realistic *Planck* noise and systematics.

4.2.1 Inpainting using Gaussian constrained realisations

One of the main issues when analysing the CMB is how to deal with foreground emissions such as dust, free-free, or synchrotron. Although component separation algorithms are able to reduce these emissions significantly, some residuals are expected to remain near the Galactic plane and in the locations of point sources, where the sky is strongly contaminated. The standard approach is to mask these regions, but, as mentioned before, this leads to other

---

[7] Note that the p-value is defined in an frequentist way as the fraction of simulations with an amplitude of the estimator equal or larger than the one obtained from the data.

difficulties. This is especially the case when dealing with polarisation data, since masking Stokes parameters introduces an undesired mixing between E- and B-mode (Tegmark & de Oliveira-Costa, 2001; Lewis, 2003). This effect has an impact when reconstructing the E-mode map, but it can be minimized by using an optimal inpainting technique. Inpainting consists on replacing the contaminated pixels by values which are somewhat consistent with the cleaned data, keeping as much as possible the statistical properties of the underlying field. Different inpainting approaches have been considered in CMB analyses. For instance, diffuse inpainting, which consists on replacing the value of the contaminated pixels by averages of those of the neighbouring pixels, works acceptably well to inpaint small regions, such as those corresponding to point sources (Planck Collaboration et al., 2020b). A second more sophisticated approach is purified inpainting, which constructs a full-sky polarisation map from a masked one where most of the pure E- and B-modes are projected out in order to minimize the leakage. It has been used in the Isotropy and Gaussianity analyses performed by the *Planck* Collaboration (Planck Collaboration et al., 2020d), including the study of the hemispherical power asymmetry. In the present work, we use an alternative approach based on GCR. This technique was introduced in Marcos-Caballero & Martínez-González (2019) for temperature, but we have extended it here for the Stokes parameters.

GCR is a method that works in the pixel domain. The idea is to fill the masked pixels by sampling from the conditional probability distribution, $p(\hat{\mathbf{d}}|\mathbf{d})$, where $\hat{d}$ is the vector of the inpainted field and d is the vector of the available pixels. Assuming that the field is Gaussian, which is a very good approximation for the CMB, all we need is the pixel covariance matrix. For the CMB signal, this can be computed from the theoretical power spectra[8] following, for instance, the Appendix A[9] of Tegmark & de Oliveira-Costa (2001), whereas for the noise plus systematics part, we need to rely on simulations. We present a full description of the method in Chapter 5. Although, in principle, this method is able to fill the contaminated pixels with a realisation that is perfectly consistent with the remaining data according to the assumed statistical properties of the underlying field, it presents two main limitations. First of all, this approach is very demanding from the computational point of view, since it requires the computation and storage of the pixel covariance matrix, whose dimension is $2N_{\text{pix}} \times 2N_{\text{pix}}$ when considering the inpainting of the Q and U Stokes parameters ($3N_{\text{pix}} \times 3N_{\text{pix}}$ if also TE correlation is considered). Hence the resolution at which we can work is limited by the required memory. Due to this constraint, we have carried out our analyses at a Healpix resolution of $N_{\text{side}} = 64$. The second limitation comes from our lack of knowledge of the part of the covariance matrix coming from noise and systematics in the *Planck* polarisation data, which needs to be characterised using simulations. In the case of the `Sevem` method for PR4, we have a total of 600 E2E simulations (600 for the CMB and 600 for the noise and

---

[8] We use the ΛCDM best fit power spectra available in the PLA.

[9] Note the small typo in equation A7 of the appendix where the proportionality constant is negative, not positive. Thus, in equation A8, the negative sign corresponds to the case where z component of the vector $\hat{r}_{ij}$ is positive, and vice versa.

systematics). As previously mentioned, to avoid overfitting[10] in the inpainting process, we need to split them into two sets of 300 independent simulations, such that the first set is used to characterise the noise contribution to the covariance matrix, while the second one is used to obtain the distribution of the amplitude of the local-variance dipole that will be compared with the data. The limited number of E2E noise simulations is not enough to achieve a good convergence of the noise covariance matrix, and consequently, matrix elements, especially the off-diagonal terms, are significantly noisy. Indeed, using semi-realistic noise simulations[11], we have found that the number of simulationes required to achieve convergence is at least of a few thousand. Another approximation is to ignore the, in general, non-Gaussian behaviour of systematics, that is not taken into account in the inpainting. However, in spite of these limitations, we find that this proposed inpainting procedure is still very useful for our purpose. Indeed, it is not necessary to reproduce the full statistical properties of the inpainted pixels in order to minimise significantly the E-to-B leakage due to incomplete sky.

4.2.2 Performance of the inpainting, confidence mask, and estimation of the p-value

The performance of the inpainting procedure can be determined by comparing the exact E-mode map obtained from full-sky Q and U maps of the E2E CMB plus noise simulations, with that obtained from the same Q and U maps that have been inpainted in the region given by the *Planck* polarisation confidence mask. In particular, we compute the map of E-mode residuals, ΔE, as the difference between the exact and inpainted maps, for the pixels outside the mask for each of the inpainted simulations. We also obtain the standard deviation of the residuals at each unmasked pixel ($\sigma_{\Delta E}$) as well as the one of the input E-mode map (σ_E). In the case of PR3, this procedure is repeated for the simulations propagated through the four component separation method, while for PR4 is done for `Sevem`. A relative error map δE is then constructed as follows:

$$\delta E = \frac{\sqrt{\sum_m \sigma^2_{\Delta E}(m)}}{\sqrt{\sum_m \sigma^2_E(m)}} \tag{4.2}$$

where m runs over the different component separation methods used for PR3 or PR4. Note that for PR4, since only `Sevem` is considered, this is simply the ratio between the dispersion of the residuals and that of the input E-mode map at each pixel. Figure 4.2 shows a comparison between the map of the relative error in the E-mode map that we get by applying inpainting

---

[10] Overfitting arise from the non convergence of the covariance matrix. If we use the same simulations to estimate the matrix and perform the analysis, the matrix will have exactly the correlations given by the sample of simulations, and this will produce an overfitting when inpainting is applied on them.

[11] Gaussian simulations that include anisotropic and correlated noise according to the noise covariance matrix obtained from the *Planck* noise E2E simulations.

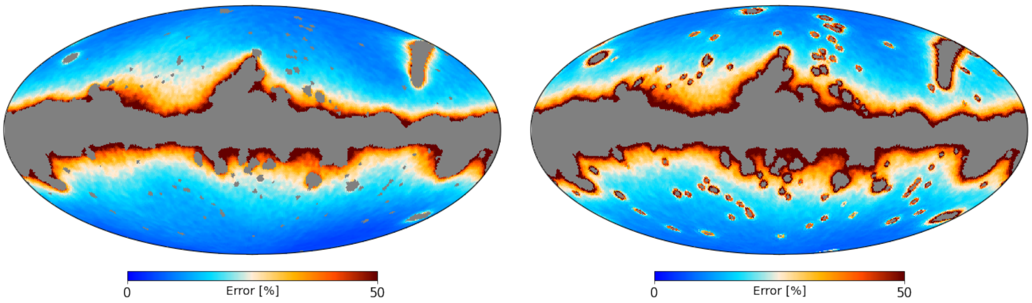

FIGURE 4.2: Comparison of relative error maps (in percentage) in the reconstruction of the E-mode for PR4. *Left panel:* error map obtained with the inpainting approach. *Right panel:* error map obtained for the simple Q and U masking approach (no inpainting). Grey area corresponds to the PR3 polarisation confidence mask.

to the PR4 data (left panel) and by directly using masked Q and U maps (right). It becomes apparent that the inpainting works quite well, especially for point sources and isolated regions outside the Galactic plane. Moreover, in the ideal case in which the covariance matrix is perfectly known, we have found that the performance is even better.

Despite having a more accurate reconstruction of the E-mode using inpainting, the residuals in regions close to the mask boundaries are still significant. One possibility to reduce these residuals is to extend the polarisation confidence mask. In particular, we define a set of extended masks by selecting different thresholds in the error map, which are determined from a compromise between the maximum admissible error and the minimum extension of the mask. Figure 4.3 shows the variation of the maximum error with $f_{\rm sky}$[12] for PR4. By removing the pixels more affected by the inpainting, an improvement in the maximum error is rapidly achieved, going down to values of $\delta E \sim 40\%$ with only a moderate reduction of the sky. However, further improvements require the removal of a significant fraction of the sky as a trade-off. The complete set of masks will be used to check the robustness of the results against the considered sky fraction, but we select as reference mask the one with the threshold of $\delta E = 40\%$ because, for PR4, it leaves a fraction of sky similar to that of Planck Collaboration et al. (2020d).

This way of generating confidence masks is similar to the one followed in Planck Collaboration et al. (2020d), although with some differences. The first one is the resolution, since in that work the mask is generated at $N_{\rm side} = 1024$ and then downgraded to $N_{\rm side} = 64$, while we generate it directly at the final resolution (mainly due to our limitation to work with high resolution maps). The second difference comes from the way in which the threshold is chosen, since in their work the residuals of both the E and B-modes are considered. We decided to focus only on the E-mode reconstruction for several reasons: B is consistent with zero in the *Planck* data, the E-mode signal is much larger than the B-mode, and finally,

---

[12] The sky fraction is actually determined from a combined mask obtained from the study of different splits. More specifically, the combined mask is obtained as the product of the individual extended masks generated for each of the considered splits (see below).

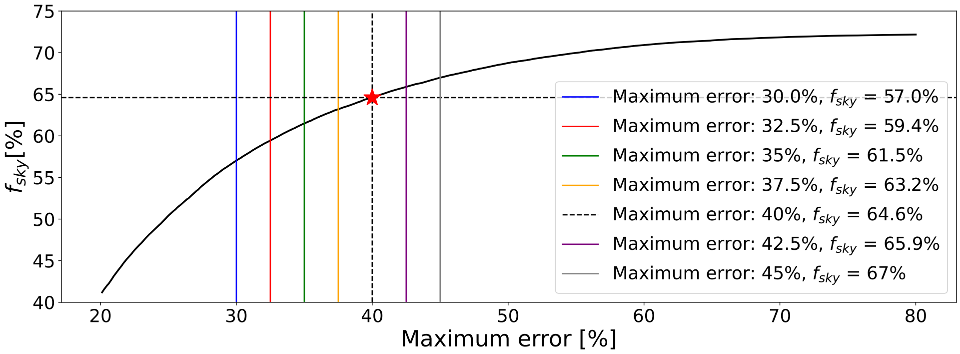

FIGURE 4.3: Sky fraction as a function of the allowed maximum error for the PR4 data set. Vertical lines represent the thresholds for the different masks used in the analysis, covering from 45% to 30% error. The red star shows the maximum error and the fraction of the sky for our reference mask, which is used for the main results of this work and plotted in the right panel of Figure 4.4.

we are only applying the estimator to E-mode maps.

Despite the increase in the error of the estimated dipolar modulation parameters due to the non-convergence of the noise covariance matrix, the amplitudes remain unbiased (see Section 4.2.4). However, splitting the simulations into two halves introduces additional uncertainty in the determination of the p-value. Then, the question is: how should we split the simulations? Although any random split is equally valid, each partition will yield a different p-value due to the slight variation in the amplitude distribution. There are also other effects that contribute to this dispersion but at a lower level. For example, the estimated noise covariance matrix is also different, which in turn affects slightly the way in which the inpainting is performed.

To get some additional insights on which is the best procedure to obtain the required p-value using the limited number of simulations, we have made some tests (see below) using semi-realistic noise simulations. These are constructed as Gaussian realisations characterised by the covariance matrix obtained from the full set of 600 E2E noise simulations[13]. Therefore, these semi-realistic simulations of anisotropic noise contain the same correlations and statistical information up to second order as the realistic *Planck* simulations. The possible presence of non-Gaussianity is the only feature that we cannot simulate. According to these simulations, a good estimation of the underlying p-value, the one that we would get if we sampled the distribution of amplitudes using the complete set of 600 simulations, can be achieved if we proceed as described below (this is detailed for PR4, but a similar procedure is applied to PR3):

1. We generate 30 random splits from the 600 E2E noise simulations. For each split we have one pair of independent sets of 300 simulations each, i.e., a total of 60 sets in total.

---

[13] More specifically we calculate the Cholesky decomposition of the covariance matrix that is then used to generate the Gaussian realisations with the required statistical properties (see e.g. (Barreiro et al., 2008)) for PR4.

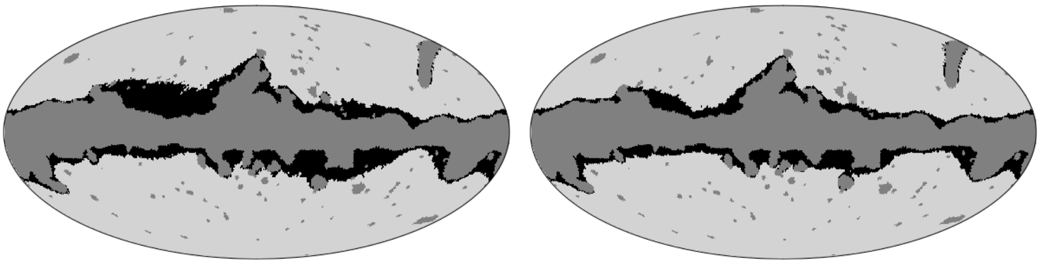

FIGURE 4.4: *Left panel:* customized mask used for the PR3 data set analysis. It leaves 61.4% of the sky unmasked. *Right panel:* customized mask used for the PR4 `Sevem` data set analysis. It leaves 64.6% of the sky unmasked. The grey region corresponds to the polarisation confidence mask provided by *Planck*, while black region corresponds to the extended area where the relative error in the E-mode reconstruction is larger than the 40%.

2. For each split, we use the first set to estimate the noise covariance matrix and we apply inpainting on the second set. We repeat the same procedure but this time estimating the matrix with the second set and applying it to the first one.

3. For each of the estimated covariance matrices (a total of 60), we construct an inpainted map from the real data.

4. For each of the 60 inpainted sets, we generate an extended confidence mask by thresholding the residuals as explained previously for the reference case (i.e., maximum δ_E of 40 per cent). These are then combined into a single extended mask as the product of all of them. We repeat this procedure for different thresholds to have a set of extended confidence masks. Figure 4.4 shows the reference extended masks for PR3 and PR4, used to obtain the main results of our analyses.

5. For each of the inpainted sets and inpainted data map (in each case inpainted using the same covariance matrix), and using the combined reference confidence mask, we apply the local-variance estimator (see Section 4.2.3) to estimate the amplitude of the dipole. Finally, we define the p-value as the fraction of simulations with an amplitude equal or greater than the one observed in the data map.

6. We get the mean of the 60 calculated p-values as our estimation of the underlying p-value. For the direction, we get the mean of each of the vector components, and then, convert it to Galactic longitude and latitude coordinates.

To establish the validity of the previous approach, the test we performed was the following one.

1. First, we generate 2000 CMB simulations modulated in Q and U according to equation 3.16. We use an amplitude (A_{DM}) and direction consistent with the ones estimated from the data, $A_{DM} = 9\%$ and $(\ell, \text{b}) = (235^\circ, -17.5^\circ)$.

2. We add a semi-realistic noise simulation to each modulated simulation (we call them modulated set).

3. We generate an independent set of 600 CMB simulations (unmodulated) and we add semi-realistic noise simulations, also independent from the previous 2000 realisations.

4. Using as reference this set of 600 simulations, we estimate the p-value of each simulation in the modulated set. For this purpose, we apply the local-variance estimator (explained in Section 4.2.3) to the ideal E-mode maps, reconstructed from the full sky Q and U maps (i.e., without applying inpainting). Just to clarify, we still apply an extended mask to the E-mode maps for the analysis. We label these p-values as the *true* ones, since they are the best possible estimation obtained from the full reference set, i.e. using all the 600 simulations.

5. Now let us focus on one of the simulations from the modulated data set. We treat it in the same way as we do for the data, and together with the 600 reference set we follow the steps described above, i.e., we generate 30 partitions of $300/300$ simulations and we compute 60 p-values for the data (including the inpainting process as described previously). The same method is followed for the other 1999 simulations from the modulated set.

6. We get the average value of the 60 p-values, finding that this estimator is a good representation of the previously estimated *true* value. Taking the distribution of differences between *true* and average values, we compute the bias as the mean. As expected, the bias largely depends on the original p-value. In particular, if we restrict ourselves to simulations whose *true* p-value is less than 8% (in the real data the p-value varies between $2-3\%$), the bias is below $1/300$.[14]

7. Although the p-values obtained from different splits are strongly correlated as they are generated from the same set, we provide the range of p-values containing 68% of the distribution as a rough estimation of the uncertainty.

The same procedure has been followed for the PR3 data set, except that, in this case, we have only 300 E2E noise simulations to construct the 30 different partitions. Therefore, for each split, we have one pair of independent sets of 150 noise simulations each (used to construct the covariance matrix). In addition, we also have 999 CMB simulations, for which we consider 900 in our study. In this way, we construct sets of 450 simulations that are used for the analysis, obtained by combining one noise simulation with three different CMB simulations.

---

[14] To test the robustness of this result, we have studied the bias using an independent set of 600 semi-realistic simulations, finding very consistent results.

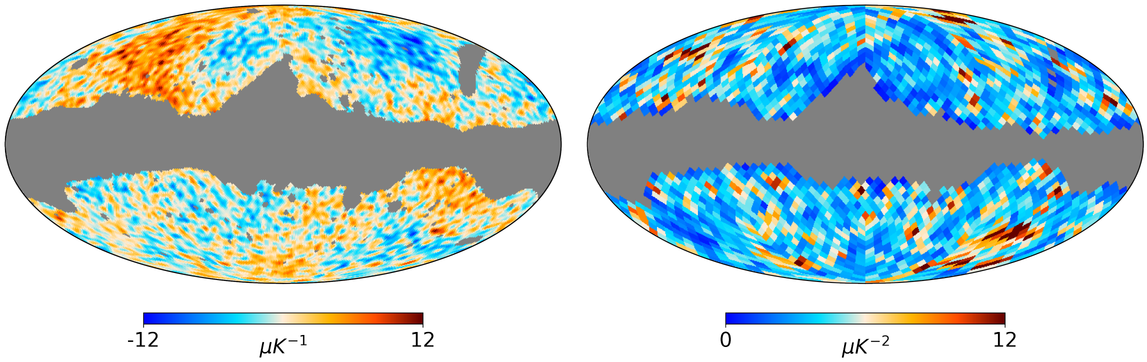

FIGURE 4.5: *Left:* `Sevem` PR4 E2E E-mode simulation after applying the transformation given by Eq. 4.3. The map is masked with an extended mask allowing a sky fraction of 63.2% *Right:* Local-variance map generated following the previous steps. The grey region corresponds to those pixels discarded since more than 90 per cent of the pixels associated to the disc are masked. A disc of 4° is used.

4.2.3 Local-variance estimator

As previously mentioned, the estimator we have used to characterise the asymmetry is the local-variance introduced by Y. Akrami (Akrami et al., 2014). The main motivation is that a dipolar modulation of the anisotropies would manifest itself as a dipolar structure in a map generated by computing variances over sky patches. The procedure to obtain the amplitude and direction of the dipolar modulation is the following:

1. This first step is exclusive to the analysis of polarisation data due to the low signal-to-noise ratio. As mentioned in Planck Collaboration et al. (2020d), if we apply directly the method to the data, then the analysis returns local-variance dipoles whose directions are not uniformly distributed across the sky (see Figure 32 in Planck Collaboration et al. (2020d)). Actually, their distribution is strongly correlated with the structure of the anisotropic noise. Therefore, we consider the following transformation directly applied on the input E-mode maps:

$$X_i' = (X_i - M_i)/\sigma_i^2 \tag{4.3}$$

   where X_i is the value of the E mode map at pixel i computed from the full-sky inpainted Q and U maps. M_i and σ_i are the mean value and standard deviation at pixel i obtained from inpainted simulations.

2. We fix a specific HEALPix low resolution and we define a set of discs of a certain radius centred in the pixel centroids of the HEALPix map. Following previous works, we use $N_{\text{side}} = 16$ and discs of 4 degrees for the present analysis.

3. We identify all the pixels of the map at the original resolution (in our case, $N_{\text{side}} = 64$ for polarisation and $N_{\text{side}} = 64$ and 2048 for temperature) that are inside each of

the discs, compute their variance only taking into account the unmasked pixels, and associate these values to the pixel in the low resolution map. This is what we know as local-variance map (LVM). If more than 90% of the pixels inside a disc are masked, then we mask the pixel in the LVM and remove it from the analysis. An example of one PR4 simulation is shown in Figure 4.5.

4. From the set of the LVM of the simulations, we estimate the mean $\bar{y}_i$ and the variance $\sigma_{d,i}^2$ on each disc, corresponding to one pixel at resolution $N_{\mathrm{side}} = 16$.

5. Finally, we fit a dipole to each of the local-variance maps by applying a weighted χ^2

$$\chi^2 = \sum_i \frac{[(y_i - \bar{y}_i) - d_0 - \mathbf{d} \cdot \hat{\mathbf{r}}_i]^2}{\sigma_{d,i}^2} \tag{4.4}$$

where y_i is the LVM at pixel i; d_0 captures the monopole component of the LVM; $\mathbf{d} = (d_x, d_y, d_z)$ is the dipole component; and $\hat{\mathbf{r}}_i$ is the unit vector pointing to the i-th pixel. Sum is over all non-masked pixels in the LVM. d_0 and $\mathbf{d}$ are obtained by minimizing the χ^2, which is analytic in this case.

4.2.4 Sensitivity and validation with simulations

Before going into the results, we assess the sensitivity of the method by considering CMB simulations that contain a dipolar modulation, taking also into account the realistic E2E PR4 noise and systematics simulations. In particular, we remark that for PR4, a transfer function that affects only the lowest multipoles of the polarisation is present in the data (see Section 4.3 and Figure 20 from Planck Collaboration et al. (2020f)). In order to take this effect into account for the E-mode map, we have computed the transfer function over the full-sky[15] for the `Sevem` data as:

$$f_\ell^{EE} = \left(\frac{1}{N_{\mathrm{sim}}} \sum_{j=1}^{N_{\mathrm{sim}}} \frac{C_{\ell,j}^{\mathrm{CMBx(CMB+N)}}}{C_{\ell,j}^{\mathrm{CMB}}} \right)^2 \tag{4.5}$$

where $C_{\ell,j}^{\mathrm{CMB}}$ corresponds to the EE power spectrum of the ith PR4 `Sevem` CMB simulation and $C_{\ell,j}^{\mathrm{CMBx(CMB+N)}}$ to the cross-spectrum between that simulation and the same one after adding the associated noise and systematics `Sevem` simulation[16]. Figure 4.6 shows the full-sky E-mode transfer function estimated for `Sevem`.

Starting with the E2E `Sevem` CMB simulations, we construct modulated maps using Eq. (3.16). Each simulation is modulated for 14 different amplitudes (going from 5% to 18%) and using a direction compatible with the one measured in the data (see Section 4.3.3), (ℓ, b) = $(235°, -17.5°)$. The modulation is applied to the Q and U maps, as a direct modulation

---

[15] Note that the transfer function is actually anisotropic, but using the full-sky transfer function is a sufficiently good approximation for our purpose.

[16] The effect of the transfer function is actually included in the noise and systematics simulation, which presents therefore a correlation with its associated CMB simulation.

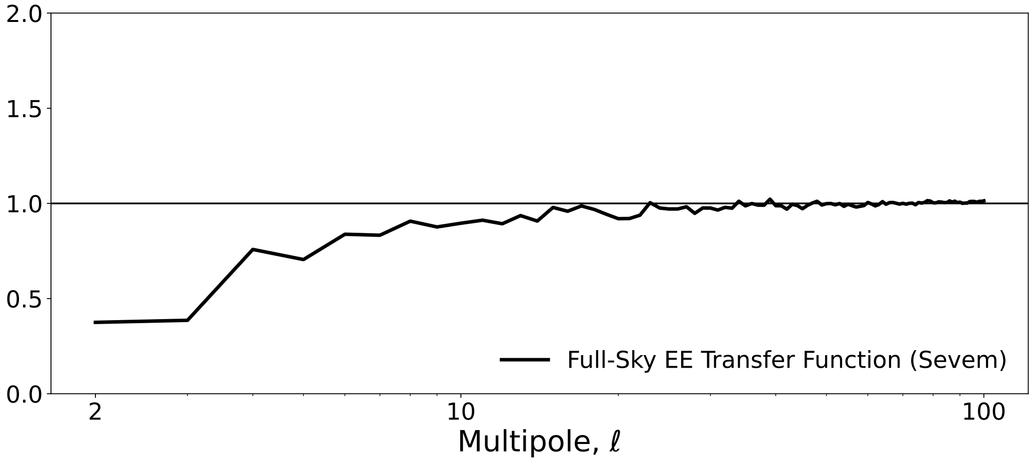

FIGURE 4.6: `Sevem` PR4 full-sky E-mode transfer function obtained from equation 4.5.

in the E-mode map is not physically allowed (Kothari, 2018). The transfer function is then applied to the modulated CMB map[17]. Finally, the E2E noise simulations are added to the simulated CMB signal[18]. All the tests presented here have been done using the customized mask with $f_{\rm sky} = 63.2\%$. Although the specific results may be slightly different using other masks, the validation and main conclusions of this subsection still hold.

The results of this subsection have been obtained by combining a number of splits of the 600 modulated E2E simulations. In particular, as previously explained, we consider 30 different splits, a total of 30 pairs of 300 E2E noise simulations. Just to make it clearer, for each considered amplitude of the dipolar modulation we enumerate the most important steps of the procedure:

1. For a given split, we have set A and set B, both of them with 300 CMB and noise simulations. We use one of them, for example set A, to compute the noise covariance matrix, and we apply inpainting on the set B. This inpainted set is what we call unmodulated analysis set.

2. From the modulated E2E simulations we keep the ones whose associated noise simulations are not included in the set A. We apply inpainting on them, using the covariance matrix estimated from set A, and we treat them in the same way as the data (we call this the modulated data set).

3. We compute the local-variance dipole amplitude and direction for the unmodulated

---

[17] Starting at $\mathrm{N_{side}} = 2048$, we downgrade the maps to $\mathrm{N_{side}} = 512$ removing the Gaussian beam of $5'$ and the pixel window function. Then, we modulate the Q and U maps, and we downgrade them to $\mathrm{N_{side}} = 64$ applying to the $\mathrm{a}_{\ell\mathrm{m}}$ the square root of the `Sevem` E-mode transfer function and the Gaussian beam of $160'$. We go through an intermediate resolution to avoid the numerical errors in the functions that convert from map to Fourier space.

[18] Note that the i^{th} CMB simulation is added to an independent noise simulation (the i^{th}+1 one) to avoid including twice the transfer function.

set and for the modulated data set, and we estimate the p-value.

4. Then, we can repeat this process but now using set B to estimate the covariance matrix. Additionally, we repeat the same procedure for the A/B sets of the other 29 splits. At the end, we get 30 p-values for each modulated E2E simulation[19]. This means that for each considered amplitude of the modulation, we have a total of 18000 local-variance amplitudes and p-values (30 splits x 600 modulated simulations), which are the values used to generate all the figures of this section. Although these values are not independent, they provide us with additional information with respect to use a single split and allow us to construct an approximated distribution of the estimated amplitudes.

The left panel of Figure 4.7 shows the distribution of the p-value of the estimated amplitude of the dipolar modulation obtained as previously explained. The results correspond to simulations modulated with an amplitude of 9%, which is close to the value found for the data (see Section 4.3.3). According to this figure, we would be able to find a p-value lower than 1% in 39.3% of the cases, while this percentage increases to 61.7% for a p-value lower than 5%. Additionally, the right panel shows similar information but for all the considered amplitudes. In particular, the percentage of times that we get a p-value equal to or lower than 1%, 5% and 32% for each amplitude is given. For *Planck* PR4 we find a p-value lower than 1% in at least 95% of the times for amplitudes larger than 16%. The amplitude reduces to 14% if we consider a p-value lower than 5%. Using the set of modulated simulations we can also calibrate the method to give a relation between the local-variance and the dipolar modulation amplitudes (see Section 4.3.3).

We have also checked how well the method performs in estimating the direction of the dipolar modulation by using the modulated simulations with an amplitude of 9% (selected taking into account the results from Section 4.3.3). In this case, we consider the ideal case of Q and U full-sky maps, to study the performance of the method itself, without including possible distortions introduced by the inpainting technique. We compare the longitude and latitude measured in the local-variance maps for the 600 modulated simulations with the input values (fixed to $(\ell, \mathrm{b}) = (235°, -17.5°)$), and compute the angular distance between them. Figure 4.8 shows the distribution of the angular distance, where bins are not uniformly distributed, but instead they are selected so the area covered by each of them is constant. As seen, the dispersion in the estimated direction is quite large, even in this ideal case, although one should take into account that the considered amplitude is relatively small.

We have carried out a final test to study the specific effect of the inpainting in both the estimated amplitude and direction of the dipolar modulation. For this, we have estimated those quantities from the 600 (unmodulated) E2E simulations in the ideal case in

---

[19] Note that for the case of the real data, since the data are independent of the simulations, we can obtain 2 p-values for each split (i.e. both data sets in the split can be used to construct the covariance matrix) and thus obtain a total of 60 p-values.

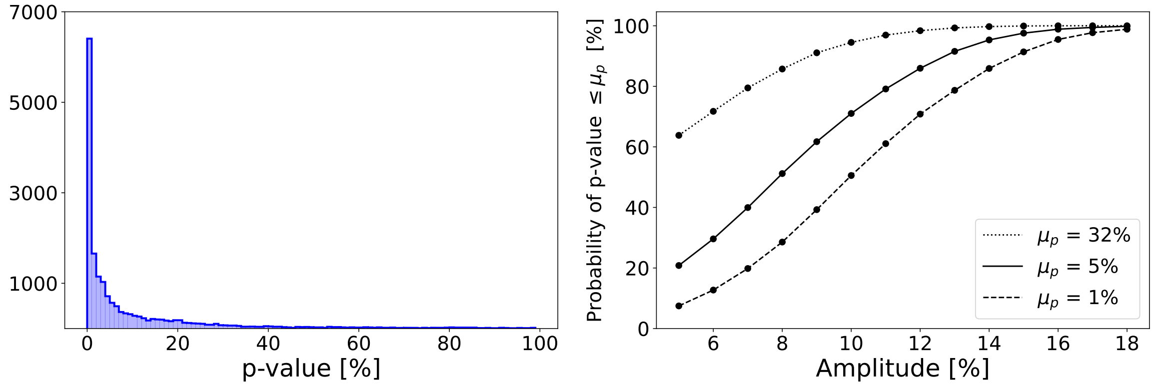

FIGURE 4.7: *Left panel:* p-value distribution of the estimated amplitude of the dipolar modulation for the modulated simulations. We use an amplitude of 9% and a fixed direction of (ℓ, b) = $(235°, -17.5°)$. *Right panel:* probability curves as a function of the dipolar modulation amplitude with the realistic PR4 noise and systematics. Assuming a dipolar modulation in the Stokes Q and U parameters, we could claim a detection with a p-value lower than 1% (dashed line) in 95% of the cases if the amplitude was larger than 16%. Solid (dotted) lines correspond to the probabilities of getting a p-value equal to or lower than 5% (32%).

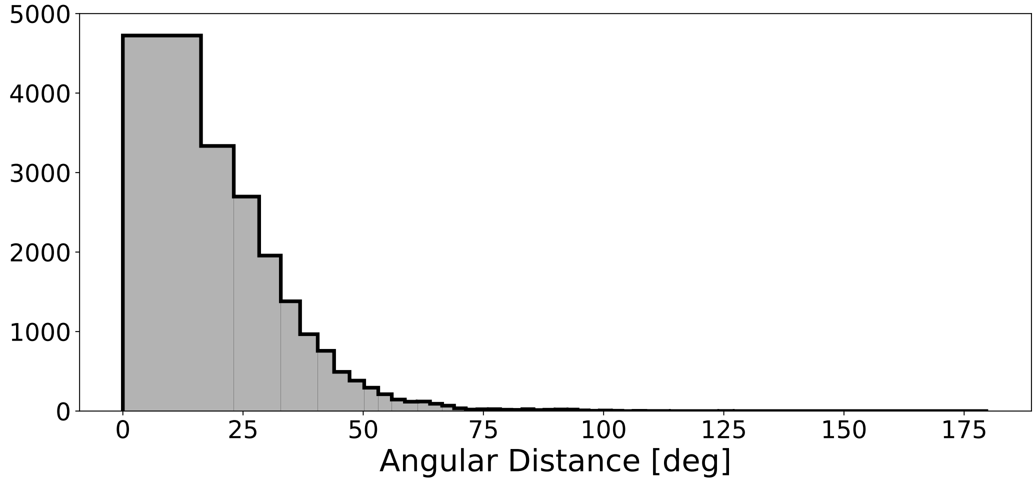

FIGURE 4.8: Distribution of angular distance between the input direction (ℓ, b) = $(235°, -17.5°)$ and the one obtained after applying the LVM estimator on the modulated (A_{DM} = 9%) PR4 simulations. The ideal case, in which the Q and U full sky maps are known, is considered.

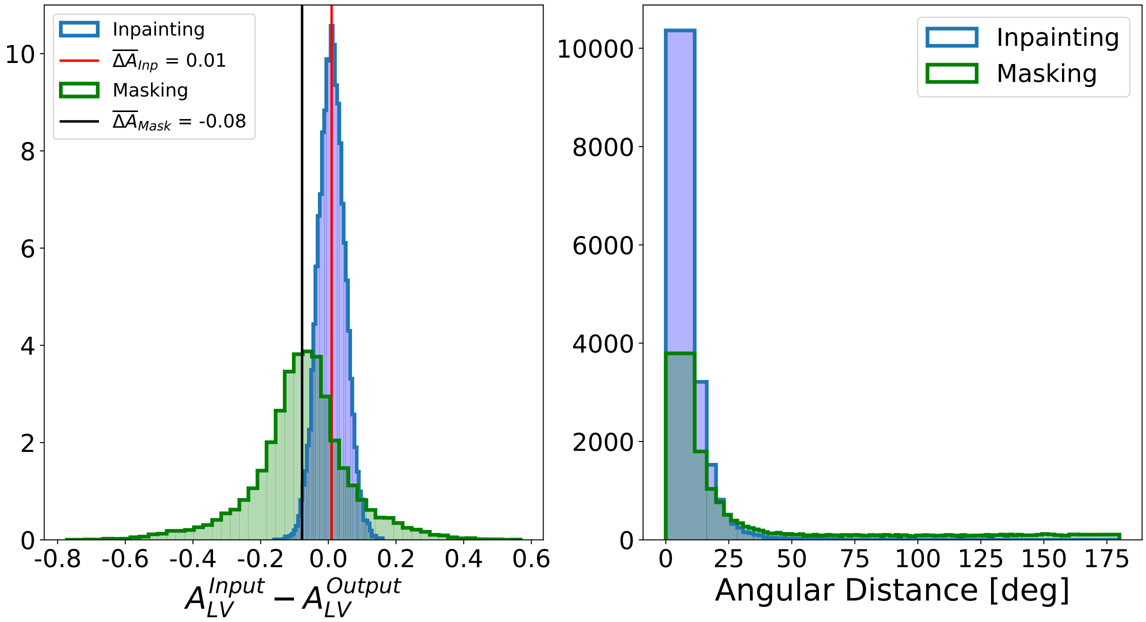

FIGURE 4.9: *Left panel:* distribution of ΔA_{LV}, defined as the difference between the amplitude measured on the inpainted (blue) or masked (green) simulations (denoted as output) and the one obtained when full-sky simulations are used (denoted as input). Red and black lines correspond to the mean of the blue (inpainting) and green (masking) distributions, respectively. *Right panel:* angular distances (in degrees) between the direction measured on the inpainted simulation and the input ones (blue). Input values are obtained from the corresponding full-sky (non-inpainted) simulations, so we remove the dispersion introduced by the estimator itself by making the difference. In green the same quantity is plotted for the case where a simple mask is applied on the Stokes parameters.

which Q and U full-sky are considered (that we call *input* case). Then we have estimated the same quantities using inpainting. For comparison, a simple masking approach, where the masked area is simply replaced by zeroes in the Q and U maps, has also been considered. The left panel of Figure 4.9 displays the distribution of the difference ΔA between the input amplitude and the one recovered from the inpainting or the simple masking approaches. It becomes apparent that the inpainting technique improves significantly with respect to simply masking the data. Indeed, on average, the latter tends to overestimate the amplitude, while no significant bias is found for the former. Furthermore, the dispersion of the distribution is significantly larger for masking (0.11) than the one we get when using our inpainting technique (0.04).

On the right panel, we also display the angular distances for the same scenarios. Again, we are not using uniform bins. From the figure, it is clear that the inpainting provides a better reconstruction of the direction than the simple masking. We also note that although the dispersion introduced by the inpainting in the estimation of this quantity is significant, it is clearly below the intrinsic dispersion of the method (see Figure 4.8). Furthermore, the median of the blue (inpainting) distribution is 10 degrees, while for the green one (masking) is almost 34 degrees.

4.3 Results

In this section, we show the results of the analysis of PR4 intensity and E-mode polarisation data. For each case, we present the p-values and directions. We also re-examine the HPA in the PR3 data set using our alternative inpainting approach to check consistency with previous results.

4.3.1 Intensity results for PR4

Following previous temperature analysis, we present the results of full-resolution, $N_{\rm side} = 2048$, PR4 data for a set of discs with radii between 4 and 40 degrees. The p-values are defined as the fraction of simulations with local-variance dipole amplitude larger than the one observed in the data. Unlike PR3, where cleaned maps were provided for four methods, for PR4, CMB maps are only available for `Sevem` and `Commander`. As expected, both of them are in good agreement with the *Planck* 2018 analysis. For `Sevem`, none of the 600 simulations have an amplitude as large as the one observed in the data for discs of 4, 6, and 8 degrees. This means that, under the ΛCDM model assumption, the probability of having such asymmetry in the temperature sky in that range of angular scales is below 0.17%. For radii larger than 8° the p-value increases systematically. We have checked with simulations that have a dipolar modulation pattern, with A = 7% and p = $(209^\circ, -15^\circ)$, that this increase of the p-value with the radius of the disc is expected. We obtain the same behaviour for `Commander`, where none of the 100 simulations shows an amplitude as large as the one measured in

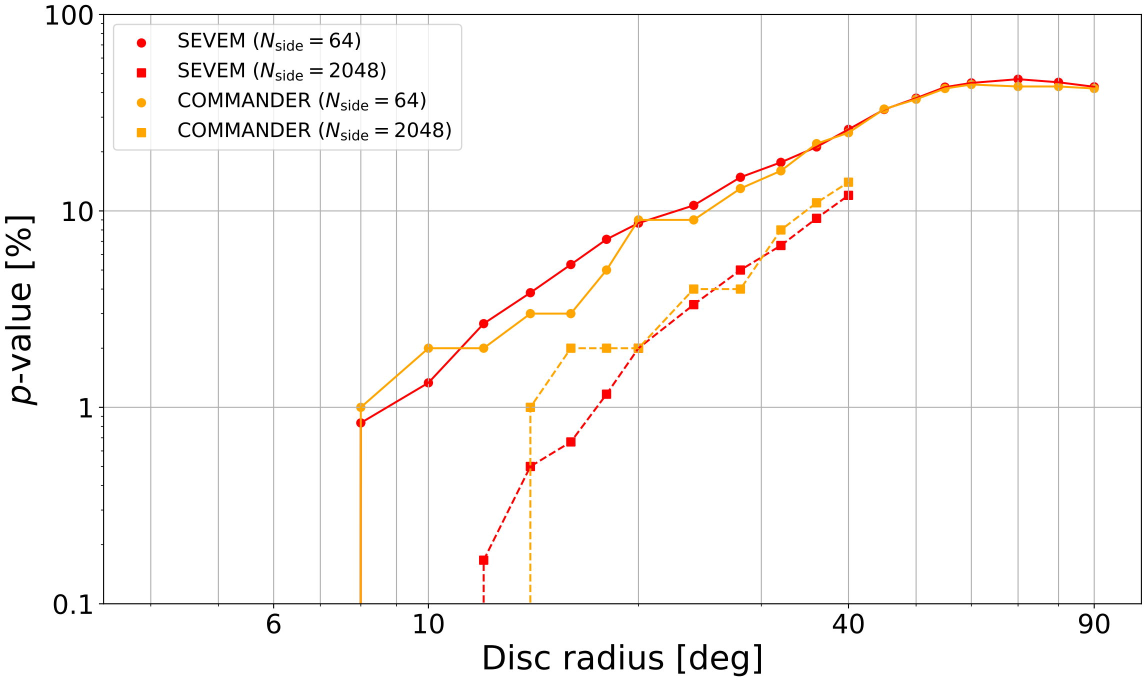

FIGURE 4.10: p-values for the asymmetry measured through the local-variance estimator for PR4 `Sevem` and `Commander` temperature maps and the two considered resolutions, $N_{\rm side}$ = 2048 and $N_{\rm side}$ = 64. The p-value is inferred by checking how many simulations have a local dipole amplitude larger than the one observed in the data. The sensitivity depends directly on the number of available simulations, corresponding to 0.17% for `Sevem` and 1% for `Commander`.

the data. Note that due to the lower number of simulations available for `Commander`, our sensitivity for the p-value is 1% for this method.

Taking into account that the asymmetry only appears at large angular scales, we repeat the analysis for low resolution maps, at $N_{\rm side} = 64$. Figure 4.10 shows the p-value as a function of the disc radius for both component separation methods and both resolutions. For low resolution maps the p-value is computed using discs with radii between 4 and 90 degrees as in Planck Collaboration et al. (2020d). In general, the significance level is lower than in the full-resolution analysis. Nevertheless, for discs of 4° and 6°, we continue to find that no simulation has an amplitude equal to or greater than that of the data.

Regarding the direction of the dipole, in Table 4.1 we provide the full-resolution and low-resolution direction for `Sevem` and `Commander` obtained by considering 4° discs. In `Sevem` the observed direction is in excellent agreement with the previously reported one (see Table 22 from Planck Collaboration et al. (2020d)), while in `Commander` there is a small discrepancy of a few degrees, which could be due to the difference in the number of simulations between releases. In particular, we have checked that the ℓ and b angles have a very large dispersion (see Figure 4.8), especially when the amplitude of the modulation is small, making consistent the values for `Commander` between both releases.

	(ℓ, b) [deg]			
	PR4		PR3	
Data	$N_{\text{side}} = 64$	$N_{\text{side}} = 2048$	$N_{\text{side}} = 64$	$N_{\text{side}} = 2048$
`Sevem`	$(208, -15)$	$(205, -20)$	$(209, -14)$	$(205, -19)$
`Commander`	$(213, -16)$	$(207, -20)$	$(209, -15)$	$(205, -20)$

Table 4.1: Local-variance dipole directions for PR4 `Sevem` and `Commander` temperature maps. Directions are measured in galactic coordinates for both resolutions and using 4° discs.

4.3.2 E-mode polarisation results for PR3

In order to compare the performance of our GCR inpainting technique versus the purified inpainting used in the *Planck* 2018 analysis (Planck Collaboration et al., 2020d), we have first carried out the analysis of PR3 data using our method. We have followed the procedure explained in Section 4.2.2. We recall that, in this case, each component separation algorithm has in total 300 noise E2E simulations each of them accompanied by 3 independent CMB skies. We have done 30 splits of two independent sets of 150 noise simulations and computed the covariance matrices as previously explained in Section 4.2.2. Combining the 150 noise simulations with the 450 independent CMB skies, for the analysis we have available 60 sets of 450 simulations.

Table 4.2 summarizes the results of this chapter. For each component-separation algorithm we provide the preferred direction of the local-variance dipole and the p-values obtained using the reference mask. In addition, in Table 4.3 we give the angular distance between the temperature and E-mode dipoles, and the p-value associated with this alignment. As one would expect, we recover results quite similar to those obtained by the *Planck* collaboration, validating also our approach. We can distinguish two groups. While for `Sevem` and `Commander` the p-value of having such a large amplitude is below 1%, for `SMICA` and `NILC` the probability is at the level of $3-4\%$. Regarding the TE alignment, we are below 6% for all the component separation methods. Again results are overall consistent between both analyses, although the result for `NILC` and `SMICA` becomes less anomalous in our case (p-values of 5.7% and 2.0% versus 1.9% and 0.9%, respectively). These small deviations can be explained by the fact that our analysis presents some differences with respect to the previous one, such as the inpainting technique, a different mask or the use of splits.

In order to test the robustness of the p-value against the mask, we have generated 4 additional masks using thresholds in the maximum reconstruction error (see Section 4.2.2) ranging from 45% to 35% and with the following allowed fractions of sky: 64.5%, 63%, 59.6%, 57.6%. Figure 4.11 presents the p-values obtained for the four component-separated E-mode polarisation maps as a function of the f_{sky}. The minimum p-value is obtained for the mask with $f_{\text{sky}} = 59.6\%$. `Sevem` and `Commander` are still below 1% and the p-value is even lower than with the reference mask ($f_{\text{sky}} = 61.4\%$). However, `SMICA` and `NILC` present still

---

[20] Planck Collaboration et al. (2020d)

Data	p-value [%] (this work)	p-value range [%]	(ℓ, b) [deg]	p-value [%] (from [20])
`Sevem`	0.22	[<1/450, 0.44]	$(232, -9) \pm 4$	0.4
`Commander`	0.70	[<1/450, 1.2]	$(222, -9) \pm 4$	0.7
`SMICA`	4.4	[1.8, 6.4]	$(225, -12) \pm 4$	5.5
`NILC`	3.4	[1.6, 5.0]	$(238, -16) \pm 5$	5.8

Table 4.2: Local-variance dipole directions and p-values for the PR3 four component-separated E-mode polarisation maps, analysed with the inpainting procedure, at N_{side} = 64, together with the 1σ interval obtained from the 60 sets as explained in Section 4.2.2. Measured directions are also showed. The error is estimated from the distribution of the angular distances between the mean direction and the directions of the 60 data sets, taking into account the contour, with azimuthal symmetry, that includes 68% of the directions. All the values have been obtained using 4° discs. Last column shows the p-value obtained in the previous work by the *Planck* collaboration (Planck Collaboration et al., 2020d).

	TE alignment (this work)		TE (from [20])	
Data	$\cos\alpha$	p-value [%]	$\cos\alpha$	p-value [%]
`Sevem`	0.91	4.2	0.86	6.9
`Commander`	0.99	0.7	0.99	0.9
`SMICA`	0.96	2.0	0.99	0.9
`NILC`	0.88	5.7	0.97	1.9

Table 4.3: Angular distance of alignment between the preferred direction of the local-variance dipole observed in intensity and the one for the E-mode polarisation data. We use the mean directions presented in Table 4.2 to compute the angular distance, and the p-value is obtained using the E2E simulations combining all the results from different splits in a single distribution. The previous results by the *Planck* collaboration (Planck Collaboration et al., 2020d) are shown in the last column .

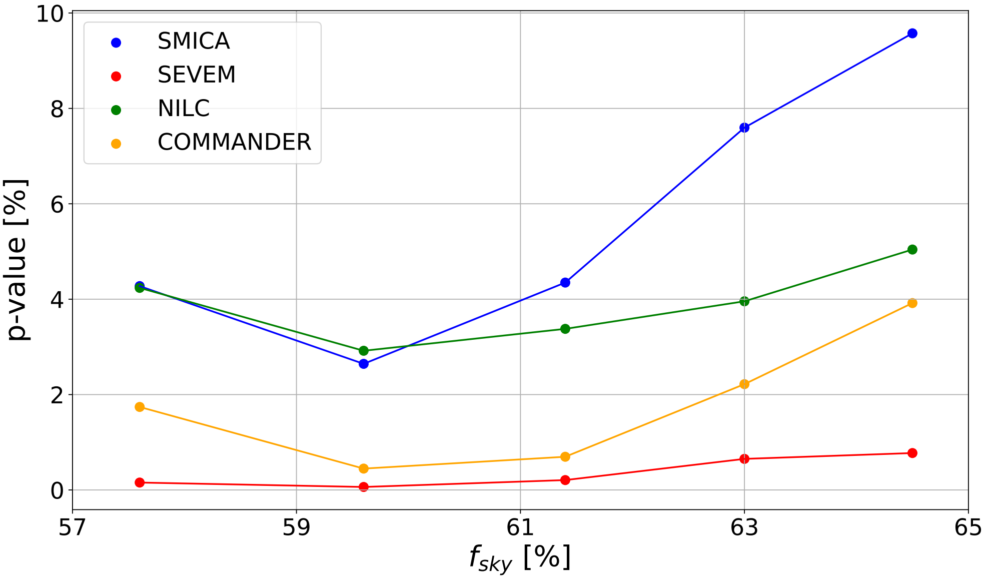

FIGURE 4.11: p-values for variance asymmetry as a function of $f_{\rm sky}$ for the four PR3 component-separated E-mode polarisation maps analysed with the inpainting procedure.: `SMICA` (blue), `Sevem` (red), `NILC` (green) and `Commander` (orange). All values have been obtained at $N_{\rm side} = 64$ and using 4° discs. We recall that our reference mask for PR3 corresponds to a $f_{\rm sky}$=61.4%.

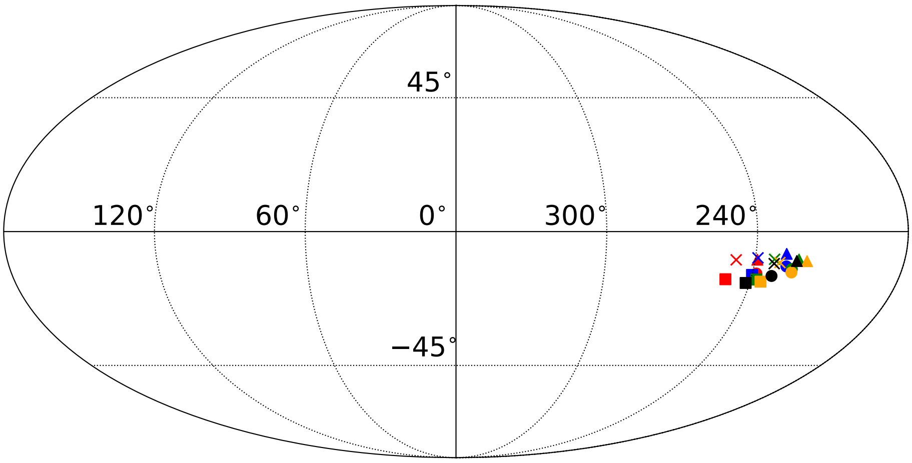

FIGURE 4.12: Local-variance dipole directions for the four PR3 component-separated E-mode polarisation maps analysed with the inpainting procedure.: `SMICA` (circle), `Sevem` (cross), `NILC` (square) and `Commander` (triangle). Different colours corresponds to different $f_{\rm sky}$: red (64.5%), blue (63%), green (61.4%), orange (59.6%), and black (57.6%). All values have been obtained at $N_{\rm side} = 64$ and using 4° discs.

p-value [%]	p-value range [%]	(ℓ, b) [deg]	TE alignment	
			$\cos\alpha$	p-value [%]
2.8	[1.8, 3.8]	$(234, -14)$ ±5	0.91	4.5

Table 4.4: Local-variance dipole direction and p-value for the `Sevem` PR4 E-mode polarisation map, analysed with the inpainting procedure, at $N_{\rm side} = 64$, together with the 1σ interval obtained from the 60 data sets. Measured directions are also showed. The error is estimated from the distribution of the angular distance between the mean direction and the directions of the 60 data sets, taking into account the contour, with azimuthal symmetry, that includes 68% of the directions. All the values have been obtained using $4°$ discs.

values around or above a few per cent for all the considered masks. Additionally, Figure 4.12 displays the coordinates of the dipole direction as a function of the $f_{\rm sky}$. All the methods and masks show certain level of robustness in the sense that all the directions are concentrated in a small region near (ℓ, b) = (230°, $-13°$). Furthermore, it seems that the latitude (b) is somehow more robust than the longitude (ℓ).

4.3.3 E-mode polarisation results for PR4

Finally, we present the results for the `Sevem` PR4 data set using the GCR inpainting [21]. We have followed the same procedure as for PR3 with a total of 30 splits of 300/300 simulations.

Table 4.4 summarizes the results, providing the preferred direction of the local-variance dipole and p-values obtained for the `Sevem` algorithm using the PR4 reference mask. The error in the direction is obtained as the 68 per cent (single-tailed) of the distribution of the angular distance between the mean direction and the direction of the 60 data sets. Therefore, it is only given to provide an estimation of the dispersion in the different considered splits. In practice the error associated to the estimated direction should be significantly larger, since this quantity is below the errors inferred from the methodology itself or the effect of the inpainting. We also give in the same table the angular distance between the temperature and E-mode dipole directions, and the p-value associated with this alignment. The analysis shows that the p-value of the amplitude for `Sevem` PR4 increases with respect to the one obtained for PR3 (see Table 4.2), increasing from 0.22 to 2.8 per cent. The p-value of the TE alignment is very similar for both pipelines and remains below 5 per cent.

Another interesting result is the level of the modulation in the data. Using the modulated simulations described in Section 4.2.4, the relation between the amplitude of the modulation (given by equation 3.16) and the one measured in the local-variance map can be estimated. Figure 4.13 shows this relation together with the amplitude observed in the data local-variance map for the reference mask. According to this, for modulations with am-

---

[21] Note that we have not used the `Commander` PR4 data since CMB and noise simulations are not provided separately (i.e. only CMB+noise simulations are provided), which makes not possible to compute independently the noise covariance matrix. Although, in principle, it would be possible to estimate the total covariance matrix, this introduces additional uncertainties in the matrix elements. Moreover, only 400 simulations are available for polarisation, increasing even further the error in the analysis.

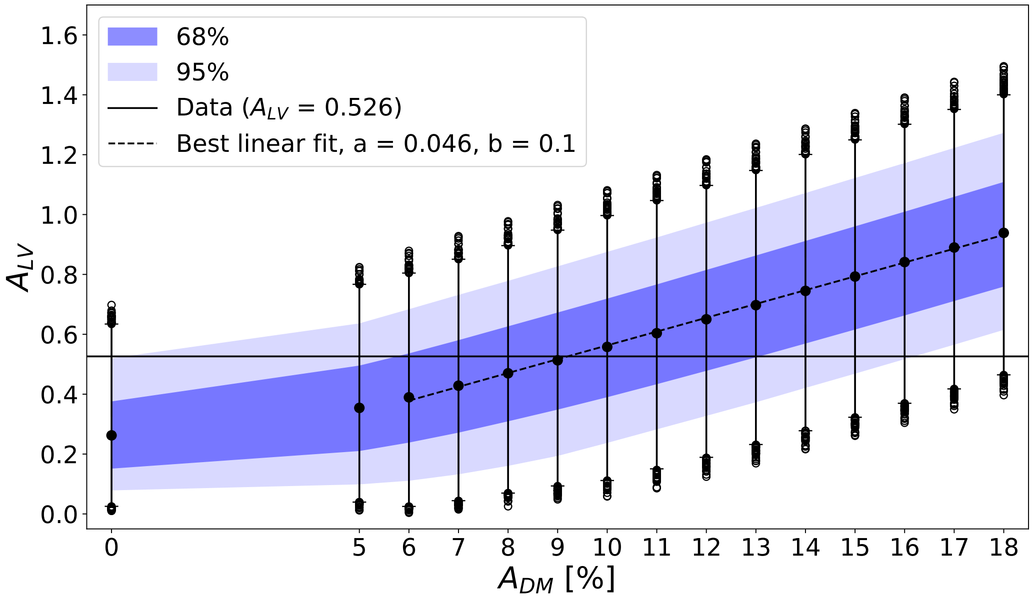

FIGURE 4.13: Relation between the dipolar modulation amplitude $A_{\rm DM}$ (input value in the modulated simulations) and the one measured in the local-variance map $A_{\rm LV}$, assuming the realistic PR4 noise and systematics. For each amplitude, all the 60 inpainted data sets are used to generate a single distribution. For amplitudes larger than 6%, there is a linear relation between both amplitudes with a slope of 0.046 (dashed black line). The horizontal black line corresponds to the mean amplitude (over the 60 data sets) measured in the data local-variance map using the PR4 reference mask. The measured $A_{\rm LV}$ value corresponds to a modulation at the level of 9%, and it is within the 68% of the distribution for the $A_{\rm DM}$ interval between 6% and 13%. Black dots in the tails of distributions represent the values outside the 99.7%.

plitudes larger than 6%, the relation is linear, although with a large dispersion due to the E-mode noise level. Assuming that the modulation model (Eq.3.16) is the correct one, the asymmetry observed in the data corresponds to a modulation at the level of 9%. Moreover, the estimated amplitude $A_{\rm LV}$ is within the 68% of the distribution for the dipolar modulation amplitude ($A_{\rm DM}$) for the range between 6% and 13%.

As before, we test the robustness of the p-value against the different masks. We have generated 6 additional masks using thresholds in the range from 45% to 30% with the following allowed fractions of sky: 67%, 65.9%, 63.2%, 61.5%, 59.4%, and 57% (we recall that the reference mask corresponds to $f_{\rm sky} = 64.6\%$). Imposing more strict conditions in the error of the recovered E-mode map (i.e. lower thresholds) while keeping a reasonable $f_{\rm sky}$ is possible in this case because a more accurate E-mode reconstruction can be obtained in PR4. This is due to both, a lower level of systematics and twice the number of simulations that have been used for estimating the noise covariance matrix. Figure 4.14 presents the obtained p-values as a function of $f_{\rm sky}$. The minimum p-value (1%) is obtained for the mask with $f_{\rm sky} = 61.5\%$. For the 5 smallest masks, allowing a sky between 67% and 61.5%, the p-value is quite stable with a value below 2.8%, and then it starts to rise with $f_{\rm sky}$. The largest mask returns a p-value of 7.4%. This behaviour could be due to the loss of informa-

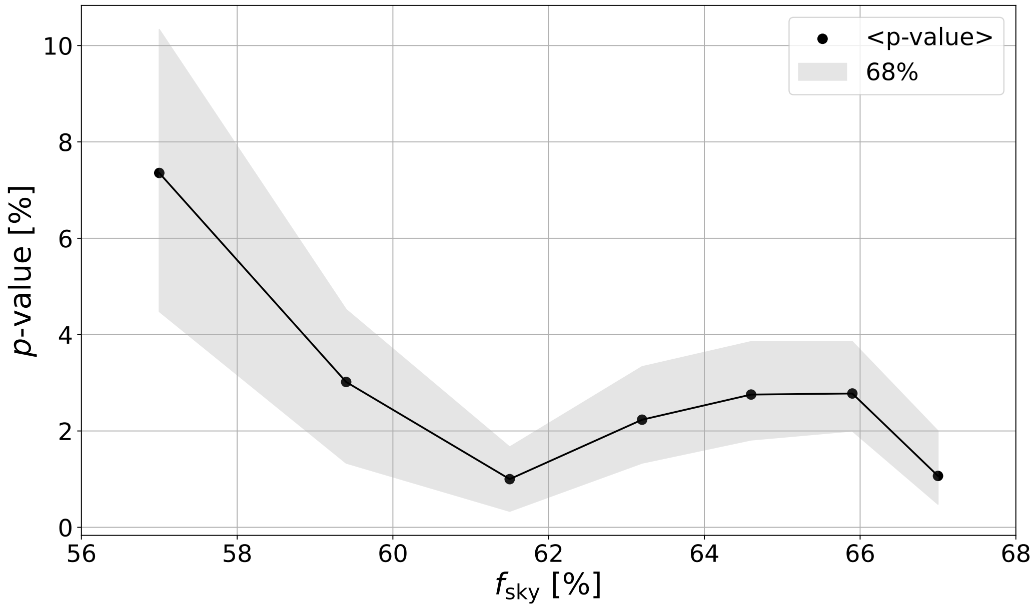

FIGURE 4.14: p-values for variance asymmetry as a function of $f_{\rm sky}$ for the PR4 `Sevem` E-mode polarisation map analysed with the inpainting procedure. The grey region corresponds to the 68% interval obtained from the 60 data sets. All values have been obtained at $N_{\rm side} = 64$ and using 4° discs.

tion as the mask increases, assuming the asymmetry had a cosmological origin. However, it could also be due to the fact that a larger mask reduces foreground contamination, what we would expect if the asymmetry were due to the presence of residuals. Additionally, Figure 4.15 and 4.16 display the directions of the local-variance dipoles and the p-value of the TE alignment, respectively, as a function of the $f_{\rm sky}$. In particular, the minimum p-value (4.6%) is obtained for the reference mask. Finally, Figure 4.17 summarizes in a Mollweide projection all the directions measured in this chapter.

We have performed a final test where we studied if the difference observed in the p-value from the PR3 and PR4 processing is consistent with that expected due to differences in the number of simulations, in the mask, inpainting, TF and noise properties. For this purpose, we construct two data sets that have the same input CMB sky but different noise properties. In particular, the first set is formed by the modulated 600 PR4 CMB simulations, plus the corresponding CMB noise (including thus also the TF). The second set consists of the same modulated PR4 simulations but adding the PR3 noise (repeating two times each simulation, because only 300 are available). Then, we run the full procedure, using the corresponding reference mask for each data set, and we compare the output p-values. Note that for the PR3 case, we have used as reference the PR3 CMB plus noise simulations, i.e. 450 simulations (450 CMB simulations and the 150 noise simulations, repeated three times, that are not considered in the matrix estimation). The distribution we get from the 600 simulations for the difference between the p-value for PR4 minus the one for PR3 is peaked

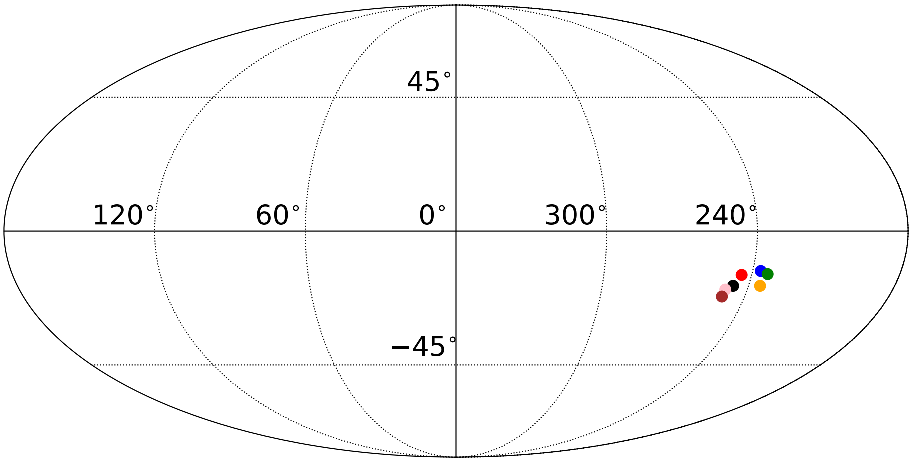

FIGURE 4.15: Local-variance dipole directions for the `Sevem` PR4 component-separated E-mode polarisation maps analysed with the inpainting procedure. Different colours corresponds to different $\mathrm{f_{sky}}$: red (67%), blue (65.9%), green (64.6%), orange (63.2%), black (61.5%), pink (59.4%), and brown (57%). All values have been obtained at Nside = 64 and using 4° discs

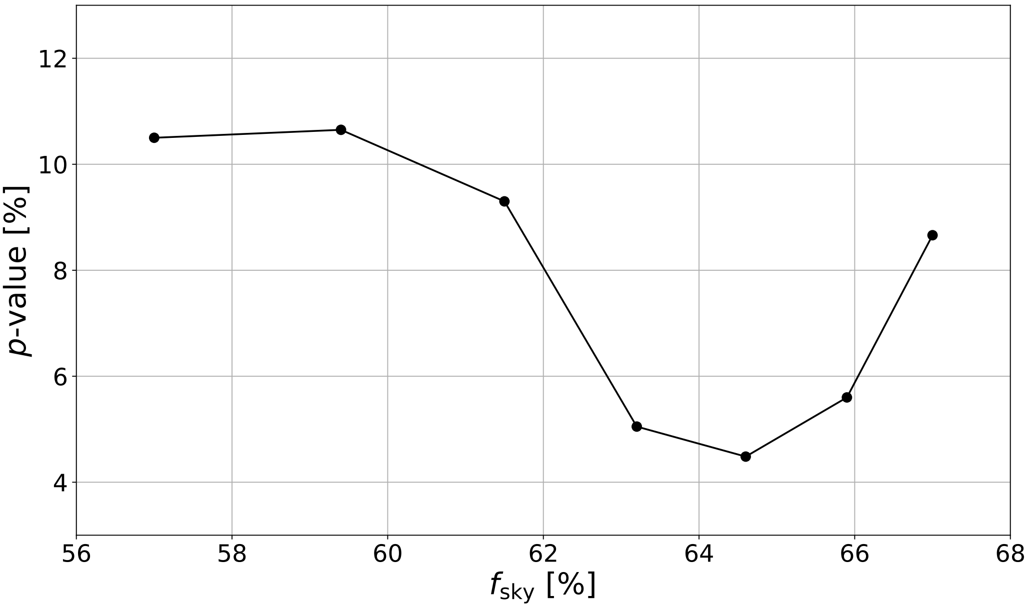

FIGURE 4.16: p-value of the alignment between the dipole axis in PR4 temperature and E-mode as a function of f_{sky}.

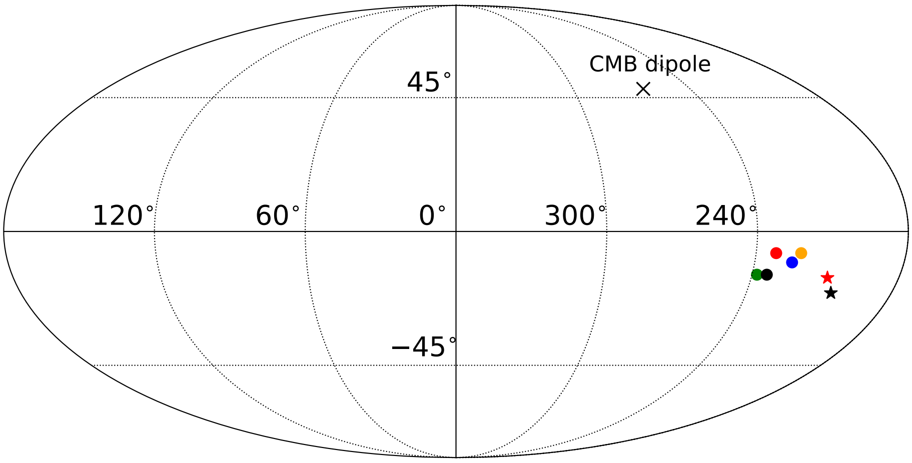

FIGURE 4.17: Local-variance dipole directions for the PR3 four component-separated E-mode polarisation maps analysed with the inpainting procedure: `Sevem` (red), `SMICA` (blue), `NILC` (green), and `Commander` (orange). The black dot corresponds to the direction measured using the `Sevem` PR4 polarisation data. For reference, we also show the CMB dipole direction, and the directions observed in the temperature data for `Sevem` PR4 at $N_{\rm side} = 64$ (red star) and $N_{\rm side} = 2048$ (black star).

at -8.9%, which means that the dipolar asymmetry can be better constrained using the PR4 processing even in the presence of the TF. The difference observed in the data (from 0.22% to 2.8%) is well within the 68% CL (which corresponds to the range -34.9% to 8.0%), and thus both values are fully consistent taking into account the differences in the data and in the analysis. Finally, we also see large tails in the distribution, which is an indication of the low signal-to-noise ratio as the amplitudes and p-values are significantly affected by noise and systematics.

5

CMB-PAInT: An inpainting tool for the CMB

Astrophysical emissions in the microwave range force us to apply component separation algorithms to recover the CMB signal. However, even in the most optimistic cases, certain regions remain strongly contaminated and must be excluded using a proper mask. Since many CMB analyses, especially the ones working in harmonic space, need the whole sky map, it is crucial to develop a reliable inpainting algorithm that replaces the values of the excluded pixels by others statistically compatible with the rest of the sky. This is especially important when working with Q and U sky maps in order to obtain E- and B-mode maps which are free from E-to-B leakage. In this work we study a method based on Gaussian Constrained Realisations (GCR), that can deal with both intensity and polarisation. Several tests have been performed to assess the validation of the method, including the study of the one-dimensional probability distribution function (1-PDF), E- and B-mode map reconstruction, and power spectra estimation. We have considered two scenarios for the input simulation: one case with only CMB signal and a second one including also *Planck* PR4 semi-realistic noise. Even if we are limited to low resolution maps, $N_{\rm side} = 64$ if T, Q and U are considered, we believe that this is a useful approach to be applied to future missions such as LiteBIRD, where the target are the largest scales.

5.1 Gaussian Constrained Realization

The inpainting technique presented in this work is a pixel domain approach based on a Gaussian Constrained Realization (GCR) (Bucher & Louis, 2012; Kim et al., 2012). The followed methodology was already described in Marcos-Caballero & Martínez-González (2019) for the temperature field, but in this work we extend it to a spin-2 field, i.e., to CMB polarisation. The idea is to fill the masked pixels by sampling from the conditional probability distribution, $p(\hat{\mathbf{d}}|\mathbf{d})$, where $\hat{d}$ is the vector of the inpainted field and d is the vector of the available pixels. The method requires a single assumption, Gaussianity of the field, which is a good approximation for the CMB data. Under this condition, only the pixel covariance matrix is needed, which can be computed given a theoretical power spectrum[1], C_l, following the next equations (see Appendix A of Tegmark & de Oliveira-Costa (2001) for the full set of equations):

$$C_{ij}^{TT} = \langle T_i T_j \rangle = \sum_{\ell} \left(\frac{2\ell+1}{4\pi} \right) C_{\ell}^{TT} P_{\ell}(z), \tag{5.1}$$

$$C_{ij}^{QQ} = \langle Q_i Q_j \rangle = \sum_{\ell} \left(\frac{2\ell+1}{4\pi} \right) [F_{\ell}^{12}(z) C_{\ell}^{EE} - F_{\ell}^{22}(z) C_{\ell}^{BB}], \tag{5.2}$$

$$C_{ij}^{UU} = \langle U_i U_j \rangle = \sum_{\ell} \left(\frac{2\ell+1}{4\pi} \right) [F_{\ell}^{12}(z) C_{\ell}^{BB} - F_{\ell}^{22}(z) C_{\ell}^{EE}], \tag{5.3}$$

$$C_{ij}^{TQ} = \langle T_i Q_j \rangle = -\sum_{\ell} \left(\frac{2\ell+1}{4\pi} \right) F_{\ell}^{10}(z) C_{\ell}^{TE}, \tag{5.4}$$

where $z = \hat{r}_i \cdot \hat{r}_j$ gives the cosine of the angle between the two pixels. The F functions are

$$F^{10}(z) = 2 \frac{\frac{\ell z}{(1-\ell^2)} P_{\ell-1}(z) - \left(\frac{\ell}{1-z^2} + \frac{\ell(\ell-1)}{2} \right) P_{\ell}(z)}{[(\ell-1)\ell(\ell+1)(\ell+2)]^{1/2}}, \tag{5.5}$$

$$F^{12}(z) = 2 \frac{\frac{(\ell+2)z}{(1-\ell^2)} P_{\ell-1}^2(z) - \left(\frac{\ell-4}{1-z^2} + \frac{\ell(\ell-1)}{2} \right) P_{\ell}^2(z)}{(\ell-1)\ell(\ell+1)(\ell+2)}, \tag{5.6}$$

$$F^{22}(z) = 4 \frac{(\ell+2) P_{\ell-1}^2(z) - (\ell-1) z P_{\ell}^2(z)}{(\ell-1)\ell(\ell+1)(\ell+2)(1-z^2)}, \tag{5.7}$$

where P_ℓ are the Legendre polynomials.

The covariance matrix will be arranged by blocks as follows:

$$M(\hat{r}_i \cdot \hat{r}_j) = \begin{pmatrix} \langle T_i T_j \rangle & \langle T_i Q_j \rangle & \langle T_i U_j \rangle \\ \langle T_i Q_j \rangle & \langle Q_i Q_j \rangle & \langle Q_i U_j \rangle \\ \langle T_i U_j \rangle & \langle Q_i U_j \rangle & \langle U_i U_j \rangle \end{pmatrix} \tag{5.8}$$

Note that the elements of this covariance matrix are referred to a specific coordinate system where the reference direction points along the great circle connecting the two points. However, what is needed is the covariance matrix referred to a global coordinate frame where the reference directions are meridians, so the following rotation matrices are applied (see Appendix A in Tegmark & de Oliveira-Costa (2001) to find how to compute the rotation

---

[1] In the present work we have used as the input power spectra the best fit to the ΛCDM mode provided by *Planck* (Aghanim et al., 2020). It can be downloaded from the *Planck* Legacy Archive: `pla.esac.esa.int`

angle[2] α):

$$\langle x_i x_j^t \rangle = R(\alpha_{ij}) M(\hat{r}_i \cdot \hat{r}_j) R(\alpha_{ij})^t, \tag{5.9}$$

$$\mathbf{R}(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2\alpha & \sin 2\alpha \\ 0 & -\sin 2\alpha & \cos 2\alpha \end{pmatrix}, \tag{5.10}$$

where $x_i = \{T, Q, U\}$.

The effect of the experimental beam and the pixel window function can be added by smoothing the power spectra in equations (5.1)-(5.4).

The maximum multipole considered in the sum of equations (5.1)-(5.4) should be equal to the largest multipole accounted for in the map. As default $3N_{\text{side}}$ is considered. This approach is only feasible for low resolution maps, up to $N_{\text{side}} = 64$, as the dimension of the total matrix is $3N_{\text{pix}} \times 3N_{\text{pix}}$. Adding a small regularizing noise to the diagonal is needed to avoid singularities and to ensure that the matrix is positive definite. The noise level depends on the resolution. For $N_{\text{side}} = 64$ a noise amplitude of 0.00001% is used. A similar regularization can be achieved by considering in the sum a maximum multipole of $4N_{\text{side}}$ or even larger. For a detailed discussion on the regularity of a CMB covariance matrix see Bilbao-Ahedo et al. (2017).

Once the matrix is computed and rotated, we reorder the columns and rows in a way that all the unmasked pixels are in the first entries and the masked ones in the last entries[3]. Then, the Cholesky decomposition allows one to sample from the desired distribution by solving the following system

$$\begin{pmatrix} \mathbf{d} \\ \hat{\mathbf{d}} \end{pmatrix} = \begin{pmatrix} \mathbf{L} & 0 \\ \mathbf{R} & \hat{\mathbf{L}} \end{pmatrix} \begin{pmatrix} \mathbf{z} \\ \hat{\mathbf{z}} \end{pmatrix}. \tag{5.11}$$

The matrix in the right-hand side is the Cholesky decomposition, where L and $\hat{\text{L}}$ are low triangular matrices and R a rectangular matrix. The number of rows and columns of the L matrix is equal to the number of unmasked pixels, while for $\hat{\text{L}}$ is the number of masked pixels.

Looking at equation 5.11 and taking into account that **L** is a lower triangular matrix, it becomes apparent that a matrix inversion is not needed and the vector **z** can be computed in a recursive way,

$$z_n = \frac{d_n - \sum_{k=1}^{n-1} L_{nk} z_k}{L_{nn}} \tag{5.12}$$

---

[2] For those who want to use the equations, there is a small typo in equation A7 of the appendix since the proportionality constant is negative, not positive. Thus in equation A8, the negative sign corresponds to the case where the z component of the vector $\hat{r}_{ij}$, the vector of the great circle connecting the two pixels, is positive, and vice versa.

[3] In the case of a full TQU covariance matrix, the order considered is: unmasked T pixels, unmasked Q pixels, unmasked U pixels, masked T pixels, masked Q pixels and masked U pixels.

If the model is coherent with the observed data this vector should be a Gaussian random vector with zero mean and unit variance. Then, a new random vector $\hat{\mathbf{z}}$, also following a $\mathcal{N}(0, 1)$ distribution, is generated, and the field $\hat{\mathbf{d}}$ is sampled, which has two contributions: the constrained part and the unconstrained or stochastic part.

$$\hat{\mathbf{d}} = \mathbf{R}\mathbf{z} + \hat{\mathbf{L}}\hat{\mathbf{z}} \tag{5.13}$$

In this procedure, inpainting is performed simultaneously on the TQU maps. Of course, if we are only interested in the temperature map, we can just compute the TT block, which will significantly reduce the computational cost and similarly if one is only interested in polarisation. However, if the TT, TQ, and TU blocks are not included, the inpainted pixels will not have the correlation between temperature and polarisation (with comes from TE correlation in the standard cosmology, or from TE, TB and EB in beyond ΛCDM models). In that scenario the size of the matrix will be reduced to $4/9$ if we just take into account Q and U and to $1/9$ if we just want T. Even if the correlation between TE is low, the Q and U pixels located outside the polarisation mask will further restrict the potential values of the temperature pixels within the mask, and vice versa for the polarisation pixels inside the mask.

In principle, the covariance matrix should include all the components present in the map to be inpainted. Dealing with the *Planck* polarisation data, the dominant component is the noise and systematics, and their contribution must be considered. If not, mismatches between the pixels outside the mask and the covariance matrix introduces artifacts in the map. As there is not any theoretical model for the noise and systematics, the only way to estimate the covariance matrix is from end-to-end (E2E) simulations. This can be a limitation of the method: the number of simulations needed for a good characterisation of the realistic anisotropic and correlated noise and systematics, is at least of a few thousands (see Appendix 8.4), while the available realistic realisations are usually limited to several hundreds.

5.2 Signal and noise simulation

In order to validate our method, we generate a single Gaussian isotropic CMB map at a resolution of $N_{\rm side} = 64$. We use the healpy[4] function `synfast`. We smooth the power spectra with a Gaussian beam of $160'$, and we consider a maximum multipole $\ell_{max} = 3N_{\rm side}$. We generate 1200 inpainting realisations from a single sky realization based on the PR3 ΛCDM best fit model. The same spectra is used to compute the pixel covariance matrix. *Planck* 2018 temperature and polarisation confidence masks define the region to be filled. At $N_{\rm nside} = 64$ they respectively leave 71.3% and 72.4% of the sky available.

---

[4] https://github.com/healpy/healpy

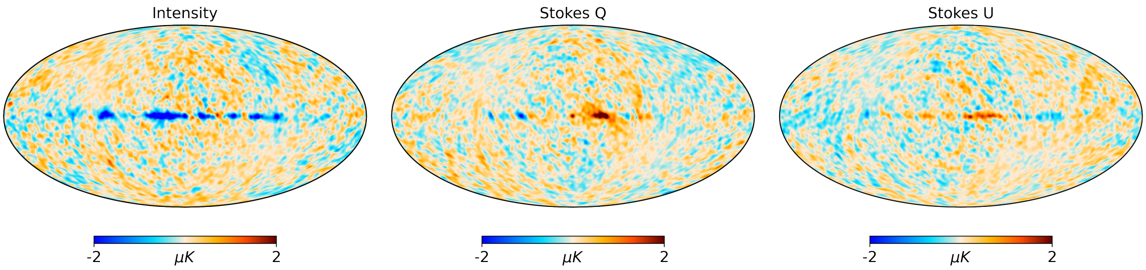

FIGURE 5.1: Input TQU semi-realistic noise simulation at $N_{\rm side}$ = 64 and convolved with a Gaussian beam of FWHM=$160'$.

In order to validate the method in the presence of correlated and anisotropic noise, we generate a semi-realistic noise simulation using characteristics from *Planck*, an ESA satellite that observed the CMB over the full-sky with an unprecedented sensitivity and frequency coverage (Planck Collaboration et al., 2011). We start by computing the covariance matrix from the 600 end-to-end (E2E) *Planck* Release 4 (PR4) (Planck Collaboration et al., 2020f) noise simulations, that include also the expected systematics, except foreground residuals, propagated through the `Sevem` component separation pipeline (Fernández-Cobos et al., 2012). Then, we generate a Gaussian random realization with the proper correlations given by the E2E simulations[5]. Following this pixel-based approach we are able to simulate not only the correlations but also the anisotropy. However, it does not include the presence of possible non-Gaussianity generated by systematic effects.

Figure 5.1 shows the input semi-realistic noise realization. The polarisation noise level used in this paper is very large compared to the noise expected for the future experiments, but it mimics the level of instrumental noise and systematics present in the *Planck* foreground-cleaned CMB `Sevem` map for PR4.

5.3 Tests and Validation

In this section, a series of tests are conducted on a set of inpainted realisations, derived from a single input sky, to assess the algorithm's performance. *Planck* 2018 temperature and polarisation confidence masks (Planck Collaboration et al., 2020b) are considered to inpaint the input maps. First, we verify that the mean and variance maps are compatible with the expected values. Then, some statistics are examined, such as the 1-point probability density function (1-PDF) in real space or the power spectrum in harmonic space confirming their consistency with the input values. Furthermore, for each inpainted Q and U maps, E- and B-mode maps are generated. By subtracting them from the corresponding input maps, we compute the mean and standard deviation of the residuals pixel by pixel. This provides

---

[5] Actually, this is also done using a Cholesky decomposition. In this case, we obtain the realization as $\mathbf{d}' = \mathbf{L}'\mathbf{z}'$, where $\mathbf{L}'$ is the Cholesky decomposition of the covariance matrix and $\mathbf{z}'$ a Gaussian random vector with zero mean and unity dispersion.

information on how the residuals are distributed and on the level of the errors introduced by the inpainting. In particular, this map can be used as a suitable reference to generate customized E- and B-mode masks for pixel based estimators.

All of these tests are applied in two different scenarios: the noiseless case where the input sky is just CMB, or the case of CMB plus a semi-realistic noise simulation (constructed as explained in Section 5.2). The second scenario can also be divided into some subcases depending on how well the pixel-pixel covariance matrix is characterised. We first study the ideal case where the correlations between pixels are well known and covariances are perfectly characterised. This is possible because the semi-realistic noise simulations are generated from an input covariance matrix. Thus, the performance of the inpainting in the presence of correlated and anisotropic noise in an ideal case can be studied. However, as previously mentioned, in a real experiment we do not expect to have a perfect knowledge of the complex properties of instrumental noise and systematics. The covariance matrix will need to be estimated from high-cost CPU simulations, limiting the number that can be produced (typically only several hundreds) and, therefore, our capacity to characterise it properly. In Appendix 8.4 we study the impact of the matrix convergence varying the number of simulations used to estimate the covariance matrix. In particular, this can be used to establish a rough estimation of the number of E2E simulations needed from future experiments. To study the convergence of the matrix the differences between the input and estimated matrices can be checked, but other variables can also be studied such as the intermediate z variables (see eq. 5.12). In presence of a mismatch between the covariance matrix and that of the unmasked pixels, an error is produced which propagates in the calculation of the z variables. Thus, these quantities are not longer $\mathcal{N}(0,1)$ variables as the dispersion increases, and it can be used as a tracer of the convergence. Additionally, the constrained part of the inpainting (first term in the right-hand side of Eq. 5.12) is also affected, and this introduces artifacts within the inpainted region. In any case, the number of simulations needed for a good convergence inferred from the study of Appendix 8.4 should be taken as a tentative number. For the realistic E2E simulations there are other effects that can contribute to the mismatch such as the non-Gaussianities, which are not simulated here.

Nevertheless, in the most ideal case we would be interested in no recovering the prominent noise and systematics in the inpainted region, but holding the statistical compatibility with the unmasked pixels. However, taking into account that not including them in the methodology generates some artifacts, there is not straightforward way to proceed. Fortunately, there is a situation where this can be avoided. If the noise is negligible compared to the signal at the map level, which is the case for the *Planck* temperature maps, its contribution can be also neglected from the matrix without having to pay the penalty of a significant mismatch. For the most general case there are other alternatives that we will leave for further studies. For example, noise and systematics can be isotropised for all the matrix elements that involve masked pixels, or directly their contribution can be removed and just take into account the noise and systematics for the unmasked pixels, where they are expected to be subdominant. The last option could lead to matrix singularity problems, which will need the inclusion of regularization noise to be solved. In any case, the optimal solution will depend on the nature of the data and the estimator to be used. For instance, in Gimeno-Amo et al. (2023) we show that performing inpainting using a noise covariance matrix estimated from a set of 300 simulations was sufficient to improve significantly the performance of our estimator with respect to a simple masking approach in Q and U.

5.3.1 Constrained contribution

In this section, we study the mean map within the inpainted region by averaging over the 1200 realisations, and we compare it with the theoretical prediction. We also show an example of inpainted maps, including the constrained and unconstrained contributions. We consider both scenarios, only CMB and adding semi-realistic noise.

From eq. 5.13 it is straightforward to obtain that the mean and covariance of the inpainted field are

$$\langle \hat{\mathbf{d}} \rangle = \mathbf{R}\mathbf{z} = \mathbf{R}(\mathbf{L}^{-1}\mathbf{d}), \tag{5.14}$$

$$\hat{\mathbf{C}} = \langle \hat{\mathbf{d}}\hat{\mathbf{d}}^{\mathbf{t}} \rangle - \langle \hat{\mathbf{d}} \rangle \langle \hat{\mathbf{d}}^{\mathbf{t}} \rangle = \hat{\mathbf{L}}\hat{\mathbf{L}}^{\mathbf{t}}. \tag{5.15}$$

The mean map is given by the constrained part and it is the dominant contribution to the inpainting in the regions close to the boundaries of the mask, where the constraints are tighter. The variance in these regions is close to zero, as these pixels are almost fully constrained and their value do not vary significantly from one inpainting realization to another.

In order to check if the mean field of the 1200 inpainted realisations ($N_{\rm sims}$) is consistent with the theoretical prediction, we define the quantity

$$\epsilon = \frac{\langle \hat{\mathbf{d}} \rangle_{\rm obs} - \langle \hat{\mathbf{d}} \rangle_{\rm th}}{\sigma\left(\langle \hat{\mathbf{d}} \rangle_{\rm obs}\right)/\sqrt{N_{\rm sims}}}, \tag{5.16}$$

which measures pixel by pixel if the observed difference is compatible with zero given the expected error. If everything is consistent, ϵ should follow a Gaussian distribution with zero mean and unit variance, $\mathcal{N}(0,1)$. This is actually seen in Figure 5.2, where we show the results for the T, Q and U components for the noiseless scenario. Similar results are obtained when the case including semi-realistic noise is considered, showing that the method also works when a well characterised anisotropic and correlated noise is added to the input CMB sky.

Figure 5.3 shows an example of one inpainted realization for the noiseless scenario: the input (top panels), inpainted (middle) and difference (bottom) maps are shown for the T (left column), Q (middle) and U (right) components. An example of the constrained and

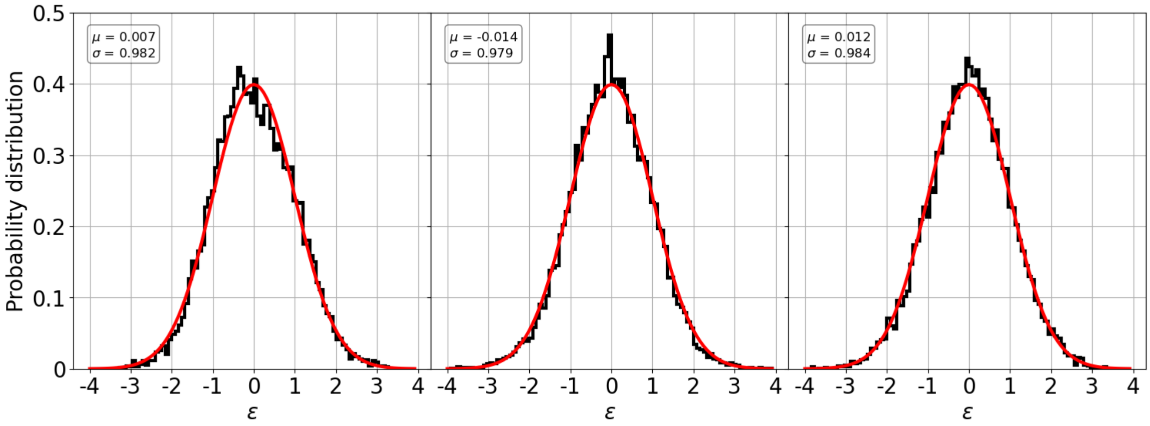

FIGURE 5.2: Distribution of ϵ variable for inpainted T, Q, and U maps. For comparison, the $\mathcal{N}(0,1)$ distribution is also given (red line).

unconstrained components is showed in Figure 5.4 for the Q component, while Figure 5.5 provides the dispersion maps for the three components. There is a clear gradient pointing from the central regions of the mask towards the boundaries where the variance tends to zero. As mentioned before, this is because the pixels in the boundary regions are strongly constrained.

Additionally, an example for the scenario with noise is given in Figure 5.6. As seen, some bright anisotropic features of the Galactic plane are actually reproduced in the inpainted maps. This is due to the fact that a perfectly characterised anisotropic covariance matrix has much more information than an isotropic one, and therefore, the sampled values are more constrained.

5.3.2 1-point probability distribution function

In this section, we study the 1-point probability distribution function (1-PDF) of the inpainted pixels. We also compute the 1-PDF for the E- and B-modes. Figures 5.7 and 5.8 provide the PDFs for the T, Q, and U components and the E- and B-modes, respectively, in the noiseless scenario. Similar results are obtained for the semi-realistic noise case, as shown in Figure 5.9.

Inpainting performs well in both scenarios from the point of view of the reconstructed 1-PDF. The difference in each bin between the input value and the average over the 1200 inpainting realisations is within 2σ for almost every point. For the semi-realistic noise scenario, the tails in the Q and U 1-PDF are larger due to the presence of noise and systematics.

5.3.3 E- & B-mode reconstruction

The quality of the recovery of the E- and B-mode maps is another crucial test. The transformation from Q and U Stokes parameters to more suitable variables E- and B-modes is not

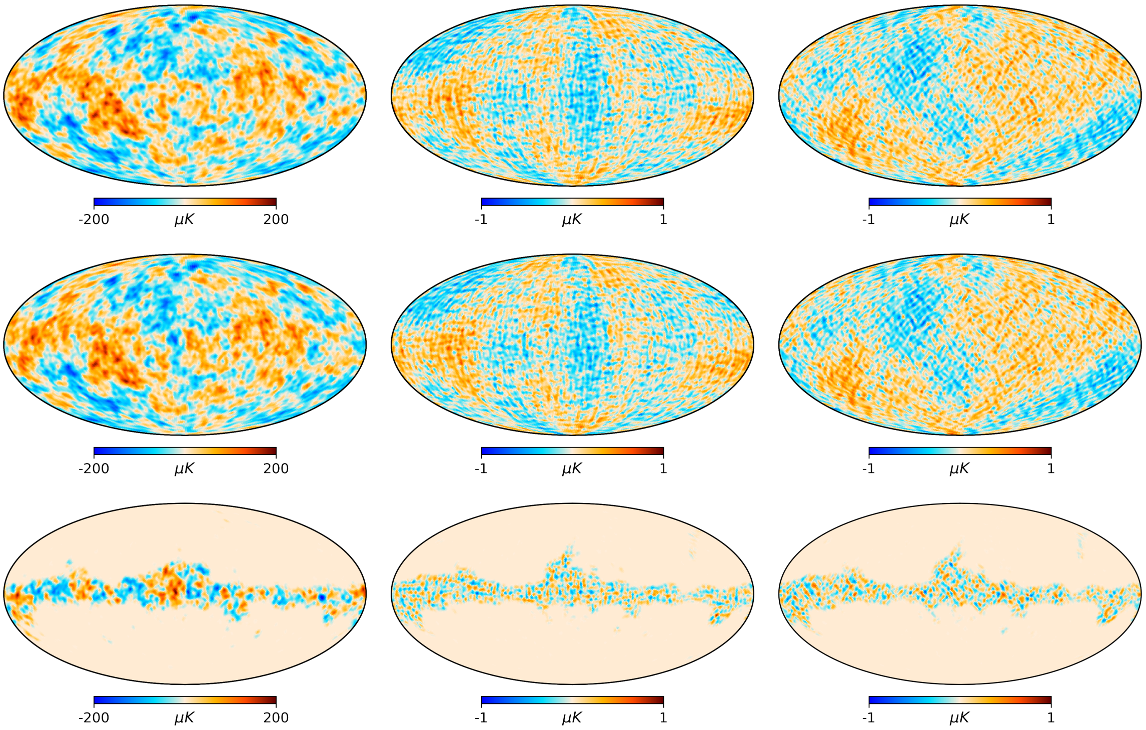

FIGURE 5.3: Example of an inpainted realization in the case where only the CMB signal is considered. First row corresponds to input T (first column), Q (second), and U (third) maps. An inpainted realization is shown in the second row, while the third one shows the difference. All the maps are at $N_{\rm side}$ = 64 and have a resolution of $160'$.

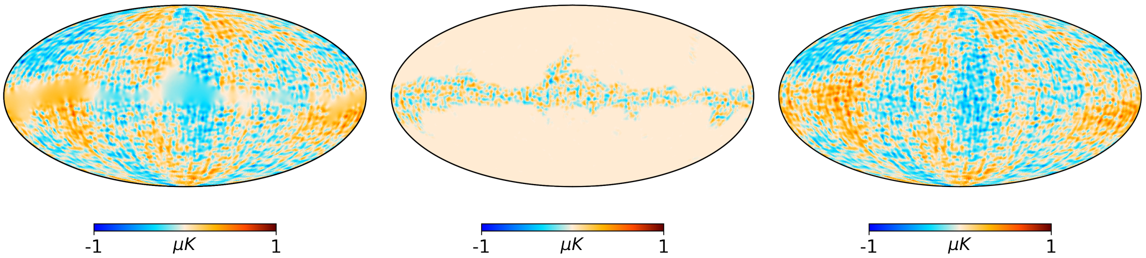

FIGURE 5.4: Example of an inpainted realization for the Q Stokes parameter in the case where only the CMB signal is considered. Left and middle panels show the constrained (deterministic) and unconstrained (stochastic) parts, while the right panel is the sum of both.

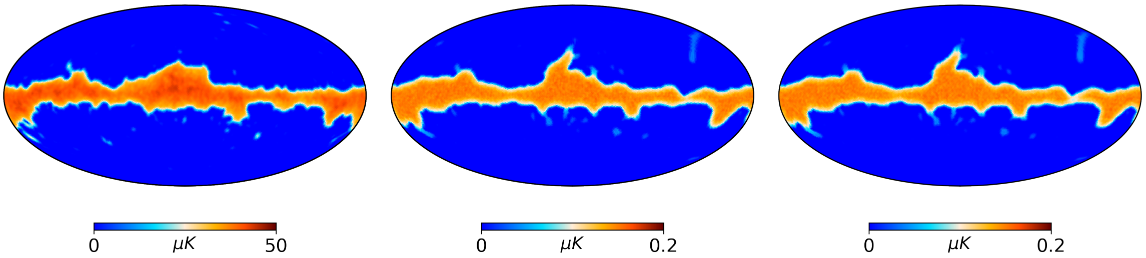

FIGURE 5.5: Dispersion of the inpainted maps estimated from the full set of 1200 inpainted realisations for the T (left panel), Q (middle), and U (right) components.

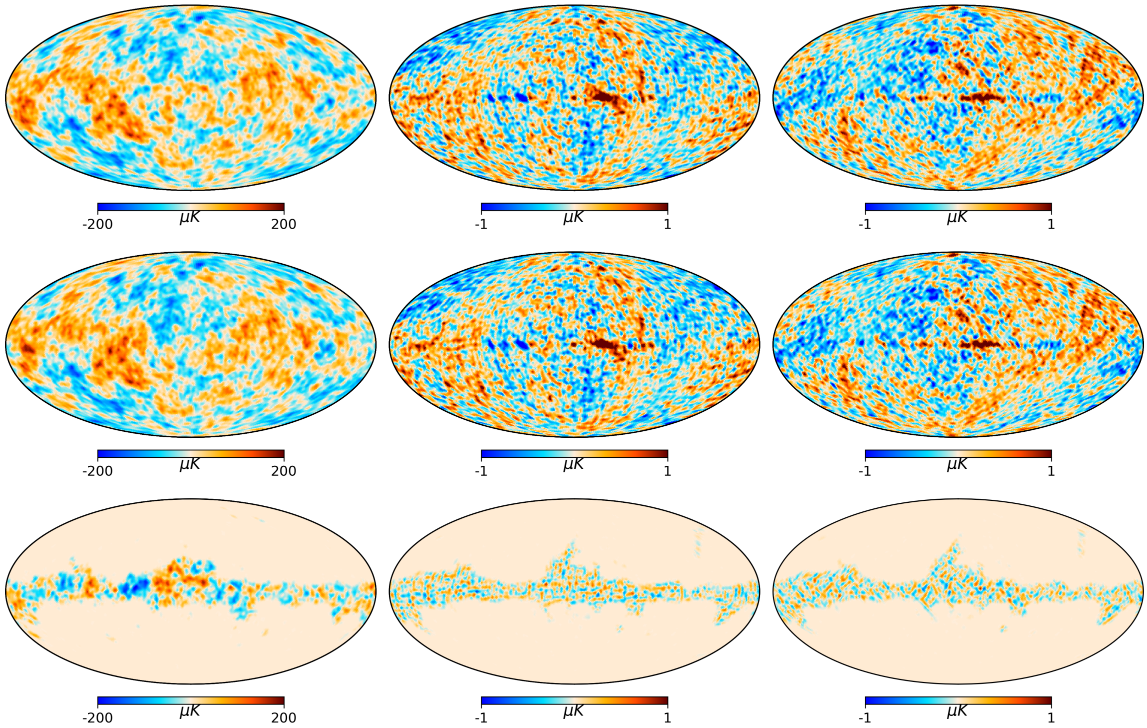

FIGURE 5.6: Example of an inpainted realization in the case where semi-realistic noise is added to the CMB signal. First row corresponds to input T (first column), Q (second), and U (third) maps. An inpainted realization is shown in the second row while the third row shows the difference. All the maps are at $N_{\rm side}$ = 64 and have a resolution of $160'$.

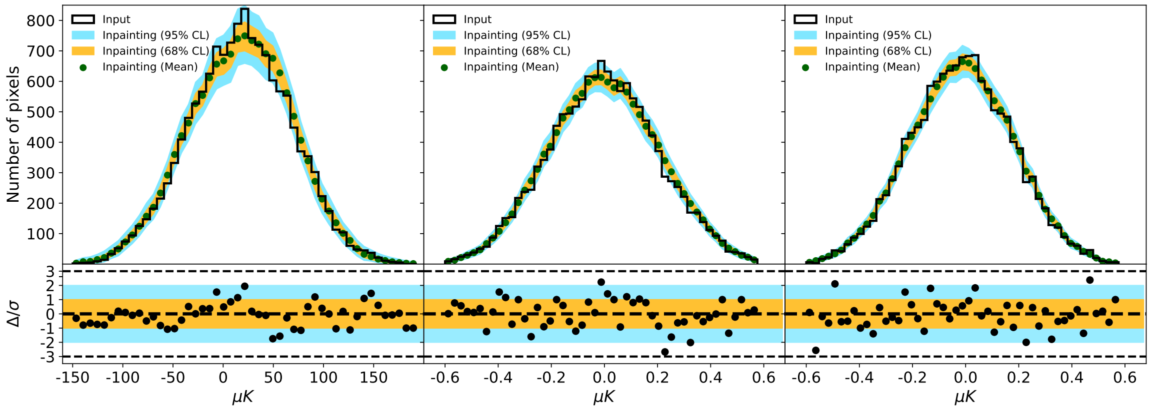

FIGURE 5.7: One-dimensional probability distribution inside the inpainted region for the T (left panel), Q (middle), and U (right) components considering only the CMB signal. The black histogram corresponds to the input map. Green dots are the average value per bin obtained from the 1200 inpainted realisations. Orange and blue contours are the 68% and 95% C.L., respectively, obtained from the distribution of the inpainted maps. Residuals are also shown in the lower panel.

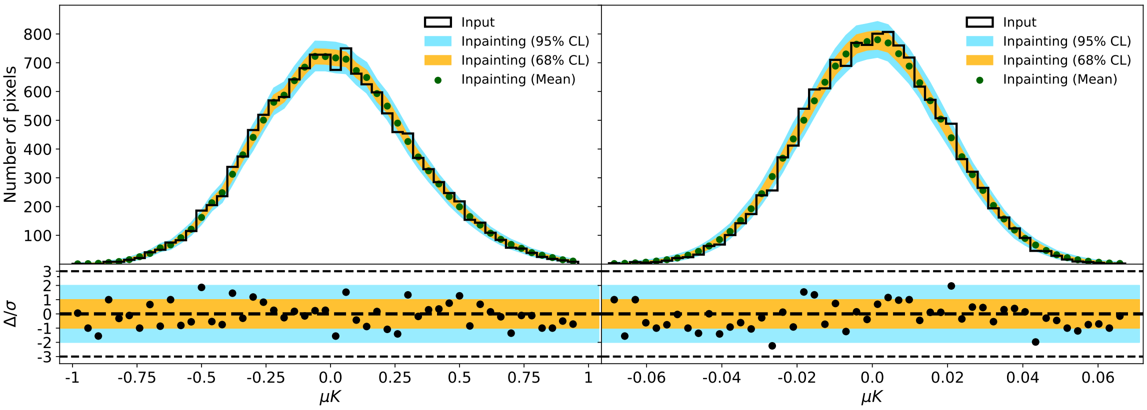

FIGURE 5.8: Same as Figure 5.7, but for E- and B-modes.

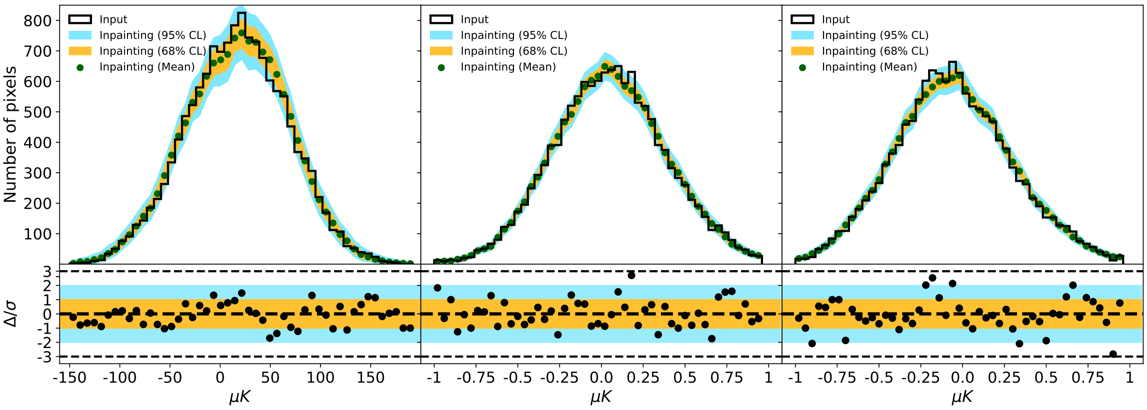

FIGURE 5.9: Same as Figure 5.7, but for the scenario where a semi-realistic noise realization is added to the CMB signal.

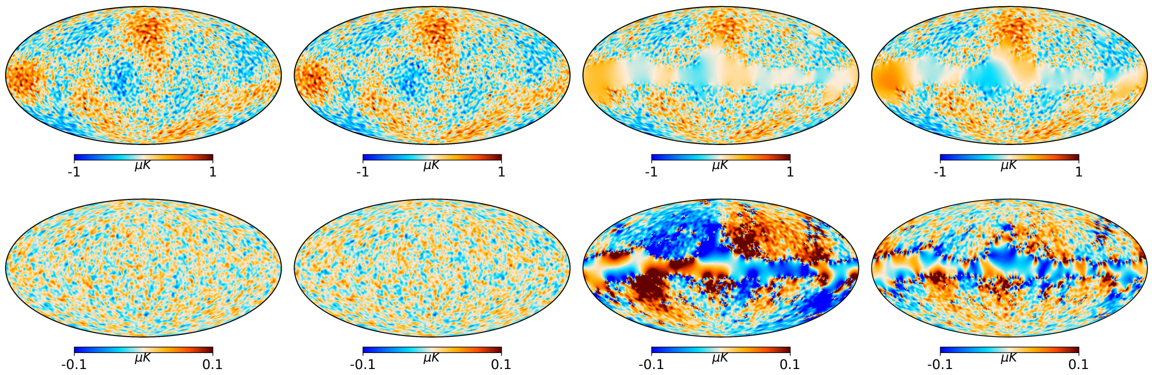

FIGURE 5.10: E (first row) and B-mode (second row) reconstruction for different approaches. First column shows the E- and B-mode maps obtained directly from the full-sky Q and U maps. The recovery obtained after applying our inpainting technique in the masked area is given in the second column. The case in which pixels inside the mask are simply replaced by zeros is shown in the third column. Finally, in the last column, a diffuse inpainting approach is applied on the input Q and U maps before obtaining the E- and B-mode maps.

local. This means that full-sky Q and U measurements are needed in order to have accurate E- and B-mode maps free from E-to-B leakage. Moreover, the reduction of this leakage is one of the main motivations of this work. Precisely, inpainting can fill the masked Q and U regions with a signal statistically compatible with the clean sky outside the mask, removing the potential foreground residuals. This approach is particularly useful in the case of pixel-based estimators that deal with the E- and B-modes maps. For harmonic-based estimators, there are alternative methodologies to deal with a mask. For instance to recover the CMB polarisation power spectra at large scale, the Quadratic Maximum Likelihood method can be used in order to reduce the E-to-B leakage. Additionally, the pseudo-$\mathcal{C}_\ell$ formalism can be used for high multipoles.

As an illustration, Figure 5.10 shows the E- and B-modes maps obtained directly from Q and U full-sky maps versus those recovered after inpainting the masked region of the Q and U maps. For comparison, we also include the E- and B-mode maps generated directly from the masked Q and U maps, and those obtained after applying a diffuse inpainting technique[6] on the Stokes Q and U parameters. In these two last cases, a strong E-to-B leakage can be clearly appreciated.

In order to assess the error in the E- and B-mode reconstructions, we compute the map of the standard deviation of the residuals. Starting from the 1200 T, Q, and U inpainted realisations, we generate the corresponding E- and B-mode maps and compute the residuals by subtracting from them the input E- and B-mode. We calculate then the standard deviation maps, pixel by pixel, which are shown in Figure 5.11. On the one hand, for the E-mode, the maximum error outside the polarisation common mask is $0.042\ \mu K$ which corresponds to around a 14 per cent of the typical amplitude of the E-mode signal, $\sigma_E \sim 0.29\ \mu K$ at

---

[6] Diffuse inpainting consists on filling iteratively the masked pixels taking the average value of the neighbour pixels until convergence is reached.

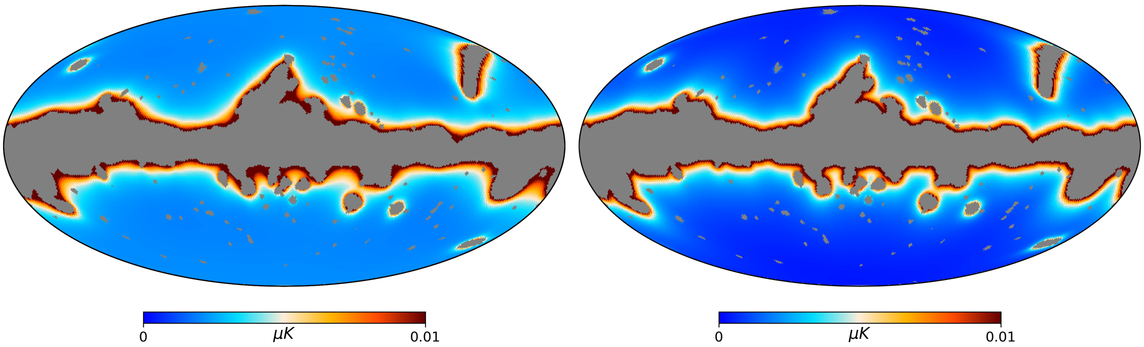

FIGURE 5.11: Standard deviation of the E- (left) and B-mode (right) residuals outside the *Planck* 2018 polarisation confidence mask at $N_{\rm side}$ = 64.

the considered resolution. On the other hand, for the B-mode, the maximum error is at the level of $0.019\ \mu K$, which is approximately the expected amplitude of the B-mode signal for $r = 0$, σ_B. However, as Figure 5.12 shows, the error decreases rapidly, and for 60% of the sky the maximum error is at the level of 20 per cent relative to σ_B. Concerning the E-mode, it is interesting to point out that the error is below 5% for more than the 71.7% of the sky, and then it goes down until it reaches a plateau, where we can not push the maximum residual to a lower value even if we extend the mask.

We get similar results for the absolute error in the E- and B-mode reconstruction when we include the semi-realistic noise realization.

5.3.4 Power Spectra

Our final test is related to the power spectra estimation. Given the input T, Q, and U full-sky maps, we calculate the TT, EE, BB and TE power spectrum, and compare it to the mean power spectra generated from the 1200 inpainted realisations. In particular, for this case where 30 per cent of the sky is inpainted, we find that the distribution of the values for each multipole of the inpainted realisations closely resembles a Gaussian distribution, even for $\ell < 30$. For larger $f_{\rm sky}$ to be inpainted, we expect the distribution to become more like a χ^2. We also calculate the 68% and 95% C.L., as well as the residuals per multipole. The residuals are computed by taking the differences between input and median values over the 1200 realisations. Then, we divide by the upper or lower sigma[7] to take into account possible asymmetries in the distributions, especially for low multipoles. Figure 5.13 shows the TT and TE power spectra, while in Figure 5.14 EE and BB are plotted. Taking into account that we have a single CMB realization, the input spectra is noisy compared to the theoretical prediction due to the cosmic variance, both of them also plotted. Indeed, the

---

[7] Low sigma is computed by integrating from the median to the 16% of the low tail, i.e. it encapsulate 34% of the probability. The other 34% is in the upper sigma which is the integration between the median up to the 84%. For instance, if the residuals are positive, which means that the input value is below the median, we use the low sigma.

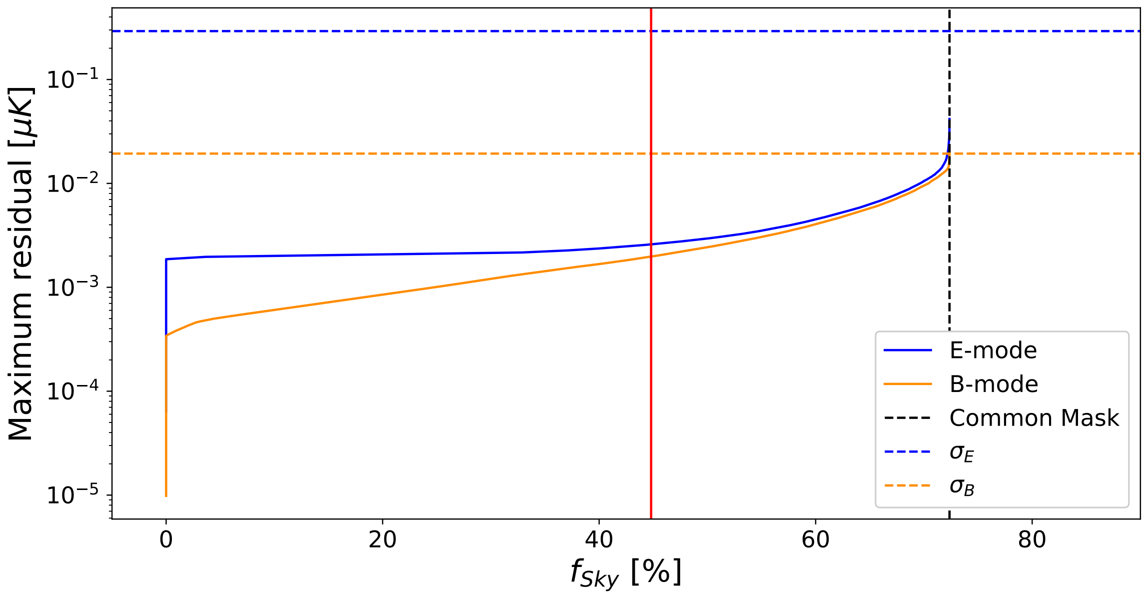

FIGURE 5.12: Evolution of the maximum residual in the E- and B-mode with respect to the available fraction of sky. The vertical black dash line corresponds to the *Planck* 2018 polarisation common mask. The horizontal blue and orange dash lines show the typical value of the E- and B-mode fluctuations, respectively. Finally, the red solid line is fixed to $f_{\rm sky}$ = 44.8%, where the relative error on the B-mode reconstruction is below 10%.

mean power spectra obtained from the inpainted simulations follow that of the input CMB, rather than that of the theoretical model, finding that almost all recovered multipoles fall within the 95% C.L. Note that for the case of the B-mode, we are considering a scenario with $r = 0$ (corresponding to the PR3 ΛCDM best fit model) and, therefore, only the contribution from lensing is present. However, we have checked that the power spectra is equally well recovered when starting with a simulation with r different from zero.

Regarding the scenario where semi-realistic noise is included, we also see a good agreement between the input TT, TE, EE, and BB power spectra and the ones recovered from the inpainted realisations. Note that, as one would expect, in this case the recovered power spectra for EE and BB is above that of the polarisation CMB signal (see Figure 5.15), due to the fact that the noise is the dominant contribution of the maps.

Finally, we perform a last test at the B-mode power spectrum level to compare between our inpainting, diffuse inpainting and a simple masking approach (i.e. put to zero all pixels inside the mask in the Q and U maps and then transform to E and B). We start by generating B-mode maps from the different approaches (GCR, diffuse inpainting and masking) and mask them with the *Planck* polarisation common mask. We compute then the corresponding power spectrum using the `PyMaster` package, the Python implementation of the `NaMaster` (Alonso et al., 2019) library, which computes the angular power spectrum of a masked field using the pseudo-$\mathcal{C}_\ell$ formalism. In particular, the pseudo-$\mathcal{C}_\ell$ are computed using a C2 apodization with 15 degrees and a uniform binning including 4 multipoles per

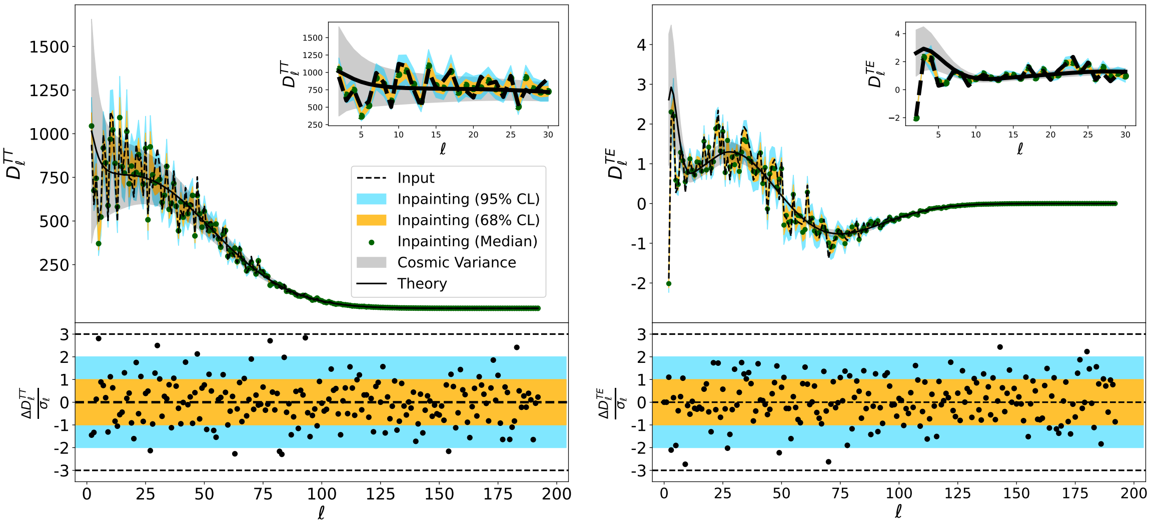

FIGURE 5.13: TT (left) and TE (right) power spectrum scaled by $\ell(\ell+1)/(2\pi)$ (D_ℓ). The solid black line shows the input theoretical model, while the grey area corresponds to the cosmic variance. The dashed black line shows the power spectrum from the input noiseless CMB realization. Green dots correspond to the average value from the 1200 inpainted realisations, while orange and blue contours are the 68% and 95% C.L., respectively. Residuals are shown in the lower panel.

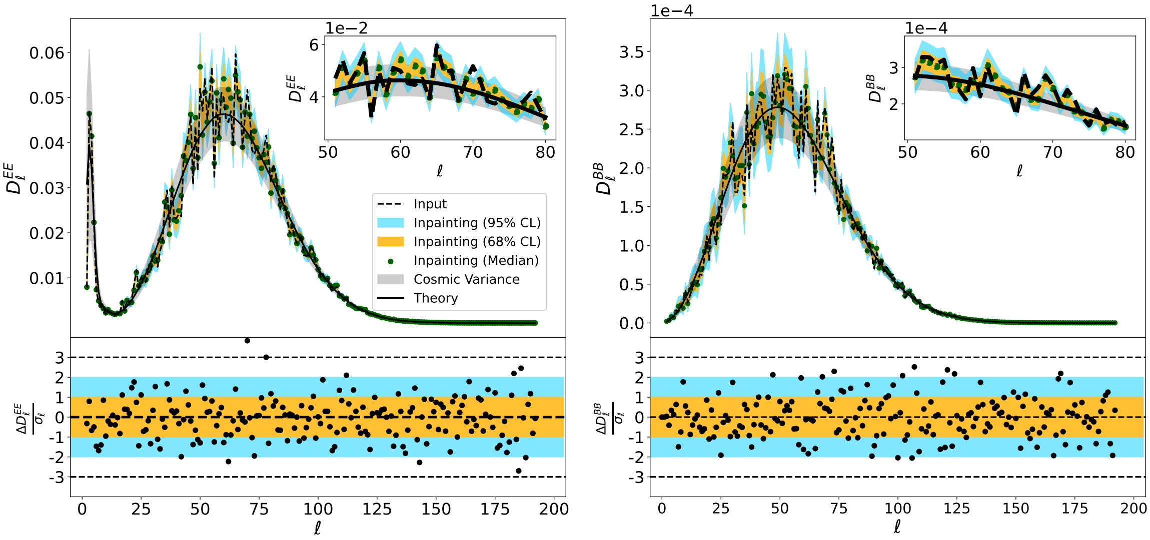

FIGURE 5.14: Same as Figure 5.13, but for the EE (left) and BB (right) power spectrum.

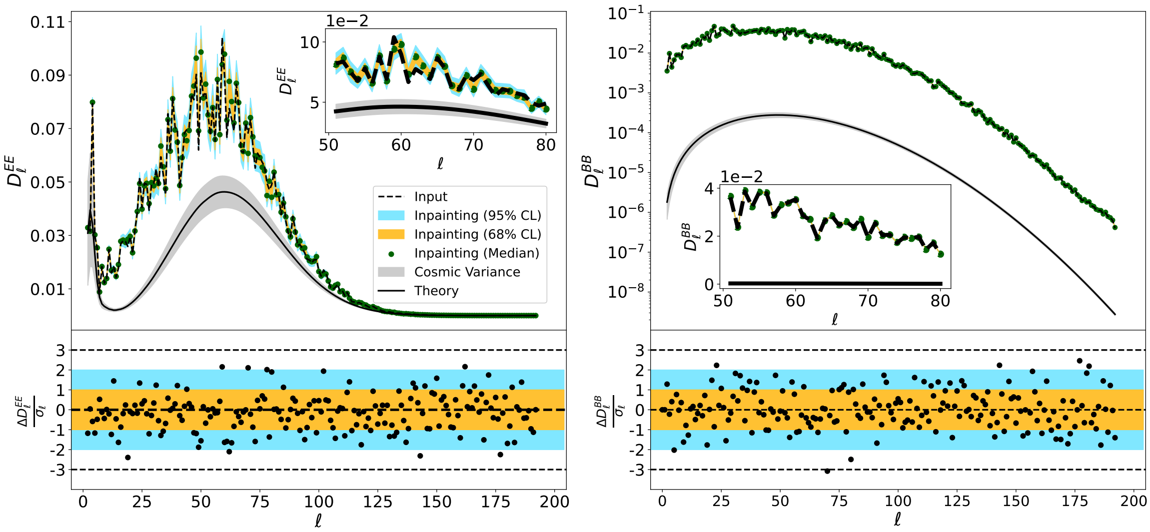

FIGURE 5.15: Same as Figure 5.14, but for the scenario in which CMB plus semi-realistic noise is considered. As reference, we also plot the power spectrum and the cosmic variance for the CMB signal.

bin.

Figure 5.16 shows the results. As expected, for a deep transition in the edge of the mask, i.e. the case where pixels inside the mask are replaced by zeros, the B-mode power spectrum is completely dominated by the E-to-B leakage (dark orange curve). This leakage can be reduced by one order of magnitude applying a diffuse inpainting as it smooths the discontinuity in the Q and U maps. However, the lensing signal is still hidden below the leakage. Our results show that the GCR is the best approach to recover the input power spectrum having residuals below the signal for all the multipole range. For comparison, we also consider the `NaMaster` pure B approach, which recovers the B-mode power spectrum starting from the masked Q and U maps. For the considered case, this approach also fails reproducing the large angular scales of the B-mode. These results show that obtaining the power spectra from an inpainted map could be used as an alternative to more standard methods. However, further work is needed to validate the usefulness of this approach. In particular, in Appendix 8.4 the robustness of the method to recover the input tensor-to-scalar ratio is tested.

5.4 CMB-PAInT

As part of this work, we have developed a Python package called `CMB-PAInT` (**Cos**mic **M**icrowave **B**ackground **P**olarization **A**nisotropies **In**painting **T**ool) to perform inpainting on an input map in the HEALPix format (Górski et al., 2005). This user-friendly package has been made publicly available at `https://github.com/ChristianGim/CMB-PAInT`.

In this section, we briefly describe the software capabilities. As an example, we also

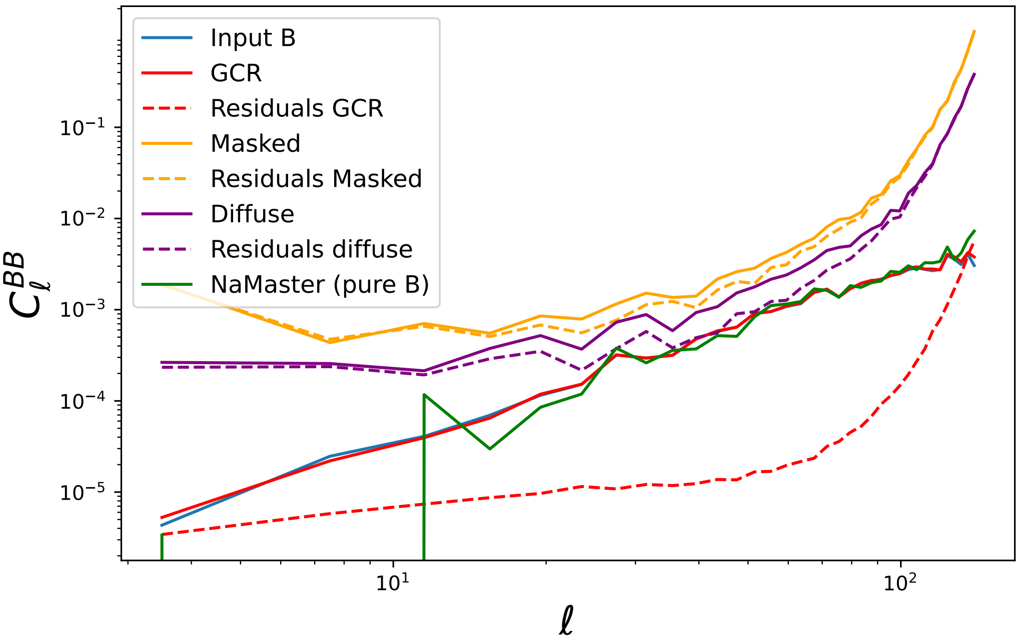

FIGURE 5.16: Pseudo-$\mathcal{C}_\ell$ for the B-mode maps reconstructed from different approaches. Residuals are computed using `NaMaster` and taking as input the difference map. *Planck* 2018 polarisation confidence mask is used. For comparison, the solid green line shows the pseudo-$\mathcal{C}_\ell$ computed from masked Q and U maps by applying the `NaMaster` pure-B technique.

give computational times for the configuration we use at NERSC[8] to perform the inpainting applications shown in Section 5.3. In particular, we consider the case where the inpainting is performed on the *Planck* temperature and polarisation common masks region (14113 and 13583 pixels to be inpainted, respectively) on T, Q, and U components at $N_{\rm side} = 64$.

`CMB-PAInT` can be used in different ways:

- To compute the pixel covariance matrix from an input angular power spectrum up to a certain ℓ_{max}. Depending on the field to be inpainted, it can compute the covariance matrix of either T, QU, or TQU.
- To compute the Cholesky decomposition from an input covariance matrix. The matrix can be just signal, previously computed from an input power spectrum, or the sum of signal plus some extra component (noise, systematics...). Given the mask it also performs the required permutations in rows and columns, i.e. it orders first the unmasked pixels and then the masked ones as explained in Section 5.1.
- To compute the z variable and inpaint the map. If the input is a single sky map, it can generate N different inpainted realisations of the same sky. If the input is a set of

---

[8] National Energy Research Scientific Computing Center (NERSC), https://www.nersc.gov/, is a primary scientific computing facility operated by the Lawrence Berkeley National Laboratory, located in California. It provides high-performance computing and storage facilities where *Planck* latest data and simulations can be found.

maps, it computes for each of them the z variable and an inpainted realization. It also includes an optional parameter, `Cons_Uncons`, to allow one saving the constrained and unconstrained parts of the inpainting process. If True, they are included in the 0 and 1 fields of the output fits file, while field 2 contains the sum.

This code can run on a NERSC-like cluster that uses `slurm` scheduling, or on a local machine, a Jupyter Notebook, or another cluster with different resource management. The only requirement is to consider the memory limitation. The code needs a configuration file that contains all the model and software parameters for running it. For example, the configuration file specifies whether to inpaint the polarisation field or not, or whether to use an external .sh file. The code creates an instance of the CMB-PAInT class and runs one of the methods based on the user input: `calculate_covariance`, `calculate_cholesky`, or `calculate_inpainting`. If the methods run on a cluster without an input .sh file, it generates a .sh file based on NERSC with the resources from the configuration file, such as number of nodes, tasks, CPUs per task, time limits, email address, or partition. The code has two levels of parallelization. It distributes the work among a number of jobs ($N_{\rm jobs}$) that are submitted together, and each job uses `mpi4py`, the MPI standard for Python, to parallelize the assigned rows or maps. This parallelization is used to calculate the covariance matrix from an input power spectrum and also to inpaint the maps. In the latter, the total number of realisations are distributed among the jobs and the tasks per job. Additionally, the code uses the `dask` package to perform Cholesky decomposition faster than the standard `numpy` implementation. The code saves intermediate products in a numpy file format, and Cholesky decomposition in a HDF5 binary data format.

Regarding computational time, the left panel of Figure 5.17 shows the time taken to compute the covariance matrix for a map of $N_{\rm side} = 64$. We use the following configurations:

1. $N_{\rm jobs} = 32$ (Single node, 32 tasks, 4 CPUs per task)
2. $N_{\rm jobs} = 32$ (Single node, 64 tasks, 2 CPUs per task)
3. $N_{\rm jobs} = 32$ (Single node, 128 tasks, 1 CPUs per task)

Since the covariance matrix is symmetric, we only need to compute the first (i) elements of the i^{th} row, which are the subdiagonal and diagonal elements. As expected, time cost increases linearly with the Job ID, because the number of operations increases in the same way. For the proposed configurations, each job computes 1536 rows ($N_{\rm pix}/N_{\rm jobs}$), which are split among the number of tasks. For instance, in configuration (1), the first task of the first job computes the rows between 0 and 47 (1176 elements), while the last task computes the rows between $1488 - 1535$ (72600 elements). It is straightforward to conclude that the bottleneck of each job is the last task, which computes the largest number of operations. In the right panel of Figure 5.17 we display the number of elements/operations that are done by

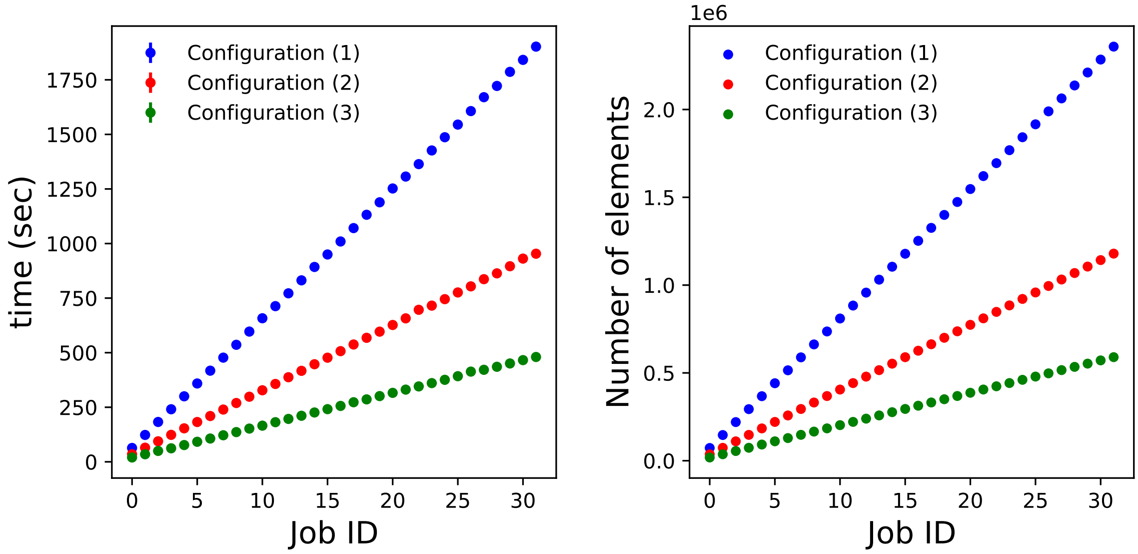

FIGURE 5.17: *Left panel:* Computational time to calculate the covariance matrix for 3 different setups. *Right panel:* Number of elements of the last task of each job for the three configurations.

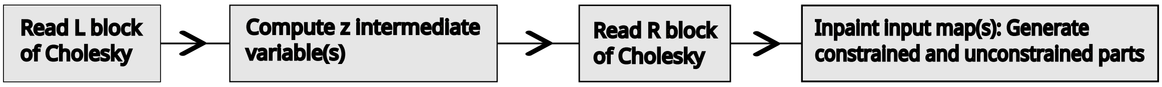

FIGURE 5.18: Diagram of the workflow in `CMB-PAInT` for the inpainting algorithm.

the last task of each job. For the configuration (2) the time cost is reduced by almost a factor of 2, as we assign more tasks per job and leave fewer elements to the last task. In this sense, the optimal configuration is (3). However, giving some CPUs per task could be necessary due to memory issues. Additionally, the code could have an extra parallelisation layer for configurations (1) and (2) if OpenMP API is used, which will improve the performance.

Figure 5.18 shows a diagram of the workflow for the inpainting method once the covariance matrix and the Cholesky decomposition[9] are computed. In this case, we try the following configuration:

- $N_{\text{jobs}} = 16$ (3 nodes per job, 10 tasks, 4 CPUs per task)

Figure 5.19 shows the time cost of each step in the workflow. The bottleneck is the reading of the L matrix, which is the largest block in the Cholesky decomposition. It takes less than 300 seconds for most ranks[10] and jobs. However, some ranks get stuck and take more than 700 seconds to read it. The next step is computing recursively the z variable, which takes around 4.5 seconds with slight variations. Reading the R matrix is faster than reading the

---

[9] It takes around 30 to 40 minutes to compute the covariance matrix and the Cholesky decomposition in a Perlmutter node for the considered example: $N_{\text{side}} = 64$, all the components (T, Q, and U) and *Planck* common masks.

[10] In the MPI context, every process that belongs to a communicator is uniquely identified by its rank, which is an integer that ranges from zero up to the size of the communicator minus one.

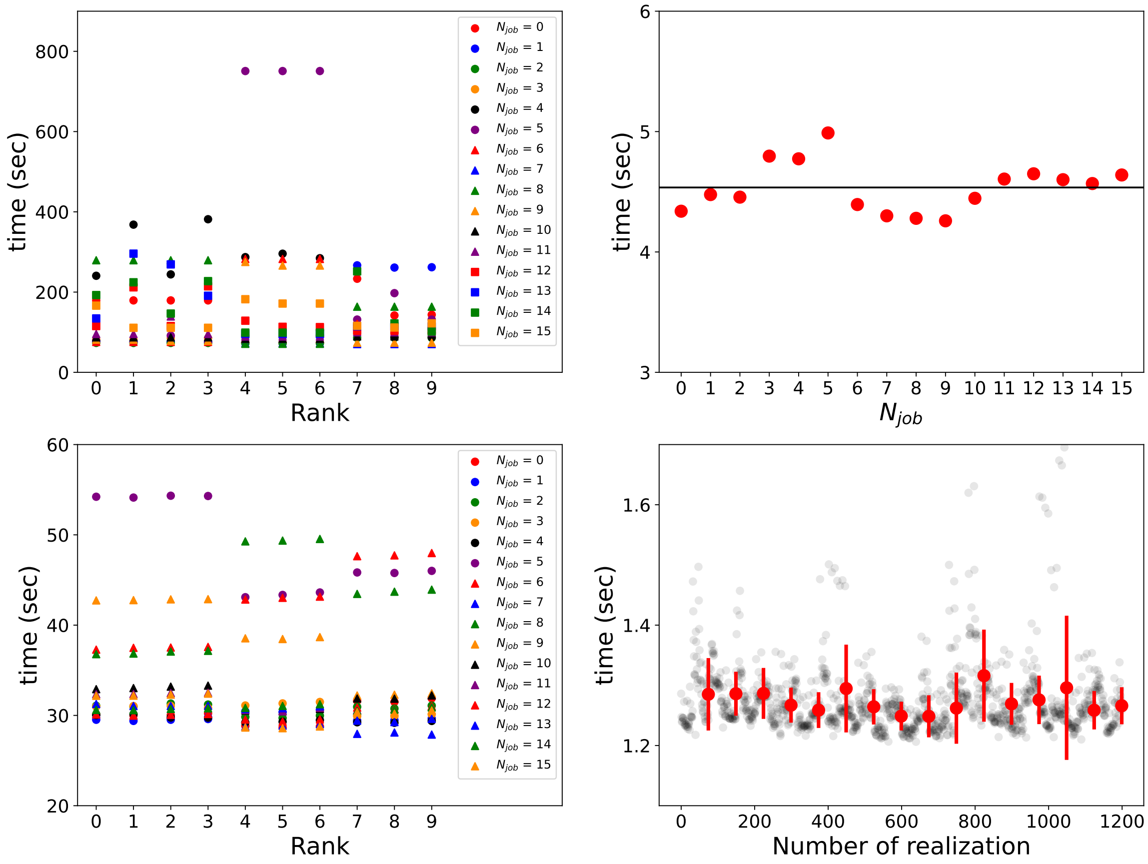

FIGURE 5.19: Computational time of each of the steps in Figure 5.18. *Upper left:* Computational time per rank and per job to read the L matrix. *Upper right:* Computational time per job to compute the z variable. *Lower left:* Computational time per rank and per job to read the R matrix. *Lower right:* Computational time to inpaint each realization (open grey dots). In red the mean time and dispersion per job (note that each job inpaints $N_{\rm sims}/N_{\rm jobs}$ realisations).

L matrix, but there is a large variation in this run. On average, this step takes around 30 seconds. The last step is inpainting a map, which takes less than 1.3 seconds on average.

6

Exploring Statistical Isotropy in *Planck* Data Release 4: Angular Clustering and Cosmological Parameters Variations Across the Sky

In this chapter, we test statistical isotropy in *Planck* Data Release 4 (PR4) by estimating the temperature and E-mode power spectra across independent sky regions. We find that the directions with higher local bandpower amplitudes in intensity are clustered for multipoles between 200 and 2000 with clustering probabilities consistently below 1% for all these scales when compared to end-to-end (E2E) *Planck* simulations; notably, this range extends beyond that reported in *Planck* Data Release 3 (PR3). However, no significant clustering is observed in the polarisation E-modes. In a complementary analysis, we search for dipolar variations in cosmological parameters fitted using the previously computed power spectra. When combining temperature and polarisation power spectra, we identify a potential anomaly in the amplitude of the primordial power spectrum, A_s, with only 5 out of 600 simulations exhibiting a dipole amplitude as large as that observed in the data. Interestingly, the dipole direction aligns closely with the known hemispherical power asymmetry, suggesting a potential link between these anomalies. All other cosmological parameters remain consistent with ΛCDM expectations. These findings highlight the need for further investigation of CMB anomalies to clarify their nature and assess their potential implications for our understanding of the early Universe.

6.1 Data and Methodology

6.1.1 Data

We use the data from *Planck* Public Release 4 (PR4)[1], which has been processed by the NPIPE pipeline described in Planck Collaboration et al. (2020f). This new pipeline reprocesses the *Planck* Low-Frequency Instrument (LFI) and High-Frequency Instrument (HFI) in a joint

---

[1] Data is available at *Planck* Legacy Archive (PLA), `https:/pla.esac.esa.int/`.

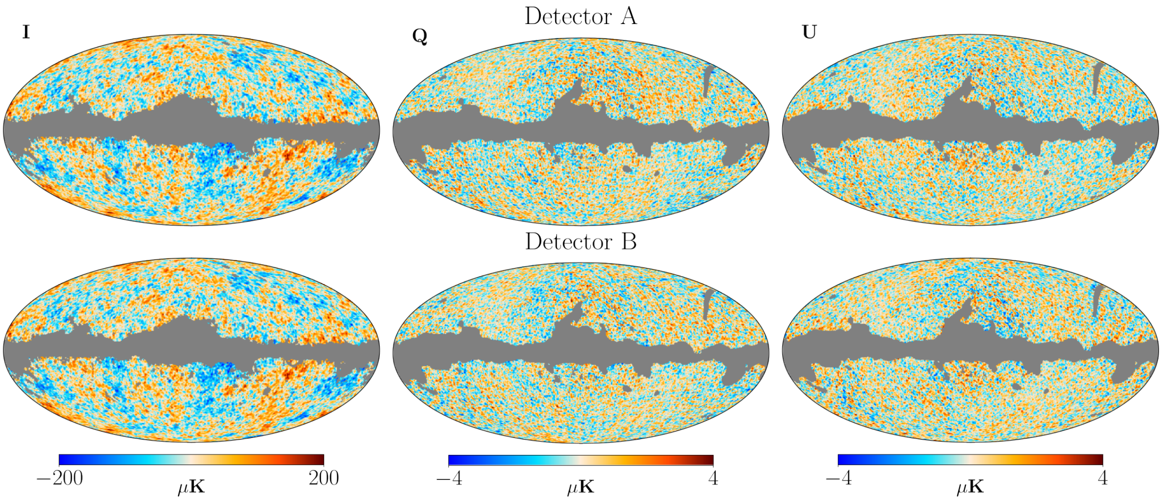

FIGURE 6.1: *Planck* PR4 detector A and B `Sevem` cleaned maps. First row shows the IQU maps, from left to right, for detector A, while the second row shows the B detector maps. All of the maps are smoothed with 1 degree FWHM Gaussian beam anx presented in Galactic coordinates. The grey area corresponds to the *Planck* PR3 confidence mask appropriate to either temperature or polarisation..

analysis, which effectively reduces the noise and systematics in frequency maps. In particular, we use the A and B detector splits cleaned with the `Sevem` (Fernández-Cobos et al., 2012) component separation method. To asses the p-values we use the 600 available *Planck* PR4 "end-to-end" (E2E) simulations, which include realisations of the CMB signal, the instrumental noise, and the systematics. These simulations try to capture all the characteristics of the full data processing such as the scanning strategy and the detector responses, and sky realisations are generated also including effects such as lensing, Rayleigh scattering, and Doppler boosting. Combining independent sets of detectors allows us to use the cross-spectrum between maps avoiding noise bias. In order to mask the Galactic residual foregrounds and the extragalactic point sources, we use the PR3 *Planck* 2018 confidence masks described in Planck Collaboration et al. (2020b). These masks leave a fraction of available sky close to 78% for both temperature and polarisation. In Figure 6.1 we show the detector A and B PR4 `Sevem` cleaned maps together with the *Planck* PR3 confidence masks.

6.1.2 Analysis Pipeline

Our pipeline can be outlined as follows:

1. Power spectrum estimation:

   We first estimate the angular power spectrum of each masked map using the well-known *pseudo-*C_ℓ `MASTER` estimator (Hivon et al., 2002). The analysis is performed on full-resolution CMB maps at $N_{\text{side}} = 2048$. We divide the sky into 12 disjoint and equal-area regions, corresponding to the base pixels of the `HEALPix` (Górski et al., 2005) pixelisation at $N_{\text{side}} = 1$ (see Fig. 6.2 for the numbering of the patches). For

Patch	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
$f^T_{\rm sky}\,[\%]$	7.2	7.3	7.8	7.5	2.3	4.0	3.4	4.9	7.6	7.5	7.3	7.0
$f^P_{\rm sky}\,[\%]$	7.7	7.3	7.7	7.6	2.4	4.5	3.5	5.2	7.8	7.5	7.4	7.5

Table 6.1: Fraction of sky ($f_{\rm sky}$) for intensity and polarisation for each of the 12 considered patches.

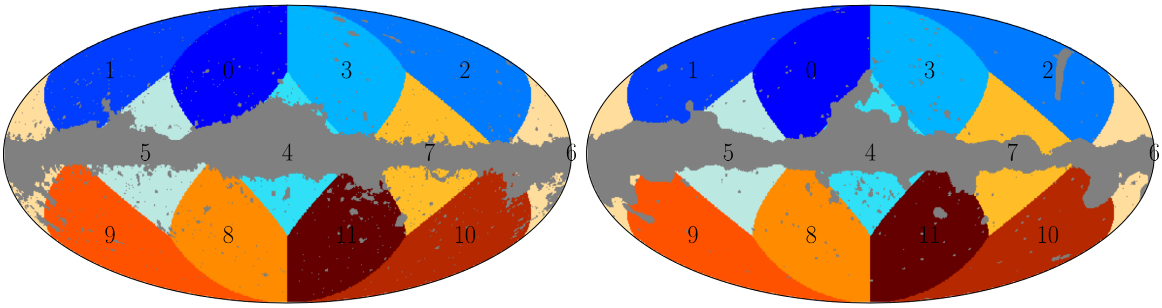

FIGURE 6.2: Numbering of the patches in a `Healpix` map at $N_{\rm side}$ = 1 resolution, along with the *Planck* confidence masks in grey. The left panel shows the intensity mask, while the right panel presents the polarisation mask.

each region, we compute a single angular power spectrum over the entire multipole range, applying binning in multipole space with a fixed bin width of $\Delta\ell = 30$. The specific bin width is chosen to mitigate the mode coupling induced by the mask[2]. The sky fraction, $f_{\rm sky}$, varies between 2% and 8% depending on the overlap between the Planck confidence mask and each base pixel region. Due to this reduced sky coverage, we have very limited information at the largest angular scales. Therefore, we exclude the first bin ($\ell = 2$–31) for most of the analyses. Table 6.1 reports the exact $f_{\rm sky}$ values for each patch, separately for intensity and polarisation. The maximum multipole used in the analysis is selected to avoid contamination from unresolved point sources and to exclude scales where instrumental noise dominates the signal. These are our default choices:

- For TT: 66 bins in the range $\ell = [32, 2011]$
- For TE: 57 bins in the range $\ell = [32, 1741]$
- For EE: 48 bins in the range $\ell = [32, 1471]$

We applied a 0.3 degree apodization to each mask, effectively reducing correlations between multipoles at small scales in the TT power spectrum. The reason behind this choice will be explained later.

2. Angular Clustering:

---

[2] This choice yields a correlation matrix that is close to diagonal while preserving a sufficient number of bins to perform a robust likelihood analysis.

We then proceed to an analysis of the *Planck* PR4 data to examine the angular clustering anomaly previously noted by the *Planck* collaboration (see Section 7.3 of Planck Collaboration et al. (2020d)). Specifically, the power spectra were computed locally in patches for various multipole ranges, and dipoles fitted to maps of the band-power estimations. Using the Rayleigh Statistic (RS), an anomalous alignment between the dipoles in the temperature data were found, with a significance level of 2σ–3σ. This behaviour is not expected in the standard cosmological model, which allows for the existence of dipolar power distributions but predicts that their directions should be completely random. Therefore, this alignment could be evidence of a deviation from statistical isotropy.

Once the binned TT, TE and EE power spectra have been estimated as described above, we adopt the methodology of Planck Collaboration et al. (2020d).

(a) For each multipole bin of the binned TT, TE, and EE spectra, we construct a `HEALPix` map at $N_{\mathrm{side}} = 1$, assigning the value of the corresponding spectrum in that bin to each pixel.

(b) A dipole is fitted to this map using inverse-variance weighting. The variances are computed from simulations. In particular, we are interested in the direction of the fitted dipole as the amplitudes are found to be fully consistent with the simulations.

(c) The estimator we use to measure the alignment is the modulo of the sum of all the normalized dipole vectors up to certain maximum bin,

$$|\hat{v}| = \sqrt{N + \sum_{i \neq j} \cos\theta_{ij}}, \tag{6.1}$$

where N is the number of dipoles, and θ_{ij} the angle between the i-th and j-th dipoles. This is essentially the Rayleigh statistic (RS), a statistical measure used to test uniformity, particularly for assessing whether a set of vectors exhibits any preferential alignment. Actually, Eq. 6.1 differs from the RS by not including any amplitude information. As previously mentioned in Planck Collaboration et al. (2020d), the amplitude of the dipole vectors are not anomalous, so they are normalized. Apodization of the masks was essential to mitigate the clustering observed among small-scale dipoles. In simulations, where dipoles are expected to be uniformly distributed, we found that the RS value starts deviating from expectation at small scales. This deviation was later identified as artificial clustering caused by correlations between small-scale dipoles. We selected a 0.3 degree apodization scale as the minimum value necessary to recover an RS value that converges to the expected value. Minimizing the apodization scale is crucial to preserve as much of the sky as possible.

(d) Finally, we asses the clustering as a function of maximum bin using a *p*-value determined as follows. We compute the RS using all the dipoles up to a certain maximum multipole bin for all the simulations and the data. Then, we define the *p*-value as the fraction of simulations with a higher RS than the one observed in the data. A small *p*-value means that the directions in the data are clustered in a way that can only be reproduced in a few simulations. Note that *p*-values are correlated as they are defined from a cumulative quantity.

3. Parameter Estimation:

In order to analyse the *Planck* temperature and E-mode polarisation maps to investigate possible dipolar variations of the cosmological parameters, we proceed as follows:

(a) We start with the previously computed binned TT, TE and EE power spectra. The choice to use these regions for estimating cosmological parameters is guided by two primary considerations: first, the patches are disjoint eliminating correlations between them, and second, this approach is computationally efficient while remaining adequate for capturing parameter variations at the dipolar level.

(b) Best-fit cosmological parameters are inferred from the measured C_ℓ values in each patch. In fact, instead of adopting a traditional MCMC method, which would be computationally expensive, we use `iMinuit`[3] to fit for the maximum likelihood values. We use a multivariate Gaussian likelihood, which it is a good approximation for our analysis choices ($\ell \geq 32$),

$$-2 \log \mathcal{L} \propto (C_\ell - \bar{C}_\ell) \cdot C^{-1}_{\ell\ell'} \cdot (C_{\ell'} - \bar{C}_{\ell'})^T, \tag{6.2}$$

where C_ℓ is the observed binned power spectrum, $\bar{C}_\ell$ is the theoretical binned power spectrum computed from `camb`[4] Boltzmann solver (Lewis et al., 2000; Lewis & Challinor, 2011), and $C^{-1}_{\ell\ell'}$ is the inverse of the covariance matrix of the power spectrum C_ℓ. The dimension of the covariance matrix is $n^{TT}_{\text{bins}} + n^{TE}_{\text{bins}} + n^{EE}_{\text{bins}}$, $n^{TT}_{\text{bins}} + n^{TE}_{\text{bins}} + n^{EE}_{\text{bins}}$). Given the size of the bins used in our analysis ($\Delta\ell = 30$), we can reasonably assume that the correlations between different bins are minimal. As a result, we neglect all off-diagonal terms in all the blocks of the covariance matrix, focusing only on the variances and the covariances between the TT, TE, and EE components within the same bin.

(c) We fit for the basic flat-space ΛCDM cosmological parameters[5] (H_0, $\Omega_c h^2$, $\Omega_c h^2$, A_s, n_s). We fix the optical depth to reionization, τ, to 0.0602, the input value

---

[3] `https://scikit-hep.org/iminuit/`

[4] `https://camb.readthedocs.io/en/latest/`

[5] Note that, in the likelihood, we include some bounds on the parameters to prevent the minimizer from exploring regions of the parameter space where `camb` breaks down.

for the E2E simulations, which is also in good agreement with the latest constraints (Tristram et al., 2024). The main reason to fix τ is that the sky fractions for the individual patches are insufficient to assess the E-mode large-angular scales, where most of the information about the optical depth is encoded. We also fix $\sum m_\nu = 0.06$ eV, and $r = 0.01$, which is also the input for the simulations. We do not expect any impact on results with this choice, as this tensor-to-scalar ratio is below the sensitivity of *Planck*. Additionally, following Planck Collaboration et al. (2016a) we include two effective foreground residual parameters (A_{ps}^{TT}, A_{ps}^{EE}), which account for residual contamination from unresolved compact objects. We assume these residuals to behave as shot-noise, modelled as $D_\ell \propto \ell^2$, where $D_\ell = \frac{\ell(\ell+1)}{2\pi} C_\ell$.

(d) We generate an $N_{\rm side} = 1$ map for each of the parameters, from simulations and data. We fit for a dipole in each of them using an inverse variance weighting approach, where variances are estimated from the 600 simulations. In this way, the Galactic low-latitude patches where the $f_{\rm sky}$ is smaller contribute less to the fit. Finally, the p-value is defined as the fraction of simulations with an amplitude of the fitted dipole larger or equal to the one observed in the data map.

6.1.3 Pipeline Validation

In order to validate our pipeline for estimating cosmological parameters, we compare the results that we obtain using `MASTER` and `iMinuit` with those reported in Tristram et al. (2024). In that work, the parameters were not inferred from high-resolution foreground-cleaned CMB maps, but instead were based on likelihoods that used the cross-spectra between pairs of frequency channels. In this work, we are not interested in the absolute value of the parameters, but on their possible dipolar variation over the sky.

Figures 6.3, 6.4, and 6.5 show the cross-detector power spectra obtained with our pipeline for the *Planck* PR4 `Sevem` cleaned maps masked with the *Planck* confidence masks, along with the corresponding best-fit spectra computed using the cosmological parameters we obtain from `iMinuit`. The error bars are computed using the 600 E2E simulations.

Tests of the pipeline with simulations revealed a bias in the TT, TE, and EE power spectrum estimation, which results in a small bias for the inferred cosmological parameters. We characterise this by computing the difference between the input fiducial power spectrum (based on the cosmology described in table 6 of Planck Collaboration et al. (2020f)) and the average of the output cross-spectra over the 600 E2E simulations propagated through `Sevem`. Subtracting this bias from each simulation, we are able to recover the input cosmological parameters. The observed bias can be explained by the mismatch between the realistic beams used in the simulations and the assumed effective beams in the analysis. Additionally, some contribution may arise from frequency-dependent effects that are not accounted for in the component separation pipeline, such as boosting. Another important point is that the bias

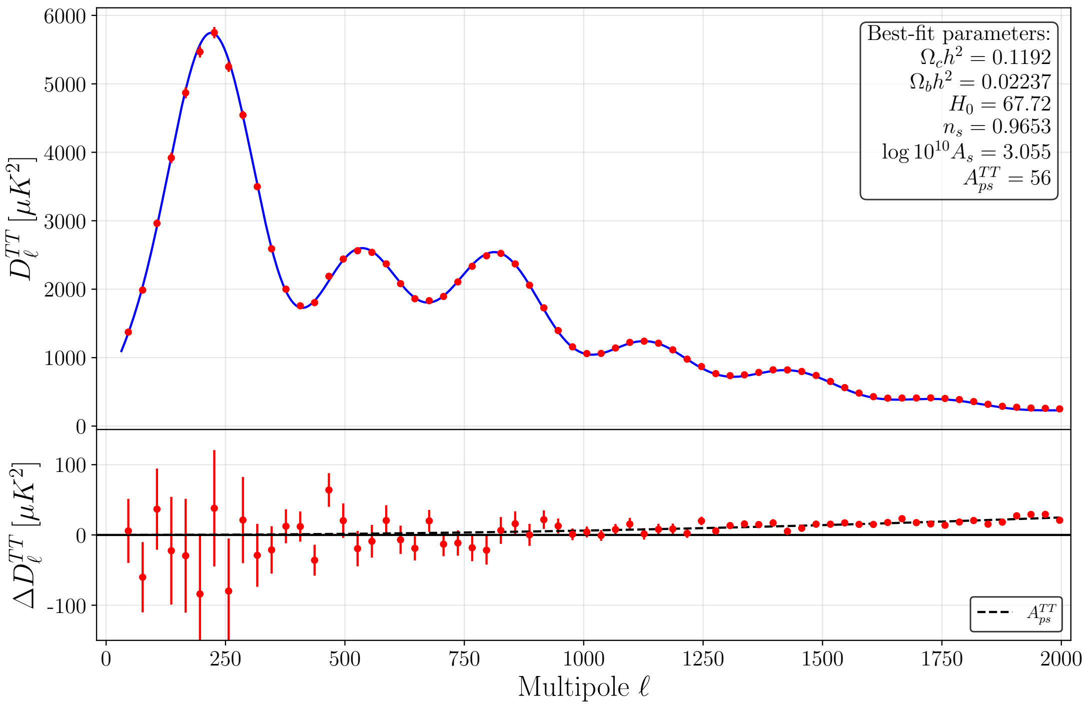

FIGURE 6.3: *Upper panel*: TT cross-power spectrum between A and B detectors of the PR4 `Sevem` cleaned maps (red symbols), and the best-fit ΛCDM model (blue solid line). Best-fit cosmological parameters are also provided in the box, here top right, along with the A_{ps}^{TT} nuisance parameter. *Lower panel*: Residuals with respect to the best-fit model. The black dashed line represents the contribution of the emission of unresolved compact objects modelled as $D_\ell \propto \ell^2$ for the fitted value of the A_{ps}^{TT} nuisance parameter.

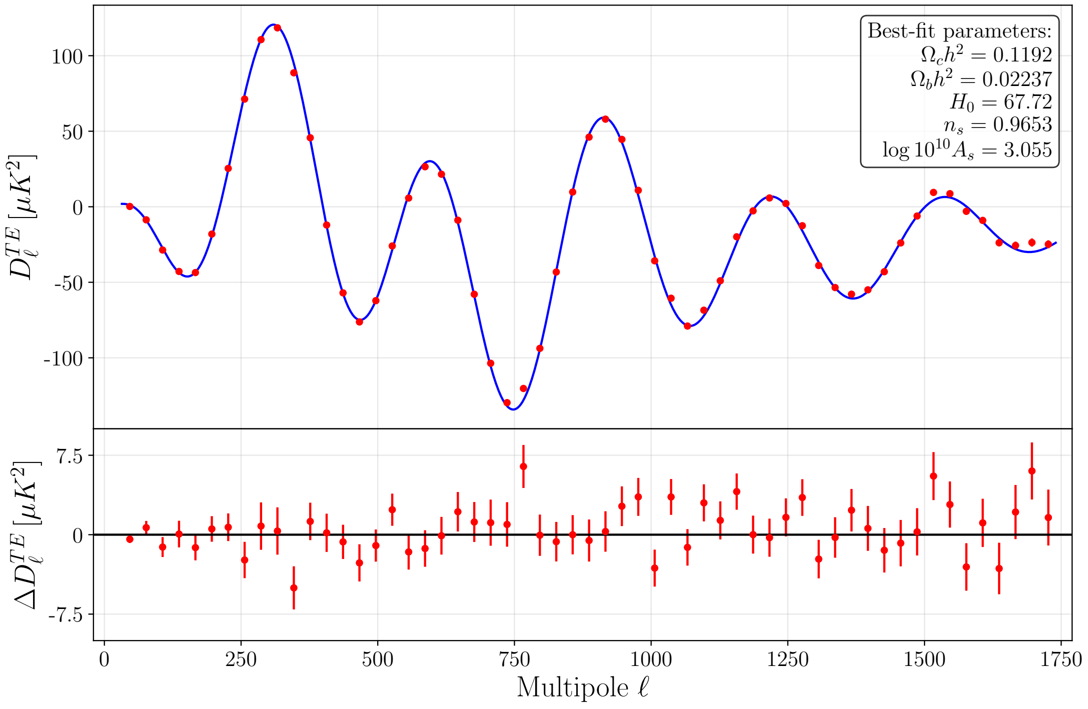

FIGURE 6.4: As Figure 6.3 for TE.

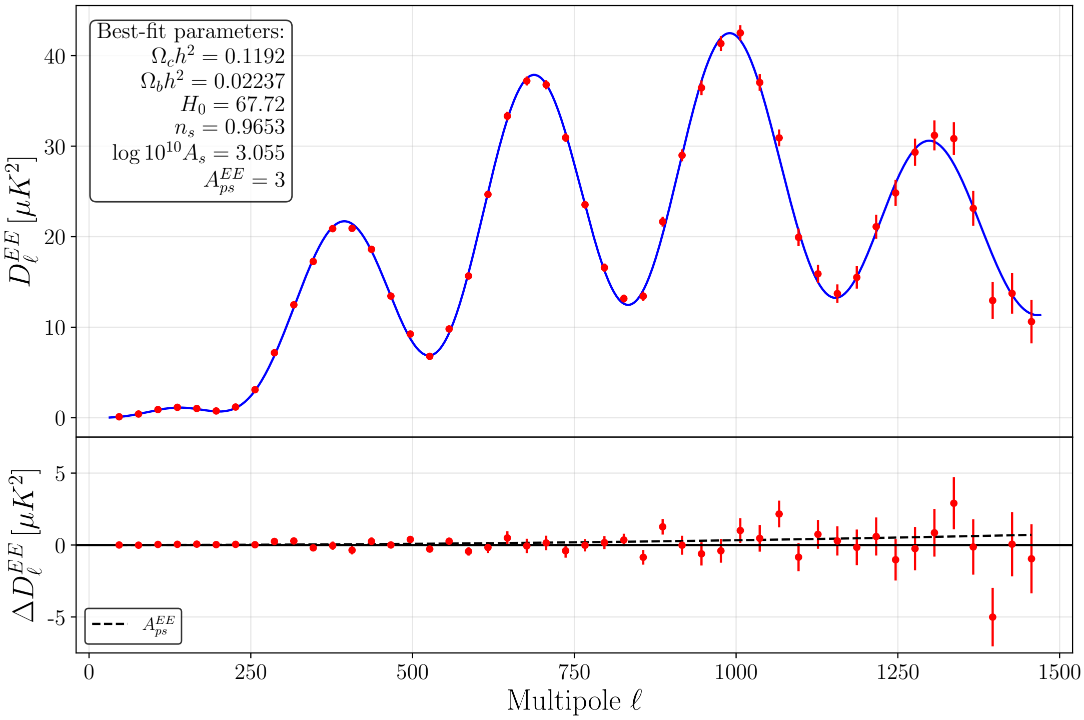

FIGURE 6.5: As Figure 6.3 for EE.

observed in the parameters in the E2E PR4 simulations is similar to that found in the data when compared to the results of Tristram et al. (2024), where the analysis is performed at the frequency map level before the component separation pipeline. By this, we mean that each parameter is biased in the same direction and by a similar amount. The bias is not a significant concern for our analysis for two reasons. Firstly, it should affect the data and simulations in the same way. Secondly, the bias is a subdominant effect when working on small patches because the uncertainties increase given the small sky fraction. Nevertheless, we have checked the robustness of our results by fitting the parameters in the patches after debiasing the power spectra, finding consistent results (see Appendix 8.4). In this case, the bias is determined for each of the patches independently. Table 6.2 summarizes the best-fit cosmological parameters, where the error bars are estimated from the 600 E2E PR4 simulations. In particular, after correcting for the bias in the power spectrum estimation, our results (third column) are very consistent with the ones reported in Tristram et al. (2024) (fourth column). For some parameters we are $\sim 1\sigma$ away from the official *Planck* values, but this is expected given that we use a more limited ℓ range, different masks, and foreground cleaned maps. Note that the error in the A_s parameter is a factor of 4 smaller than in Tristram et al. (2024). This is due to the fact that we are fixing the optical depth at reionization, which is highly correlated with the amplitude of the scalar primordial perturbations. We get similar results for the cosmological parameters when considering the weighted average over the 12 patches.

Parameter	$TTTEEE$ (no debiasing)	$TTTEEE$ (debiased)	PR4 ($TTTEEE$)
H_0	66.78 ± 0.50	67.72 ± 0.50	67.64 ± 0.52
$\Omega_b h^2$	0.02212 ± 0.00013	0.02237 ± 0.00013	0.02226 ± 0.00013
$\Omega_c h^2$	0.1209 ± 0.0011	0.1192 ± 0.0011	0.1188 ± 0.0012
$\ln(A_s \cdot 10^{10})$	3.057 ± 0.0033	3.055 ± 0.0033	3.040 ± 0.014
n_s	0.9598 ± 0.0036	0.9653 ± 0.0037	0.9681 ± 0.0039
A_{ps}^{TT}	55 ± 4	56 ± 4	-
A_{ps}^{EE}	0 ± 1	3 ± 1	-

Table 6.2: Best-fit cosmological parameters. The second and third columns present the results from our pipeline without and with bias correction (no debiasing and debiased, respectively). These values are obtained using the *Planck* PR3 confidence masks. The fourth column shows the latest cosmological parameters derived in ref. Tristram et al. (2024) from the TT, TE, and EE power spectra using the PR4 dataset. The last two rows provide the results for the nuisance parameters which are given by their values at $\ell = 3000$. The error bar in A_s for the first and second columns are much smaller because τ is fixed.

6.1.4 Bayesian Approach

We perform an MCMC analysis on the debiased data as an extra validation of our pipeline. For this purpose, we use `cobaya`[6] (Lewis & Bridle, 2002; Lewis, 2013; Torrado & Lewis, 2021). Figure 6.6 shows the posteriors for each of the parameters together with the value inferred from `iMinuit`. We use flat priors in the same region where minimization is performed, and compute the posteriors for two scenarios: fixing τ (red contours), and leaving it free but with a Gaussian prior $\mathcal{N}(0.06, 0.006)$ (blue contours) applied. This test reveals three key conclusions. First, the minimum found by `iMinuit` is fully consistent with the position of the peak in the posterior for all parameters, showing the robustness of the minimizer. Second, the width of the posterior aligns remarkably well with the standard deviation of the minimum values obtained from the 600 simulations. Finally, the posterior width for A_s is significantly reduced when τ is fixed.

6.2 Results

6.2.1 Angular-Clustering

In this section, we show the results for the angular-clustering analysis. The three panels in Figure 6.7 show the TT, TE, and EE dipole directions determined in bins of $\Delta\ell = 30$ for the A/B detector splits of the `Sevem` PR4 data. The plots are rotated in such a way that the centre of the image is located at $(\ell, b) = (205, -20)$ in Galactic coordinates. This is the direction of the HPA found in the temperature data in Gimeno-Amo et al. (2023). The left column of Figure 6.8 presents the corresponding RS values as a function of $\ell_{\max}$, while the right column shows the associated p-values. We consider as our reference case that obtained without including the first bin ($2 \leq \ell < 32$), which is given by the green line

---

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

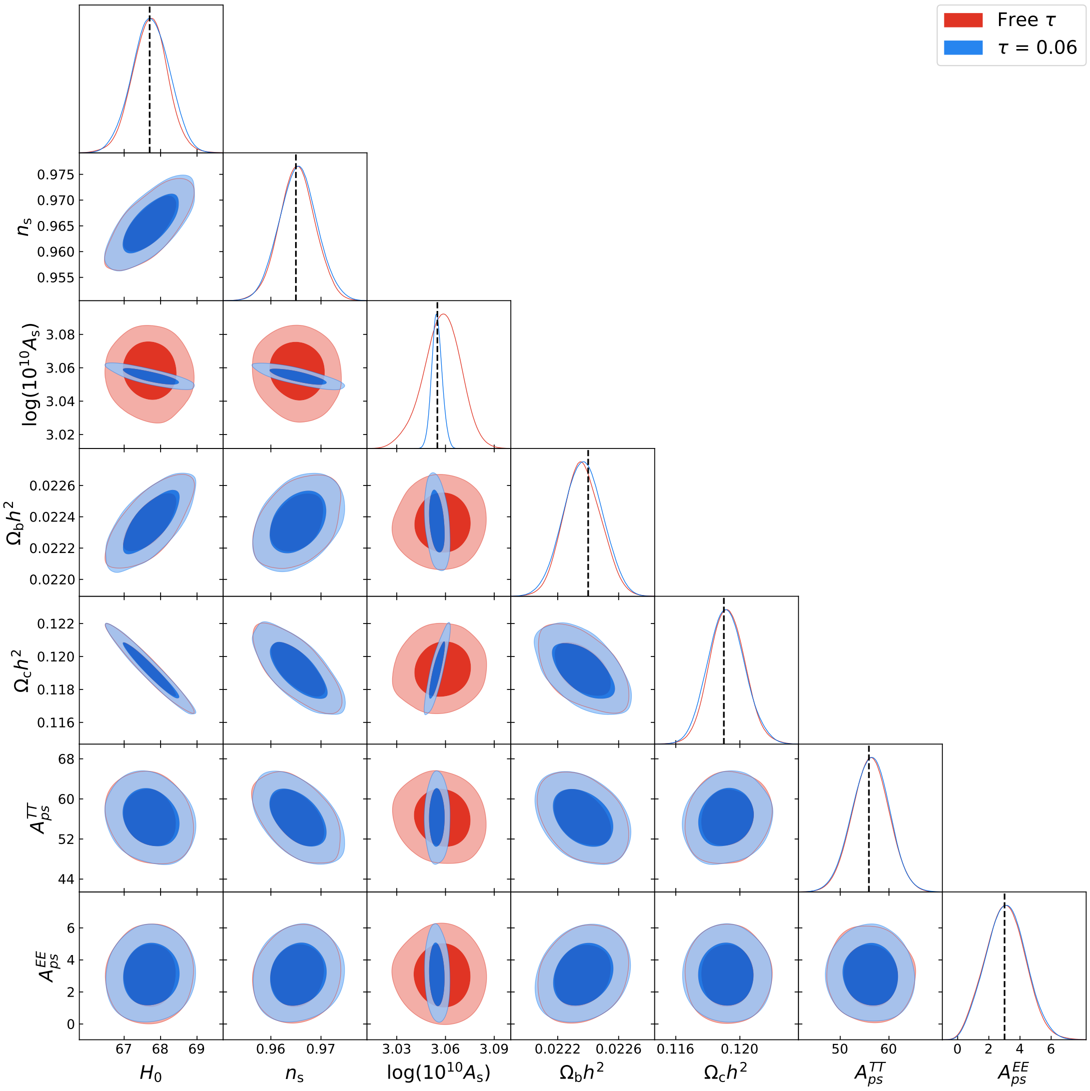

FIGURE 6.6: Constraints on the base ΛCDM model using the TT, EE, and TE cross-spectra between detector A and B for the `Sevem` PR4 cleaned maps. A binned power spectrum, starting from ℓ = 32, and a Gaussian likelihood are used. Diagonal plots are the marginalized parameter constraints. Black dashed lines correspond to the parameters obtained with `iMinuit`. Two scenarios are considered: fixing $\tau = 0.06$ (in red) and leaving τ as a free parameter (in blue). Contours contain 68% and 95% of the probability.

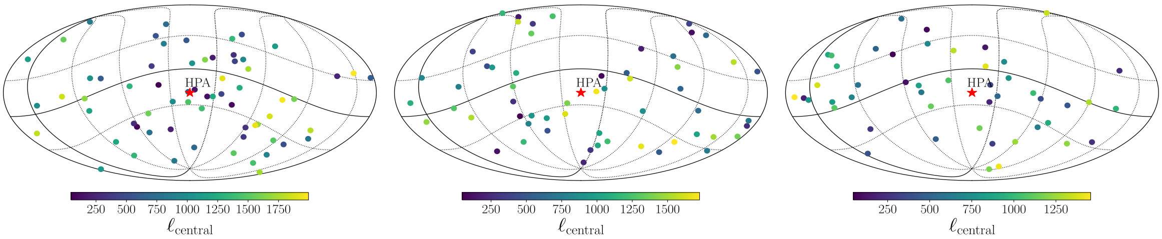

FIGURE 6.7: Dipole directions for the binned cross power spectrum ($\Delta\ell = 30$) between the detector A and B splits of the `Sevem` PR4 cleaned maps, from ℓ = 2 to ℓ = 2011 (TT, left), 1741 (TE, middle), and 1471 (EE, right). Dipoles are fitted to $N_{\text{side}} = 1$ maps, with each pixel assigned the value of the local cross power spectrum in the corresponding bin, using inverse-variance weighting. The direction for a specific multipole bin is coloured according to the central value of the bin, as shown in the colour bar. The maps are rotated in such a way that the centre is located in (l, b) = (205, -20) in Galactic coordinates, which is the preferred direction for the HPA in temperature data (marked with a red star). Graticule shows the Galactic reference frame.

in the left and right panels. In the same figure, we include the expected RS curve for an isotropic field, which has been computed from random directions uniformly distributed in the sky. Additionally, the theoretical expectation, at first and second order, for the average value across the simulations is included,

$$\left\langle\sqrt{N+2x}\right\rangle \approx \sqrt{N} - \frac{1}{4N^{3/2}}\sigma_x^2 + \mathcal{O}(\left\langle x^3\right\rangle), \tag{6.3}$$

where $x = \sum_{ij}\cos\theta_{ij}$, and σ_x^2 is the variance of x. We are also assuming $< \sum_{ij}\cos\theta_{ij} >= 0$.

The significance of the temperature alignment is compatible with previous results up to $\ell_{\max} \approx 750$. From $\ell_{\max} \approx 200$ the p-value is essentially below 1% up to the maximum multipole considered. We also consider two additional cases in which we slightly modify the first bin included in the analysis. The black curve in the left panels corresponds to the case where the first bin is included, thus incorporating all the information from $\ell = 2$ to 2011. In contrast, the grey line represents the case where we exclude the first three bins. In both cases, the results appear to be robust, with the p-value remaining below 1% for most multipoles. Only two regions, around $\ell \approx 300$ and $\ell \approx 1200$, show a slight increase in the p-value. Since the p-value is a cumulative quantity, this could be due to the dipoles in these bins being oriented far from the clustering direction, effectively contributing negatively to the RS. In a previous *Planck* paper (Planck Collaboration et al., 2020d), the p-value was found to increase rapidly from $\ell_{\max} \approx 1000$. We observe a similar increase in the case where we do not apodize the mask. Figure 6.9 presents the correlation between the bins in temperature for both scenarios, with and without apodization. For small scales, we see an increase of correlations for the non-apodized case, which produces an artificial clustering in simulations, and thus reduces the significance. Additionally, we note that for $\ell_{\max} < 100$ the temperature p-values are not anomalous, which is inconsistent with the HPA reported in the analysis of large angular scales. As mentioned in Planck Collaboration et al. (2020d), this could simply be due to the high variance of the estimator in this region.

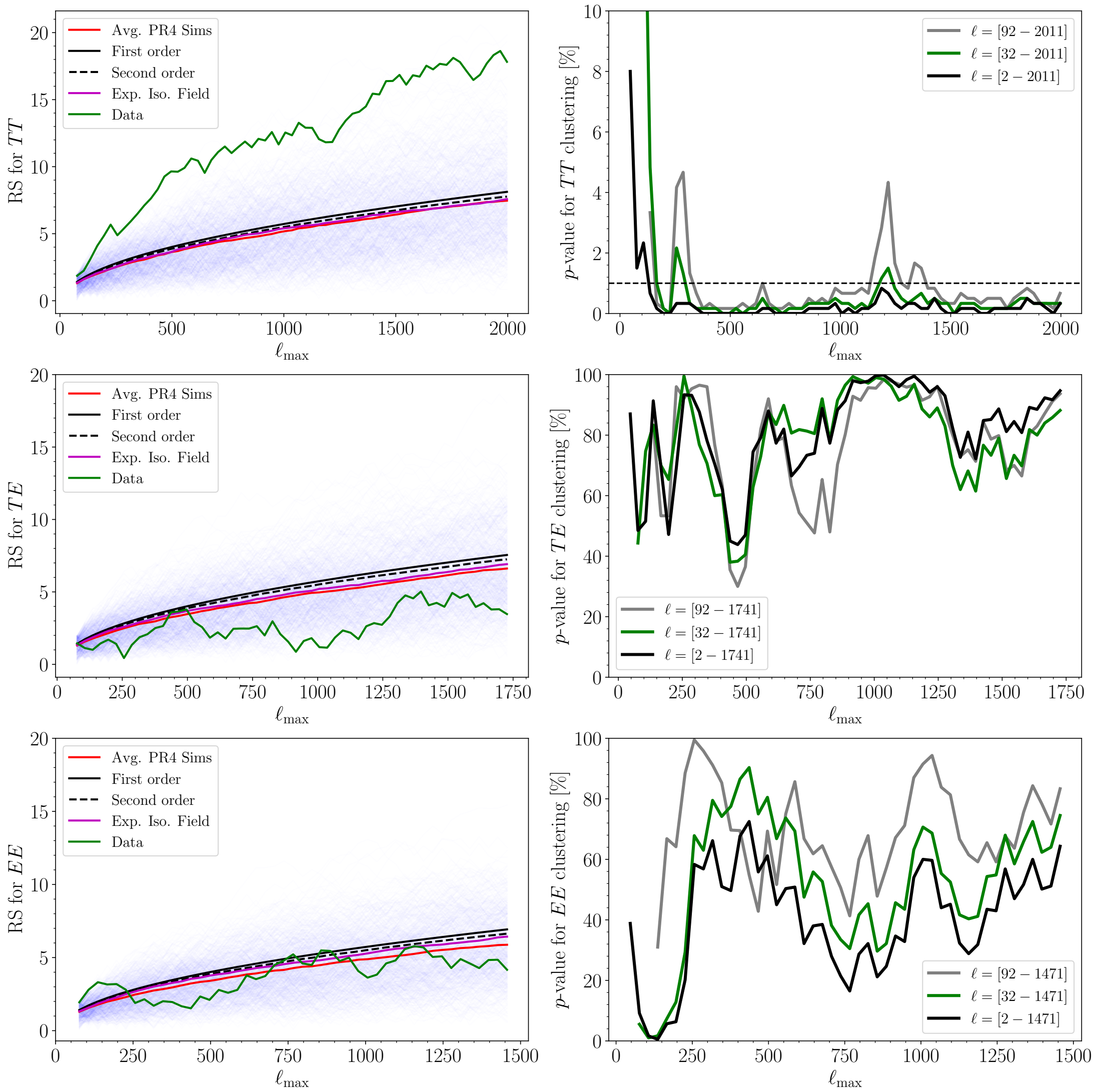

FIGURE 6.8: *Left column*: RS estimator for TT (top panel), TE (middle), and EE (lower). Blue curves correspond to each of the 600 E2E PR4 simulations, while the red curve is their average. Results for the data are shown by the green curve. The purple curve is the expected value for an isotropic field. The solid (dashed) black curve is the theoretical expectation at first (second) order. *Right column*: Derived p-values for the angular-clustering of the cross-power distribution obtained from the PR4 detector split maps as a function of $\ell_{\max}$. The p-values are derived from the fraction of E2E PR4 simulations with an RS equal to or larger than that observed in the data, hence small p-values would correspond to an unexpected alignment between dipole directions. The black curve shows the results starting from ℓ = 2, while the green and grey curves start in ℓ = 32 and ℓ = 92, respectively. For the TT case, the black dashed line represents the 1% p-value.

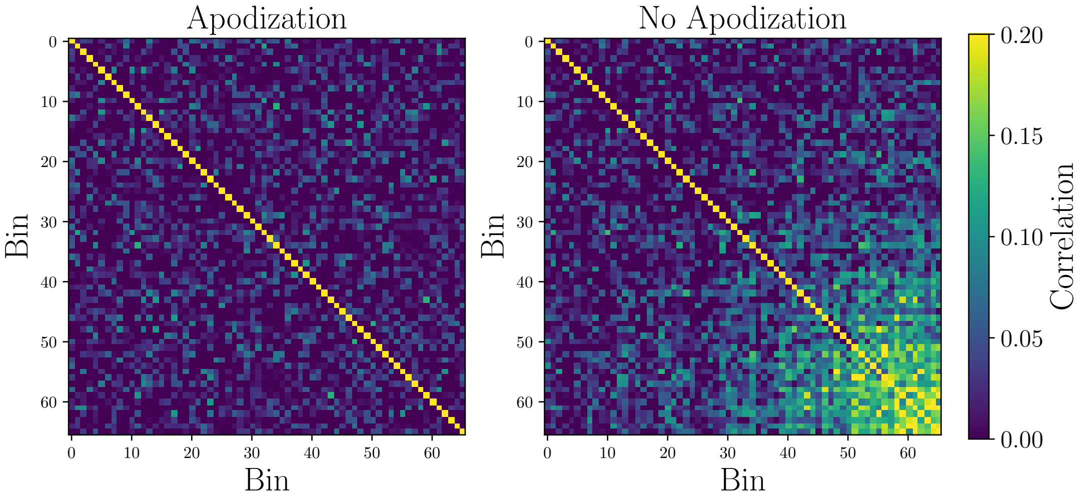

FIGURE 6.9: Correlation matrix between bins of $\Delta\ell$ = 30 for the temperature power spectrum. The left and right panels show the results for the apodization (0.3°) and no apodization cases, respectively.

An important step during the dipole fitting is the subtraction of the mean field at each bandpower. We realized that the mean fields exhibit a dipolar feature pointing in the direction of the CMB dipole. This could be due to the Doppler boosting effect (Planck Collaboration et al., 2016b; Challinor & van Leeuwen, 2002), which is also simulated in the FFP10 realisations. We assessed the impact of the mean field subtraction on the recovered dipole directions. The distribution of dipoles exhibits significant non-uniformity when the mean field is not subtracted prior to the fitting procedure. However, this effect is corrected once the mean field is removed.

It is evident from the right panel of Figure 6.8 that some p-values for the TE spectrum are close to 100%. This observation was also reported in the previous *Planck* analysis. A high p-value means a low value for the RS statistic, which could also be anomalous. To assess whether such low values are unusual, we examine the maximum p-value of the RS statistic for the TE data, which is 99.5%, and scan over the simulations to see in how many of them we are able to find such high p-values[7], but not restricting ourselves to any $\ell_{\max}$ range (to take into account the look-elsewhere effects). In fact, we find that the maximum p-value in the data is exceeded in 16% of the simulations. In other words, 16% of simulations show at least one p-value above 99.5%. Furthermore, we observe in the data that in the multipole range $\ell_{\max}$ between 900 and 1050, p-values consistently fall above 95%. We analyse the simulations finding that approximately 6% of them exhibit a range of 7 consecutive bins with p-values above the 95% threshold. These results indicate that neither of these features is statistically anomalous, and can be explained by look-elsewhere effect.

For the EE polarisation signal, the p-value is at the 1% level only for a single bin ($\ell_{\max} =$

---

[7] Note that the p-value of the simulations is computed by removing the given simulation from the set and using the remaining 599 as the reference.

$62-91$). For the case where the first bin is included the p-value reaches a minimum p-value of 0.5% again for a single bin ($\ell_{\max} = 92 - 121$). However, we do not see any anomalous behaviour in that range for either TT or TE. Considering that in PR3 the p-value in the E-modes remained close to 1% for several consecutive bins, whereas here it is observed in a single bin, which is not statistically significant, most likely the PR3 results were more affected by systematics at low-ℓ in polarisation. We conclude that the hint of an anomaly in the E modes observed in PR3 has disappeared in the PR4 data.

Following the analysis performed in Planck Collaboration et al. (2020d), we also tested whether the directions of the EE dipoles are aligned with the TT dipoles. Here we made a small change in the statistic described in Section 6.1. In order to avoid using the information from TT and EE alone, we simply use the mean of the cosine of the angles between all pairs of dipoles, where one is TT and the other one is EE, i.e. $\left\langle \sum_{ij} \cos\theta_{ij} \right\rangle$, where $\theta_{ij} = v_i^{TT} \cdot v_j^{EE}$.

Figure 6.10 shows the p-values for TT-EE alignment. The main motivation for studying such alignment in previous *Planck* works was the existence of a multipole range below $\ell_{\max} = 250$ where the p-value for both TT and EE was below 1%. In this work, such a situation does not occur, but we still find it interesting to perform the same analysis.

The black curve in Figure 6.10, corresponding to the including the first bin, exhibits a pattern similar to that shown in Figure 40 of Planck Collaboration et al. (2020d) up to $\ell_{\max} \approx 1000$. This indicates that TT and EE appear to be clustered towards a similar direction at the level of 1% over a wide range of $\ell_{\max}$. This is highly unexpected if TT and EE are completely independent, even if they were clustered individually. However, we know that a non-zero TE spectrum induces a correlation between T and E. In the *Planck* paper, this was explored with simulations to assess whether the TT-EE alignment is expected in the case where both TT and EE are individually clustered. In particular, they examine all simulations having a minimum p-value below 1% for both TT and EE in overlapping multipole ranges, and then check the correlation between TT and EE directions. They just found 2 simulations satisfying that criterion, with neither showing a high correlation between directions. Furthermore, we also explore the TE-EE alignment (magenta line in Figure 6.10), and do not find anomalous features. The results seem to lose statistical significance once the first bins are removed (see green and grey lines in Figure 6.10).

Another interesting result is shown in the right panel of Figure 6.10. In this case, only the cosines between the TT and EE vectors within the same bin are considered, that is, when i = j. No anomalous alignment behaviour is observed, as the p-value never drops below 10%. This suggests that the anomaly mainly arises due to the off-diagonal terms.

Our results show a clear anomaly in the TT clustering. On the other hand, TE and EE seem to be compatible with the *Planck* E2E PR4 simulations, and only exhibit anomalous behaviour over a very narrow multipole range, which suggests a look-elsewhere effect. Similarly, the TT-EE alignment indicates, under certain conditions, a clear anomaly with p-values below 1%. However, the interpretation is not entirely clear, as the statistical signif-

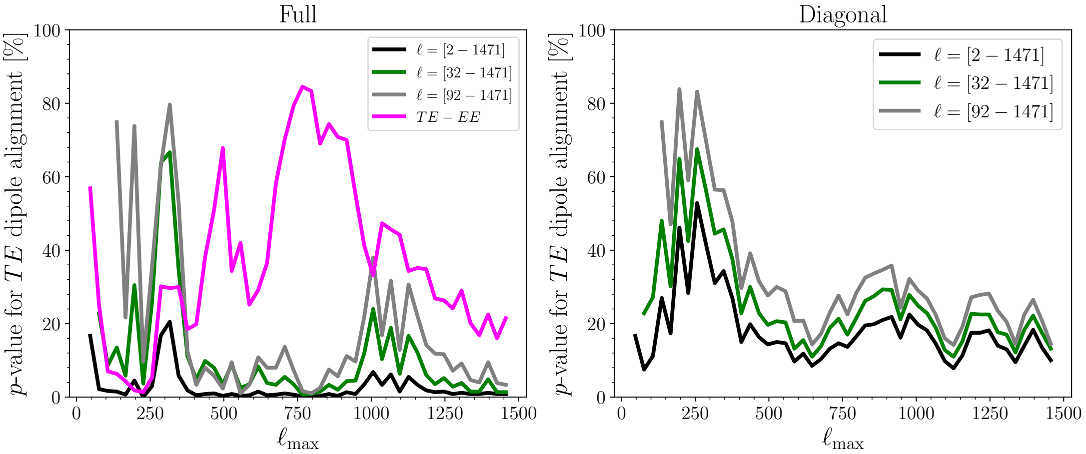

FIGURE 6.10: Derived p-values for the angular-clustering between TT and EE dipole directions. Black, green, and grey colours corresponds to the same cases as in Figure 6.9. The magenta line shows the corresponding correlation between the TE and EE dipole directions in the ℓ = [2-1471] range. The p-values are derived from the fraction of E2E PR4 simulations with a RS value equal to or larger than the one observed in the data. The results shown in the left panel are obtained by averaging the cosines between all the dipole directions (full), whereas the right panel uses only the cosines between dipoles within the same bin (diagonal).

icance seems to depend slightly on whether the first bins are included or not. Moreover, if only the cosines between dipoles within the same bin are considered, the statistical significance is greatly reduced.

6.2.2 Analysis of Cosmological Parameters

In this section, we show the results for the analysis of the cosmological parameters. We will consider as reference the case where we do not subtract the bias (see Section 6.1.3) at the power spectra level. Figures 6.11 and 6.12 show the cosmological and nuisance parameters, respectively, for patch 4 and 10. These plots provide, for the patch with the smallest sky coverage and one of those with the largest, a direct comparison between the debiased and no debiasing cases. For reference, we also include the only temperature scenario, which has been run without including the TE and EE power spectra. As mentioned in Section 6.1.3, the bias on the power spectra could have a significant impact on the cosmological parameters for large f_{sky}. However, for smaller sky fractions, the bias is subdominant compared to the statistical uncertainties. In particular, Figure 6.13 shows the distribution of the 5 cosmological parameters for all patches for the two no-debiasing cases: only temperature (top panel) and including polarisation (bottom panel). The impact of the bias on the parameters can be directly observed in these plots. The parameters are not strongly affected compared to the large error bars. Note also that most of the patches are biased in a similar way, so the bias is not expected to produce a dipolar pattern. Nevertheless, such a pattern would be removed in the dipole fitting process by subtracting the mean field. This is the main reason why the results are robust against the debiasing (see Appendix 8.4). Note also the depen-

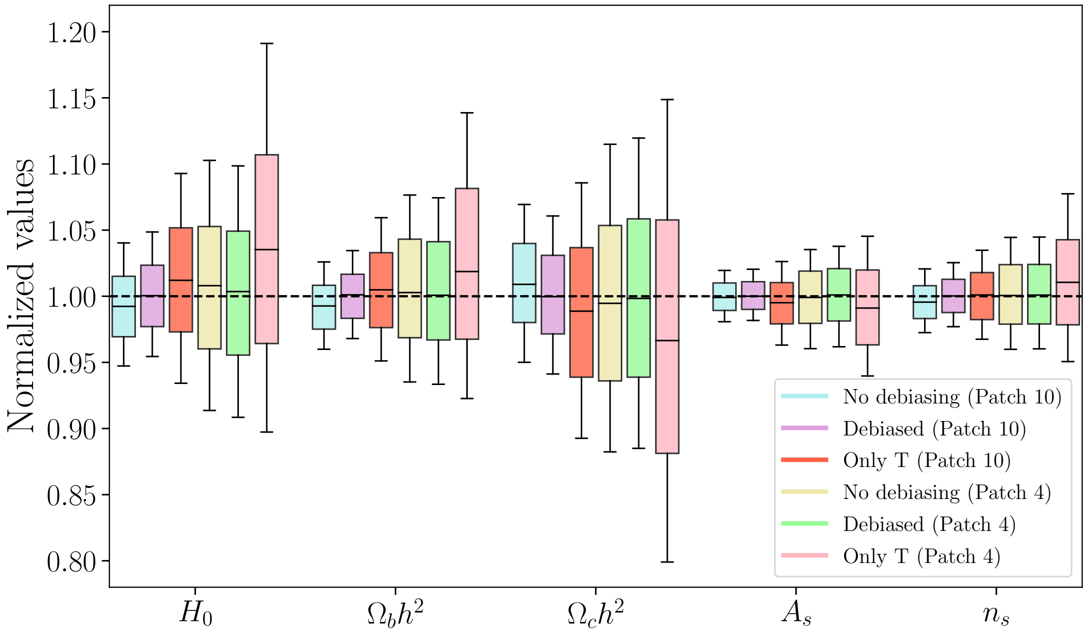

FIGURE 6.11: Distribution of the cosmological parameters for the 600 E2E PR4 simulations for two different patches, patch 4 ($f_{\rm sky} \approx 2\%$) and patch 10 ($f_{\rm sky} \approx 7\%$). For comparison three different scenarios are included: temperature only (no debiasing), including polarisation (no debiasing), and the debiased case (corrected for the bias in temperature and polarisation). The distributions are normalized to the input values. The boxes represent 68% of the probability, while the large error bars include 95.4%.

dence of the error bars with the patch index and sky fraction (see Table 6.1 and Figure 6.2). It is clear from Figure 6.11 that the bias only produces a shift of the distribution while, as expected, a lower sky fraction leads to a broadening of it. In Appendix 8.4 we perform a set of robustness tests by running the pipeline for different analysis choices, including the debiased cases and cuts in the $\ell_{\rm min}$ and $\ell_{\rm max}$, showing that the results are in general quite stable. Note in Figure 6.12 that, in the no debiasing case, the nuisance parameters tend to have a negative average value. Since point sources have not been simulated, these parameters should be zero. This reflects how they attempt to absorb the effect of the bias.

An intriguing aspect of the analysis is how Doppler boosting affects the cosmological parameters, particularly A_s. Before fitting a dipole to the parameter maps, we calculate the mean field and standard deviation for each parameter using all simulations, following the same procedure used for the power spectra in the angular clustering analysis. Notably, the mean field for A_s exhibits a dipolar feature aligned with the CMB dipole, as shown in the left panel of Figure 6.14. We then fit dipoles to the mean fields. The right panel of Figure 6.14 displays the fitted dipoles along with the directions of the dipoles obtained from the mean fields of each bandpower of the power spectrum (red dots). For reference, we also include the direction of the CMB dipole from Planck Collaboration et al. (2020f). A clear alignment is observed between the A_s mean field dipole, the CMB dipole, and the

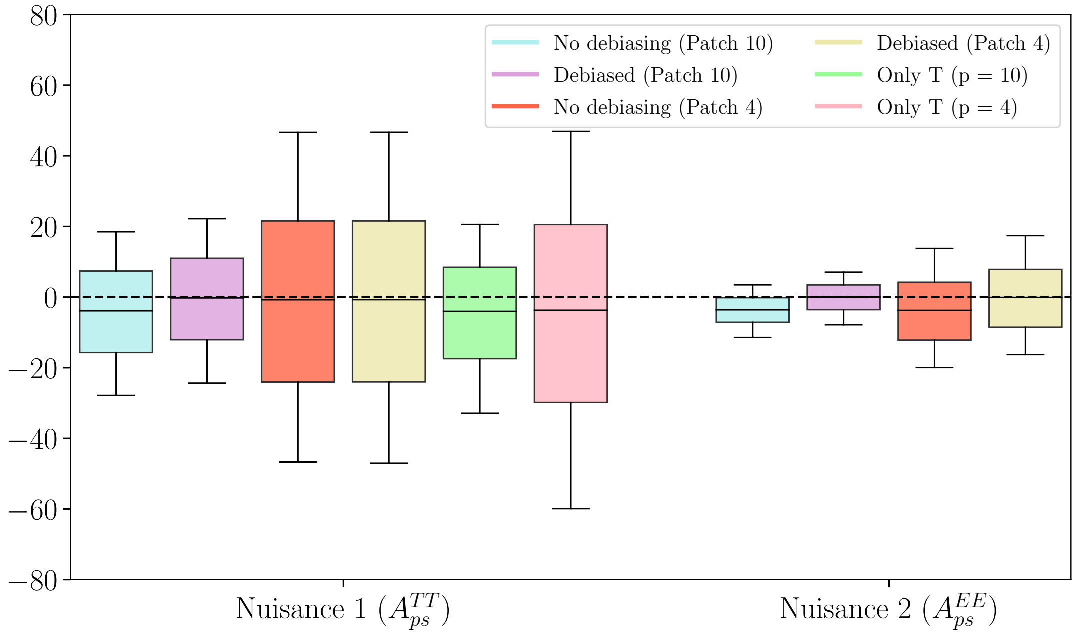

FIGURE 6.12: Same as Figure 6.11 for the nuisance parameters.

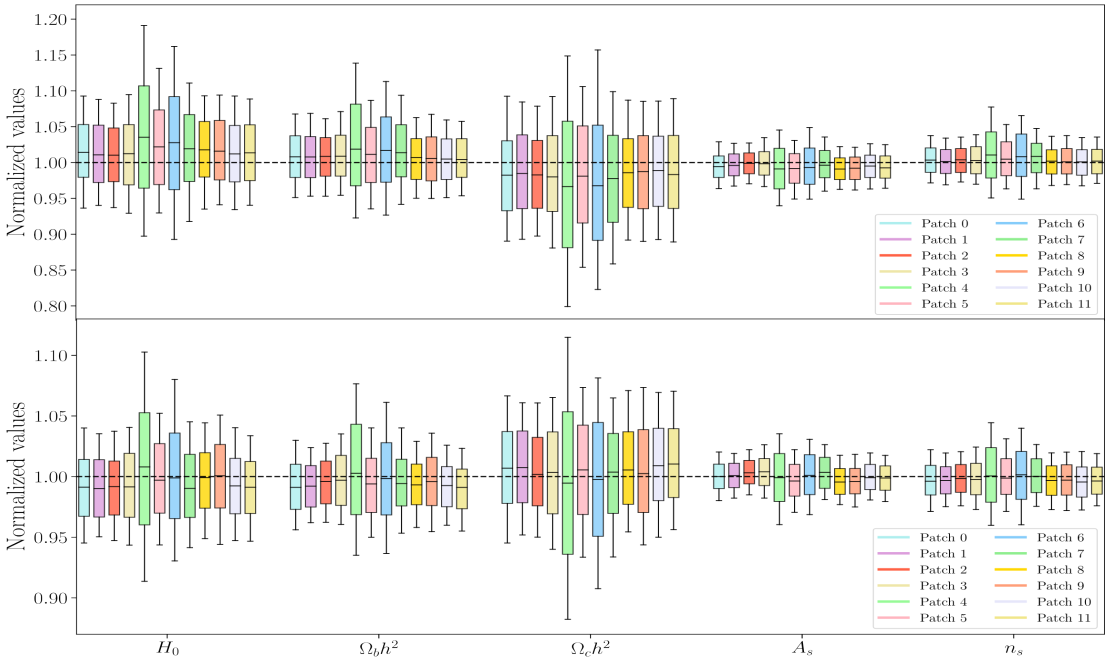

FIGURE 6.13: Distribution of the cosmological parameters computed from 600 E2E PR4 simulations for the 12 patches in the temperature only case (top panel) and including polarisation (bottom panel). In both cases, no-debiasing results are shown. The distributions are normalized to the input values. The boxes represent 68% of the probability, while the large error bars include 95.4%.

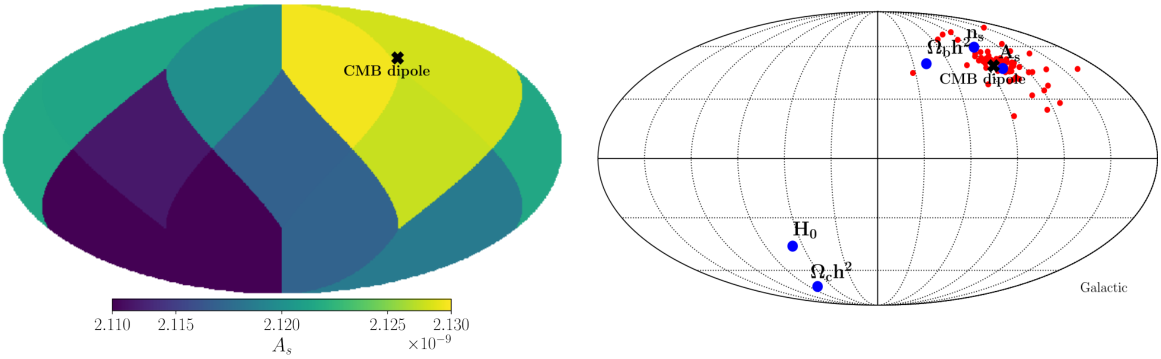

FIGURE 6.14: *Left panel*: Mean field for the A_s cosmological parameter, obtained by averaging the results from the 600 E2E PR4 simulations for each of the 12 patches. *Right panel*: Directions of the dipoles fitted on the mean field for each of the five cosmological parameters (blue dots). The dipole directions for the mean fields of the TT power spectrum bandpowers are also shown (red dots). In both panels the black cross corresponds to the CMB dipole direction.

dipoles of the bandpower mean fields. However, this effect is mitigated by subtracting the mean fields prior to the fit. As shown later, the directions in the data are not aligned with the CMB dipole; in particular, the direction for A_s is nearly 55 degrees away. This indicates that Doppler boosting does not influence our estimator. Furthermore, after debiasing, the mean field for A_s no longer displays a dipolar pattern, suggesting that Doppler boosting is a contributing factor to the observed bias.

Figure 6.15 presents the distribution of dipolar amplitudes for the five cosmological parameters. The green distribution is obtained including polarisation, while the blue distribution is found using only temperature information. The black vertical lines represent the data, the solid line for the temperature plus polarisation case, and dashed line for temperature only. It also includes the probability-to-exceed (PTE) values, representing the percentage of simulations with an amplitude equal to or larger than that observed in the data. In particular, for the temperature plus polarisation case, the amplitudes for H_0, $\Omega_b h^2$, $\Omega_c h^2$, and n_s are fully consistent with the ΛCDM predictions. However, there are only 5 simulations from 600 with a larger amplitude for the A_s dipole than observed in the data. For the temperature-only analysis, none of the five parameters exhibits an anomalous amplitude. Two effects contribute to this difference. First, the distribution of amplitudes is broader for the temperature-only scenario, consistent with the increased uncertainty in the parameters. Second, the amplitude observed in the data is slightly larger, a trend common to all parameters except $\Omega_b h^2$.

The right panel in Figure 6.16 shows the dipole directions in the data for all the cosmological parameters. In particular, A_s, which is the only one showing an anomalous amplitude, is closest to the HPA. The HPA directions used in this analysis are those reported in Gimeno-Amo et al. (2023). For reference, we present both HPA directions: one inferred from temperature data alone and the other derived from polarisation E-modes alone. We also show the direction for the A_s dipole without the mean field correction. Even if this is

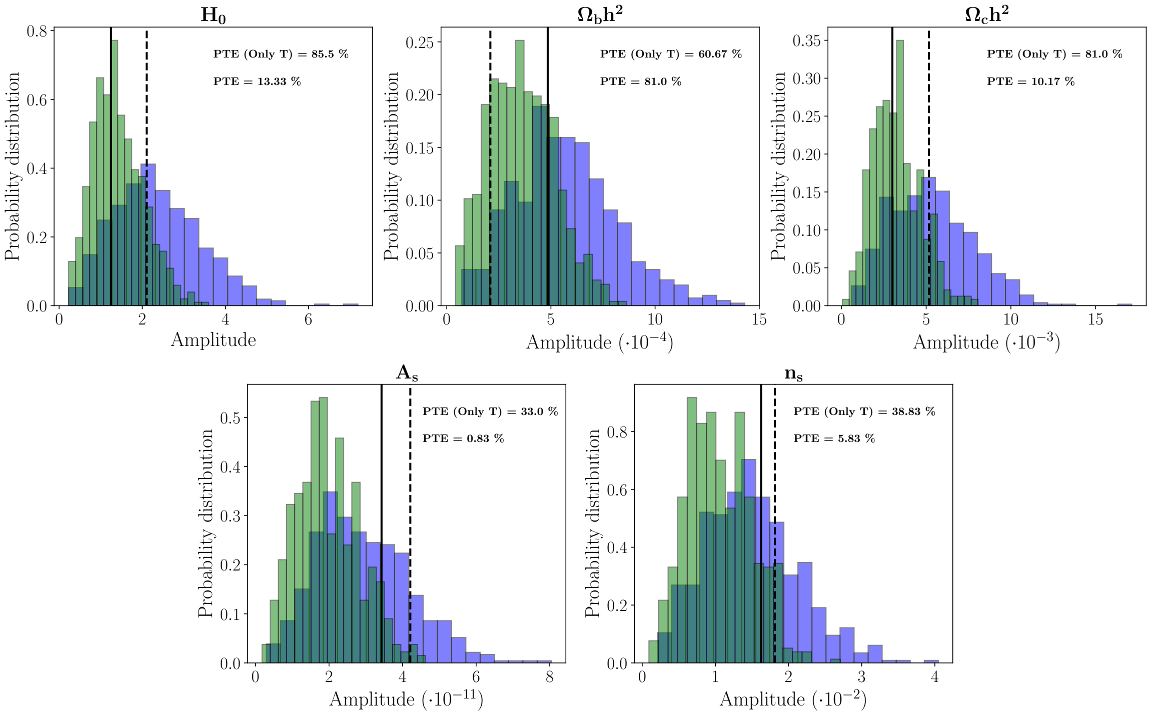

FIGURE 6.15: Probability distribution of dipole amplitudes for the 600 E2E PR4 simulations for the five cosmological parameters. Blue distribution represents temperature-only case, while green distribution includes also the polarisation information. The black lines correspond to the values observed in the data. Solid line for temperature-only case and dash line for temperature plus polarisation. In each case the probability to exceed is also provided.

displaced slightly towards the CMB dipole direction, it is not significantly affected by the Doppler boosting effect, contrary to what happens with the simulations (left panel). Moreover, the PTE is still below 1%. This means that the dipole exhibited by the A_s parameter is stronger than the one produced by the Doppler boosting.

Figure 6.17 shows the directions in data for all the cosmological parameters for both scenarios, temperature only (red dots) and including polarisation (blue dots). An interesting fact is that even if the amplitude of A_s is no longer anomalous, the direction is still close to the HPA direction. In particular, the distance between the A_s dipole direction observed in the temperature-only fits and the HPA for temperature data alone is quite similar to the distance between the A_s dipole inferred including polarisation and the HPA for the polarisation E-mode signal alone. Additionally, we see that the directions for temperature only are not far away from the directions obtained including polarisation.

It is well-established that A_s and τ are degenerate, and the uncertainty in A_s increases when τ is not fixed. As a result, the analysis presented here examines the conditioned distribution for A_s, rather than the distribution marginalized over the τ value. However, we are unable to quantify by how much our results, particularly the PTE for A_s, might change if τ is not fixed. The primary limitation is that we cannot extract meaningful information

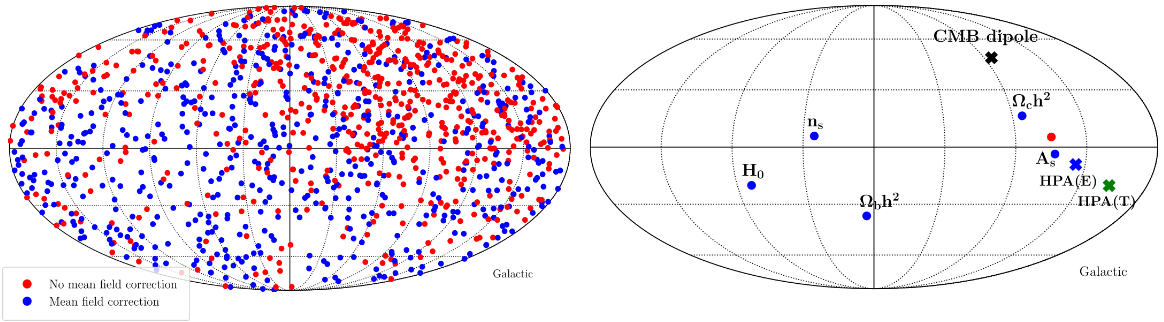

FIGURE 6.16: *Left panel*: Dipole directions of the A_s parameter for the 600 E2E PR4 simulations before (red dots) and after (blue dots) mean field subtraction. *Right panel*: Dipole directions (blue dots) of the five cosmological parameters observed in the combination of temperature and polarisation data. The dipole direction for A_s before mean field subtraction is also shown (red dot). The black cross corresponds to the CMB dipole direction, and the blue and green crosses represent the HPA directions measured in T and polarisation E-modes, respectively.

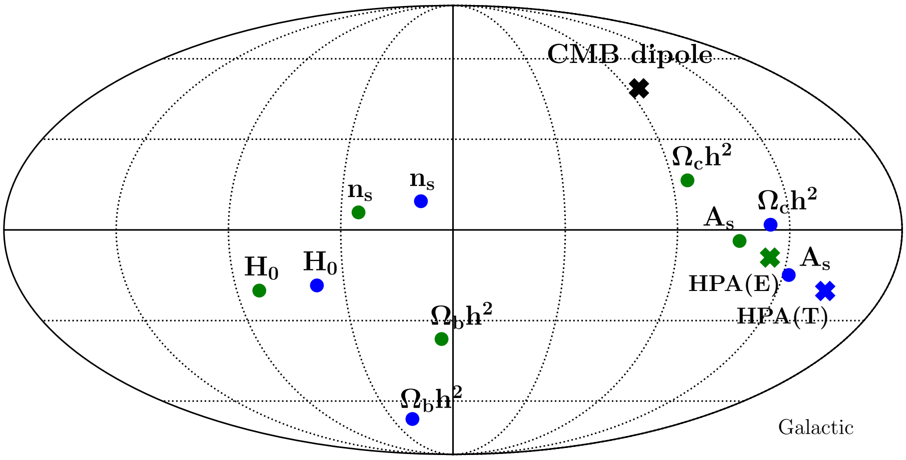

FIGURE 6.17: Dipole directions of the five cosmological parameters observed in the data. Green dots correspond to temperature plus polarisation case, while blue dots are from temperature-only scenario. The black cross corresponds to the CMB dipole direction, and the blue and red cross represent the HPA directions measured in T and polarisation E-modes, respectively.

about τ from such small sky patches.

To address this, we performed a set of tests. First, we run the pipeline with τ as a free parameter, imposing a reasonable bound in `iMinuit` between 0.03 and 0.09, i.e., ten times the *Planck* error bar on τ. In almost all simulations, the τ values clustered near the boundaries of the allowed range, reflecting that no information can be determined about τ alone. However, the distribution of $A_s e^{-2\tau}$ seems to be similar to that obtained by fixing τ. In other words, the data can constrain the combination fairly well regardless of whether τ is fixed. We also run this case for the data, finding similar behaviour. The final p-value for the combination $A_s e^{-2\tau}$ is at the level of 2%. Therefore, it seems that there is an anomaly that is best captured by the combination $A_s e^{-2\tau}$, but once τ is fixed, it is propagated to A_s, because both quantities are related by a constant.

Additionally, we run the pipeline with a Gaussian prior on τ. Unlike the Bayesian approach, the minimizer can not deal with the prior and consistently converged to the maximum of it, yielding results essentially identical to those obtained with fixed τ. Finally, we consider the case in which the first bin, covering multipoles between 2 and 31, is included in the likelihood. Although this bin carries information about τ, the inferred values are still clustered near the bounds. This is expected, given the small $f_{\rm sky}$, which produces large error bars. Moreover, on these scales, the Gaussian approximation on the likelihood is no longer a good approximation, even if the bin size makes it more Gaussian following the central limit theorem. For all of these reasons, we decided to focus on the case where τ is fixed.

7

Conclusions

> *"What we call the beginning is often the end. And to make an end is to make a beginning. The end is where we start from."*
>
> T.S. Eliot

The Cosmic Microwave Background has provided humanity with a deep understanding on the Universe in which we live. While the ΛCDM model offers a remarkably successful framework that accurately describes a wide range of cosmological observations, there are still numerous questions to be addressed, and some measurements that lack a satisfactory explanation. The nature of dark energy and dark matter, even if they account for approximately 95% of the total energy content of the Universe, is still unknown. Moreover, the current expansion rate, H_0, inferred from early- and late-time probes exhibits a tension exceeding 5σ. The origin of a set of unexpected features observed in temperature large angular scales, the so-called CMB anomalies, remains unexplained. While the statistical significance of these anomalies is still an open discussion, there are no compelling models that could reconcile the observations with the standard modelThis PhD thesis aims to shed light on some of these anomalies by analysing them in the most recent *Planck* Data Release, which has effectively reduce the systematics effects on large angular scales in polarisation. In particular, two chapters are dedicated to the study of the statistical isotropy, while one chapter focuses on the methodological development of an inpainting technique.

In the next sections, the conclusions of Chapters 4, 5, and 6 are presented. At the end, a summary of the main results and a comment on the future work is provided.

7.1 Chapter 4

In this chapter, we have applied the local-variance estimator to the latest *Planck* release data set (PR4) in both temperature and polarisation, performing a frequentist analysis of the isotropy assumption. In temperature, we have not found any simulation from a set of 600 with an amplitude as large as the one observed in the data for 4°, 6°, and 8° discs. This means that the data exhibits a dipolar-like behaviour across the sky at a level below 0.17% for that range of angular scales. We have performed the analysis at two different resolutions, $N_{\rm side} = 64$ and $N_{\rm side} = 2048$. For low resolution maps the direction points towards (ℓ, b) = (208°, -15°), while for full resolution the direction is (ℓ, b) = (205°, -20°). These results are in good agreement with previous reports, and cannot be explained by residual systematics since a similar detection is found in WMAP.

Regarding polarisation, we have implemented an alternative inpainting approach in order to minimize some effects that affect the reconstruction of the E-mode map from the measured Stokes parameters. This technique works optimally under certain conditions, as shown in Chapter 5, that are not completely satisfied in the *Planck* data. The main problem is the low number of realistic noise simulations, which does not allow to compute in an accurate way the noise covariance matrix. However, we have checked in this chapter that the inpainting is not biasing the parameters of the analysis and that also improves significantly the results compared with a simple masking approach.

First, we have applied the inpainting technique to the PR3 data set and obtained results in good agreement with those reported by the *Planck* collaboration using a different inpainting algorithm. The small differences found can be explained by the large intrinsic dispersion of the estimated dipole direction, as well as to differences in the considered mask and inpainting algorithm. In particular, we found that both the p-values and the directions depend on the choice of mask, with the exception of `Sevem`, whose p-value is always below 1%.

For PR4, we performed the analysis only on the `Sevem` cleaned maps, since it is the only method providing both cleaned maps and simulations that are suitable for the analysis. We get a p-value that lies between $1\% - 3\%$ for sky fractions in the range from 67% to 59.4%. For smaller $f_{\rm sky}$ the p-value increases rapidly, getting up to 7.4% for $f_{\rm sky}$ = 57%. If the asymmetry is of cosmological origin, this behaviour is somehow expected due to the loss of information when the number of available pixels is reduced. However, we cannot discard that this effect is related to the reduction of foreground contamination in the CMB map when increasing the mask, as one would expect if residuals were the cause of the observed asymmetry. Finally, the results are probably even more sensitive to the sky fraction and geometry of the mask, due to the fact that the estimated direction of the dipole is close to its boundary.

It is interesting to note that the amplitude of the dipolar modulation for intensity ($\sim 7\%$)

and polarisation ($\sim 9\%$) are of similar level and point in a similar direction. If the modulation in intensity were simply a statistical fluke such that the standard model still holds, we would expect a similar pattern in the E-mode map, taking into account the correlation between T and E. However, since this correlation is small (typically around 10 per cent), a much lower dipole amplitude in polarisation would be expected and, therefore, this seems inconsistent with the values found. Even more, we have checked with E-mode maps constrained to intensity modulated simulations, that this correlation is expected to increase the p-value of the alignment by only about 0.5% when realistic E2E noise is included. In any case, given the large uncertainties associated to the estimated amplitudes, improved polarisation data is needed before any further conclusion can be established.

Regarding the difference between the p-values for PR3 (0.22%) and PR4 (2.8% for our reference case), we have verified that both values are compatible when accounting for differences between data sets, including the number of simulations, mask, inpainting, noise properties, and transfer function. One important consideration is the mask. Although we fix the same threshold in the maximum error of the reconstructed E-modes, this results in a different sky fraction for PR3 and PR4, which can influence the estimated p-values. In any case, other factors intrinsic to the data could also contribute to these results. For instance, if the anomaly were due to systematic effects, a larger p-value would be expected in PR4, given the improvement in this regard in the data. However, if the anomaly were of cosmological origin, the increase in the p-value could also be due to the presence of the transfer function, which suppresses the power at low multipoles in the PR4 polarisation data. Therefore, no clear conclusion can be drawn, except that a modest detection of asymmetry is still present in the E-mode polarisation maps.

Unfortunately, the sensitivity of the data is still not sufficiently good to obtain robust results. In particular, the model-dependent test presented in this chapter shows that the dipolar modulation model parameters exhibit large uncertainties. Using the *Planck* polarisation E2E simulations, we find that a detection with a p-value below 1% would be possible in 95% of cases only if the amplitude of the modulation were at the level of 16%, which is well above the observed value. Therefore, a significant improvement in the sensitivity and control of systematics in the polarisation data at large angular scales is needed before robust results can be achieved in relation to a possible violation of isotropy. Future experiments, such as LiteBIRD, that will be able of measuring cosmic variance limited E-modes over a very large fraction of the sky, are expected to provide key information on the HPA.

7.2 Chapter 5

In this chapter, we have presented an inpainting technique based on Gaussian Constrained Realisations, that can be applied to CMB temperature and polarisation data. The algorithm uses the Cholesky decomposition to sample from a conditional probability distribution. We have also developed a `Python` user-friendly package, `CMB-PAInT`, which is publicly available at `https://github.com/ChristianGim/CMB-PAInT`. We used this package to obtain all the inpainted realisations of this chapter.

In order to asses the performance of the methodology, a series of tests have been done in two different scenarios: (1) CMB signal only, and (2) CMB signal with semi-realistic noise simulations based on *Planck* Data Release 4. In particular, we check that the constrained part of the inpainted maps is consistent with that expected from the model. We have also studied the one-dimensional probability distribution of T, Q, and U, as well as those of the E- and B-modes. In both scenarios, they agree well with the input values within the expected errors. Our methodology is also able to reconstruct accurately systematics and noise from an anisotropic field if they are included in the covariance matrix. However, this requires that all the correlations between pixels of the anisotropic field are perfectly characterised. Otherwise, artifacts would appear due to a mismatch between the statistical properties of the pixels outside the mask and the assumed model. This can be verified by inspecting the intermediate z variables, which should follow a normal distribution if everything is consistent.

For certain applications, further studies may be needed in order to minimise the presence of prominent systematics in the inpainted region while, at the same time, not introducing a statistical mismatch between observed pixels and the covariance matrix. In any case, the optimal strategy will depend on the nature of the data and estimator to be studied.

The most interesting tests involve the E- and B-mode reconstruction and the corresponding power spectra. Both tests show that we can remove well the E-to-B leakage. For the first scenario, at the map level, we are able to reconstruct the E-mode map with a relative error below 5% for a sky fraction of 71.7 per cent, covering almost all the sky outside the polarisation confidence mask (which allows 72.36 per cent of the pixels). In the case of the B-mode, the relative error is around 10% for $f_{\rm sky} = 45\%$, due to its weaker signal. At the power spectra level, we reproduce the input TT, TE, EE, and BB power spectra up to $\ell_{max} = 192$, covering the full range of the reionization and the recombination peaks of the B-mode. The residuals between the input and reconstructed spectra are consistent with zero within the dispersion obtained from the 1200 inpainted realisations. In the second scenario, the method reproduces the input power spectra that were strongly affected by systematics and noise. Furthermore, we perform a comparison between our methodology and other techniques (diffuse inpainting, masking and the NaMaster pure-B approach), finding that the GCR was the only method able to recover the input B-mode spectrum for all the considered multipole range.

This inpainting approach is limited to low resolution maps due to computational memory requirements, but it is enough to target the polarisation largest scales, which is the main goal of future observations, searching for the primordial B-mode. In view of the present results, we believe that this will be a useful and powerful algorithm for the analyses of future

CMB experiments, such as LiteBIRD (LiteBIRD Collaboration et al., 2023).

7.3 Chapter 6

In this chapter, we perform two complementary analyses on the *Planck* PR4 CMB maps (in particular, detector splits), which were cleaned using the `Sevem` component separation method, in order to test the statistical isotropy of the Universe. Our approach employs well established techniques for both power spectrum estimation (`MASTER`) and cosmological parameter fitting (`iMinuit`). In both analyses, we use the binned TT, TE, and EE power spectra computed over 12 independent sky regions, corresponding to the 12 pixels of the $N_{\rm nside} = 1$ `HEALPix` scheme. These patches overlap with the *Planck* confidence mask, so the effective $f_{\rm sky}$ for each region varies between 2% and 8%.

The first analysis focuses on the angular clustering feature. Previous *Planck* releases reported an unexpected alignment in directions derived from temperature power distribution maps over a wide range of angular scales, up to $\ell \approx 1000$. We confirm this alignment over a broader multipole range. The p-value, defined as the fraction of simulations exhibiting a Rayleigh statistic greater than that observed in the data, is below 1% for multipoles between $\ell = 200$ and 2000. We suggest that the apparent discrepancy with earlier results for $\ell > 1000$ may be due to artificial clustering caused by the lack of apodization in the previous analysis, which introduces correlations between adjacent multipole bins. Additionally, the hint of an anomaly in the E-modes that was present in PR3 disappears, with only a couple of bins showing p-values below 1%. In particular, this occurs in the same multipole range where TT and the alignment between TT and EE begins to exhibit anomalous behaviour. However, the interpretation remains unclear, as removing the first bin or considering only the cosines between dipoles within the same bin reduces the statistical significance. Given that the anomalous behaviour in the E-modes takes place within a narrow multipole range, it is most likely a consequence of the "look elsewhere" effect. A similar reasoning can be applied to a few bins close to 100% for TE.

In our second analysis, we examine the potential presence of dipolar variations in the cosmological parameters across the sky. Using `iMinuit`, we maximise the likelihood in each of the 12 independent patches using the previously computed binned TT, TE, and EE binned power spectra. Given the limited sky fraction available in each of these patches, we can not access the large scale E-modes and τ needs to be fixed. Fitting a dipole on the resulting maps reveals that all parameters, except for A_s, are consistent with the standard cosmological model. A hint of anomalous behaviour is detected in A_s, with only 5 out of 600 simulations showing an amplitude as extreme as that observed in the data. This anomaly appears to be associated with the combination of A_s and τ, $A_s e^{-2\tau}$, as it is the quantity best constrained by the data. Once τ is fixed, the anomaly propagates to A_s. The direction of the dipole is close to that of the hemispherical power asymmetry, and it is also located in the region where the intensity power spectrum bandpower dipoles are clustered, suggesting a potential link between these anomalies. Although the anomaly remains robust to variations in the choice of ℓ_{min} and ℓ_{max} (see Appendix 8.4), it disappears when the TE and EE spectra are excluded.

An additional outcome validating our pipeline is the successful detection of the Doppler boosting effect in simulations, which is observed both at the power spectrum level and in the spatial distribution of A_s parameter. This effect can be effectively encoded and subtracted in the mean field prior to dipole fitting, and it is unlikely to be the origin of the anomaly given that the direction observed in A_s is approximately 55 degrees away from the CMB dipole direction.

Our results do not agree with some previous analyses that claim a strong evidence for a violation of the cosmological principle of isotropy, as indicated by variations in cosmological parameters across the CMB sky. This may be attributed to methodological differences. In particular, our analysis uses completely independent sky regions and relies on covariances fully derived from the *Planck* provided end-to-end simulations. Finally, the evidence for an anomaly in $A_s e^{-2\tau}$ combination is modest and not entirely conclusive. Future analyses with improved polarisation data will be crucial to further clarify these findings.

7.4 Summary and future work

An HPA is confirmed in the temperature data from PR4 using the local-variance estimator, with a p-value below 1%. After applying an inpainting technique based on Gaussian constrained realisations to the Stokes Q and U parameters, the reconstructed E-mode map exhibits a similar asymmetry at the $1\% - 3\%$ level, with an axis closely aligned with that found in temperature. Interestingly, by fitting a dipole to the spatial distribution of the power in each bandpower of the temperature angular power spectrum, we find an anomalous clustering of preferred directions, which also align remarkably well with the HPA axis. No analogous clustering is observed in the polarisation data. Motivated by these findings, we explored dipolar signatures in the spatial distribution of cosmological parameters. Most parameters are consistent with ΛCDM expectations, except for A_s, which shows an anomalous dipole amplitude at the $\sim 1\%$ level, with a direction again closely aligned with the HPA. Notably, Mukherjee et al. (2016) identified A_s and n_s as the parameters most susceptible to acquiring spatial dependence in the presence of a cosmic hemispherical asymmetry, in agreements with our results.

While there is no doubt about the observed asymmetry in temperature in Chapters 4 and 6, the large uncertainties in the polarisation E-mode results prevent any robust conclusive statement. The origin of these uncertainties are both the large-scale systematic effects present in the *Planck* data and the low signal-to-noise ratio.

Since the *Planck* temperature measurements are already cosmic variance limited up to

$\ell \sim 1600$, future temperature data are not expected to significantly impact the study of CMB anomalies, except potentially through an increased sky fraction. However, a substantial improvement is expected in polarisation analyses. A natural and promising continuation of this work would be to apply the same methodology to the forthcoming high-precision polarisation data.

8

Conclusiones

> *"Lo importante es no dejar de hacerse preguntas."*
>
> Albert Einstein

El Fondo Cósmico de Microondas ha dotado a la humanidad con un gran conocimiento sobre el Universo en el que vivimos. El modelo ΛCDM ofrece un marco teórico que describe de manera precisa las observaciones. A pesar de ello, todavía existen un gran número de incógnitas que carecen de una explicación satisfactoria. La materia y energía oscuras son aún desconocidas a pesar de contribuir al 95% del contenido total de energía del Universo. Además, existe una discrepancia entre las estimaciones de la constante de Hubble obtenidas a partir de observaciones del Universo temprano y aquellas derivadas del Universo tardío. El origen de las anomalías a gran escala del FCM carece de explicación. Si bien la significancia estadística de estas anomalías sigue siendo objeto de debate, no existen modelos convincentes, basados en física fundamental, que puedan reconciliar las observaciones con el modelo estándar. Esta tesis doctoral tiene como objetivo arrojar luz sobre algunas de estas anomalías mediante su estudio en los datos más recientes de *Planck*, los cuales han sido procesados de tal forma que se han reducido los efectos sistemáticos a grandes escalas angulares en polarización. En particular, dos capítulos de la presente tesis están dedicados a estudios de isotropía estadística, y uno dedicado al desarrollo metodológico de una técnica de *inpainting*.

Las siguiente secciones presentan las conclusiones de los capítulos 4, 5 y 6. Al final, se incluye un breve resumen de los principales resultados y se comentan las perspectivas de trabajo futuro.

8.1 Capítulo 4

En este capítulo hemos realizado un análisis frecuentista de la hipótesis de isotropía estadística aplicando el estimador de varianza local al conjunto de datos más reciente de *Planck* (PR4), tanto en temperatura como en polarización. En el caso de temperatura, ninguna de las 600 simulaciones presenta una amplitud del dipolo de varianza local tan elevada como la observada en los datos para discos de 4°, 6° y 8°, lo que indica un comportamiento de tipo dipolar con un *p-value* inferior al 0.17% en dicho rango de escalas angulares. El análisis se ha realizado para dos resoluciones, $N_{\rm side} = 64$ y $N_{\rm side} = 2048$. En los mapas de baja resolución, la dirección del dipolo se orienta hacía $(\ell, b) = (208^\circ, -15^\circ)$, mientras que en los mapas de máxima resolución se encuentra en $(\ell, b) = (205^\circ, -20^\circ)$. Estos resultados están de acuerdo con estudios previos y no pueden atribuirse a efectos sistemáticos residuales, dado que los datos de WMAP muestran una detección similar.

En lo que respecta a la polarización, se ha implementado una técnica alternativa de *inpainting* con el objetivo de mitigar ciertos efectos que dificultan la reconstrucción del mapa de modos E. Tal y como se detalla en el capítulo 5, esta técnica ofrece un rendimiento óptimo bajo determinadas condiciones que, sin embargo, no se cumplen en los datos de *Planck*. La principal limitación radica en la escasez de simulaciones realistas de ruido, lo que impide estimar con precisión la matriz de covarianzas del ruido. No obstante, se ha comprobado que el procedimiento de *inpainting* no introduce sesgos en los parámetros analizados y que, además, mejora de forma significativa los resultados frente a una estrategia basada en utilizar una máscara.

Primero, hemos aplicado la técnica de *inpainting* a los datos de polarización de PR3, y hemos comprobado que los resultados son compatibles con los publicados por la colaboración, a pesar de haberse empleado un algoritmo de *inpainting* distinto. Las pequeñas discrepancias observadas pueden atribuirse a la dispersión intrínseca que muestra la estimación de la dirección del dipolo, así como a las diferencias en la máscara y el metodo de *inpainting*. En particular, se ha comprobado que tanto los valores p como las direcciones obtenidas dependen de la máscara considerada, con la única excepción del caso `Sevem`, cuyo valor p se mantiene siempre por debajo del 1%.

En el caso de los datos de PR4, el análisis se ha realizado únicamente con los mapas de `Sevem`. Se obtiene un valor p comprendido entre el 1% y el 3% para fracciones del cielo entre el 67% y el 59.4%. Para cielos más enmascarados, el valor p aumenta rápidamente, alcanzando un 7.4% cuando la fracción del cielo es del 57%. Si la asimetría tuviera un origen cosmológico, este comportamiento sería en cierto modo esperable, dado que la pérdida de píxeles observados implica un aumento en el error estadístico. No obstante, no puede descartarse que este efecto se deba, en realidad, a una menor contaminación galáctica al ampiar la máscara, como cabría esperar si la asimetría fuese causada por residuos de dicha contaminación. Por último, es probable que los resultados sean incluso más sensibles tanto a la fracción de cielo como a la geometría de la máscara, dado que la dirección estimada del dipolo se encuentra próxima al borde de la región enmascarada.

Resulta relevante señalar que la amplitud de la modulación dipolar en intensidad ($\sim 7\%$) y en polarización ($\sim 9\%$) es similar, y que ambas direcciones están casi alineadas. Si la modulación en intensidad fuera una fluctuación estadística, se esperaría un patrón similar en el mapa de modos E, debido a la correlación existente entre T y E. Sin embargo, dado que la correlación es débil (del orden del 10%), cabría esperar una amplitud significativamente menor en polarización, lo que resulta dificil de reconciliar con los valores observados. Además, al emplear mapas de modos E condicionados a simulaciones moduladas en intensidad, se ha comprobado que dicha correlación solo incrementa el valor p de la alineación en aproximadamente un 0.5% al considerar ruido realista. En cualquier caso, debido a las elevadas incertidumbres asociadas a las amplitudes, será necesario disponer de mejores datos de polarización antes de poder extraer conclusiones robustas.

En cuanto a las diferencias entre los valores p obtenidos para PR3 (0.22%) y PR4 (2.8% en nuestro caso de referencia), se ha verificado que ambos son compatibles al tener en cuenta factores como el número de simulaciones, la máscara empleada, las diferencias en las propiedades del ruido o la función de transferencia. Un aspecto a tener en cuenta es la máscara. Aunque se fije el mismo umbral en el error máximo de los modos E reconstruidos, esto da lugar a fracciones del cielo distintas en PR3 y PR4, lo que puede afectar a las estimaciones. En cualquier caso, también podrían intervenir factores intrínsecos a los propios datos. Por ejemplo, si la anomalía se debiera a efectos sistemáticos, cabría esperar un mayor valor p en PR4, dado la mejora de este tipo de efectos. Sin embargo, si el origen fuera cosmológico, el aumento del valor p podría deberse a la función de transferencia, que atenúa la señal de la gran escala en los datos de polarización de PR4. Por lo tanto, no puede extraerse una conclusión definitiva, más allá de que se mantiene una detección moderada de asimetría en los modos E de polarización.

Lamentablemente, la sensibilidad de los datos actuales aún no es suficiene para obtener resultados concluyentes. En particular, uno de los tests realizados en este capítulo muestra que los parámetros del modelo presentan grandes incertidumbres. Utilizando las simulaciones realistas de *Planck*, se ha comprobado que solo sería posible una detección con un valor p inferior al 1% en el 95% de los casos si la amplitud de la modulación fuera del orden del 16%, un valor muy superior al observado. Por tanto, se requiere una mejora significativa tanto en la sensibilidad como en el control de los efectos sistemáticos en polarización para poder establecer resultados sólidos en relación a una posible violación de isotropía. Se espera que futuras misiones como LiteBIRD, capaces de medir modos E limitados por varianza cósmica en una fracción amplia del cielo, proporcionen información clave sobre la HPA.

8.2 Capítulo 5

En este capítulo se ha presentado una técnica de *inpainting* basada en realizaciones gaussianas condicionadas, que puede ser aplicada tanto a datos de temperatura como de polarización. El algoritmo emplea la descomposición de Cholesky para rellenar los píxeles enmascarados con valores generados a partir de una distribución gaussiana condicionada al resto de los datos observados. Asimismo, se ha desarrollado un paquete de `Python`, `CMB-PAInT`, el cual está disponible en `https://github.com/ChristianGim/CMB-PAInT`. Este paquete ha sido utilizado para generar todas las realizaciones empleadas en este capítulo.

Con el fin de evaluar el rendimiento de la metodología, se han realizado una serie de pruebas en dos escenarios distintos: (1) únicamente se incluye la señal del CMB, y (2) se añade ruido semi-realista basado en las simulaciones de PR4. Se ha comprobado que la parte condicionada de los mapas con *inpainting* es coherente con lo esperado según el modelo. También se ha analizado la distribución de probabilidad unidimensional de T, Q y U, así como la de los modos E y B, observándose en ambos casos una buena concordancia con los valores de entrada dentro de los márgenes de error esperados. Nuestra metodología también permite reconstruir de forma precisa sistemáticos y ruido anisótropo y correlacionado si se incluyen estos también en la matriz de covarianzas. No obstante, esto solo es posible si se conocen con precisión todas las correlaciones entre píxeles. De no ser así, podrían aparecer artefactos debido a discrepancias entre las propiedades estadísticas de los píxeles fuera de la máscara y las asumidas por el modelo. Este posible desajuste puede verificarse inspeccionando las variables intermedias z, que bajo condiciones ideales deben seguir una distribución normal.

Para ciertas aplicaciones, podrían ser necesarios estudios adicionales con el fin de minimizar la presencia de sistemáticos prominentes en la región reconstruida mediante *inpainting* sin introducir, al mismo tiempo, una discrepancia estadística entre los píxeles observados y la matriz de covarianza. En cualquier caso, la estrategia óptima dependerá de la naturaleza de los datos y del estimador que se desee analizar.

Las pruebas más relevantes son las relacionadas con la reconstrucción de los modos E y B, así como sus respectivos espectros de potencia. En ambos casos se demuestra que es posible eliminar eficazmente la contaminación de E a B. A nivel de mapa, se logra reconstruir el mapa de modos E con un error relativo inferior al 5% para una fracción del cielo del 71.7%, practicamente tode el cielo fuera de la máscara de polarización. En el caso de los modos B, el error relativo es del orden o inferior al 10% para el 45% del cielo. A nivel de espectro de potencias, se recuperan con precisión los espectros TT, TE, EE y BB de entrada hasta $\ell = 192$, cubriendo por completo el rango correspondiente a los picos de reionización y recombinación de los modos B. Los residuos son compatibles con cero dentro de los errores esperados para 1200 realizaciónes. En el segundo escenario, el método es capaz de recuperar los espectros de entrada incluso en presencia de ruido y efectos sistemáticos. Asimsismo, se ha llevado a cabo una comparación con otras metodologías que incluyen, entre otros, el *diffuse inpainting* o el metodo de purificación de `NaMaster`. Los resultados muestran que el método basado en realizaciones condicionadas es el único que consigue recuperar el espectro de modos B en todo el rango de multipolos considerado.

Este enfoque de *inpainting* está limitado a mapas de baja resolución debido a los requisitos de memoria. No obstante, resulta suficiente para analizar las escalas más grandes en polarización, que constituye el principal objetivo de futuras observaciones orientadas a la busqueda de los modos B. A la luz de los resultados obtenidos, consideramos que este será un método útil y eficaz para los análisis de los futuros experimentos del CMB, como puede ser LiteBIRD.

8.3 Capítulo 6

Con el objetivo de poner a prueba la isotropía estadística del Universo, en este capítulo se llevan a cabo dos análisis complementarios. Para ello se utilizan los mapas de detectores de los datos más recientes de *Planck*, limpiados por `Sevem`. Empleamos técnicas conocidas tanto para la estimación del espectro de potencias (`MASTER`) como para el ajuste de parámetros cosmológicos (`iMinuit`). En ambos análisis se utilizan los espectros de potencias TT, TE y EE agrupados en intervalos de multipolos (*bines*), y calculados sobre 12 regiones independientes del cielo. Estas regiones corresponden a los 12 píxeles definidos por el esquema `HEALPix` con $N_{\text{side}} = 1$. Al incluir la máscara de *Planck*, la fracción efectiva del cielo en cada una de las regiones varía entre un 2% y un 8%.

Trabajos previos de la colaboración *Planck* identificaron un alineamiento anómalo entre las direcciones derivadas de mapas de distribución de potencia en temperatura, abarcando un amplio rango de escalas angulares. En el primer análisis cuantificamos este alineamiento en los datos más recientes de *Planck* a través del estadístico de Rayleigh, confirmando la existencia de un agrupamiento sobre un intervalo aún más amplio de multipolos. En particular, el valor p, definido como la fracción de simulaciones con un estadístico de Rayleigh superior al observado en los datos, es inferior al 1% para multipolos comprendidos entre $\ell = 200$ y $\ell = 2000$. El uso de una máscara no apodizada podría explicar la discrepancia observada con respecto a trabajos anteriores para $\ell > 1000$, dado que esto puede inducir correlaciones entre bines adyacentes y dar lugar a un agrupamiento artificial. Asimismo, el indicio de anomalía observado en los modos E de polarización en PR3 desaparece en estos nuevos datos, con solo un par de bines mostrando valores p inferiores al 1%. Esto sucede para el mismo intervalo de multipolos en el que TT y el alineamiento entre TT y EE comienzan a mostrar un comportamiento anómalo. No obstante, su interpretación no es concluyente, ya que carece de robustez cuando se realizan pequeñas modificaciones en el estimador. Dado que la anomalía en los modos E de polarización se restringe a un número muy limitado de bines, lo más probable es que se trate de una fluctuación estadística. El mismo razonamiento se aplica al espectro de TE que muestra algunos valores p próximos al 100%.

En el segundo análisis se estudia la posible presencia de variaciones dipolares en la distribución espacial de los parámetros cosmológicos. Para ello, se estiman los paramétros en cada una de las 12 regiones maximizando la función de verosimilitud con `iMinuit`. En particular, la función de verosimilitud utiliza los espectros calculados en el análisis previo. Dado que la fracción del cielo en cada parche es muy limitada, no es posible acceder a los multipolos más bajos y, en consecuencia, es necesario fijar el valor de τ.

Al ajustar un dipolo sobre los mapas resultantes, se observa que, salvo A_s, el resto de parámetros son consistentes con las predicciones del modelo ΛCDM. En el caso de A_s, solo 5 de las 600 simulaciones muestran una amplitud igual o superior a la observada en los datos. Esta anomalía parece estar presente en la combinación $A_s e^{-2\tau}$, que es la cantidad mejor determinada por los datos. Al fijar τ, la anomalía se manifiesta en A_s. La dirección del dipolo es próxima a la de la HPA, y se sitúa en la región del cielo donde se agrupan los dipolos ajustados sobre los mapas de potencia. Esto sugiere una posible conexión entre estas anomalías. A pesar de que la anomalía siga presente al realizar diferentes cortes en $\ell_{\rm min}$ y $\ell_{\rm max}$, desaparece al excluir los espectros de potencias que involucran los datos de polarización.

La detección, en simulaciones, del *boosting* generado por el efecto *Doppler* es un resultado que valida nuestra metodología. Este efecto no puede ser el origen de la anomalía observada, dado que la dirección del dipolo de A_s se encuentra aproximadamente a 55 grados de la dirección del dipolo del FCM.

Nuestros resultados difieren de los obtenidos en estudios recientes que afirman haber encontrar evidencia de variaciones significativas de los parámetros cosmológicos a lo largo del cielo. Esta discrepancia podría deberse a diferencias metodológicas. En particular, nuestro análisis utiliza regiones del cielo completamente independientes y la matriz de covarianzas es derivada íntegramente a partir de simulaciones realistas de *Planck*. Por último, la evidencia de una posible anomalía en la combinación $A_s e^{-2\tau}$ es moderada y no concluyente.

8.4 Resumen y trabajo a futuro

Se confirma la presencia de una HPA en los datos de temperatura de PR4 mediante el estimador de varianza local, con un valor p inferior al 1%. Tras aplicar una técnica de *inpainting* a los parámetros de Stokes Q y U, el mapa reconstruido de modos E muestra una asimetría similar con una significancia entre el 1% y el 3%, y un eje casi alineado con el observado en temperatura. Sorprendentemente, al ajustar un dipolo sobre la distribución espacial de la potencia en banda del espectro angular de potencias de temperatura, se observa un agrupamiento anómalo de las direcciones, alineadas de forma notable con el eje de la HPA. No se detecta un agrupamiento análogo en los datos de polarización. En base a estos resultados, se investiga la distribución espacial de los parámetros cosmológicos. La mayoría de ellos son compatibles con las predicciones del modelo ΛCDM, a excepción de A_s, que presenta una amplitude dipolar anómala con una significancia estadística del orden del 1%, con una dirección nuevamente próxima a la de la HPA. Cabe destacar que en Mukherjee et al. (2016) identificaron los parámetros A_s y n_s como los más susceptibles de adquirir una dependencia espacial en presencia de una asimetría hemisférica, en concordancia con nuestros resultados.

Aunque no hay duda sobre la asimetría observada en temperature en los capítulos 4 y 6, las grandes incertidumbres en los resultados de los modos E de polarización impiden extraer conclusiones robustas. Estas incertidumbres se deben tanto a los efectos sistemáticos a gran escala como a la baja relación señal-ruido inherentes a *Planck*.

Dado que las medidas de temperatura de *Planck* ya están limitadas por la varianza cósmica hasta $\ell \sim 1600$, no se espera que futuros datos de temperatura tengan un impacto significativo en el estudio de las anomalías del FCM, más allá de un aumento en la fracción de cielo limpio de contaminantes. Sin embargo, se espera una mejora sustancial en los análisis de polarización. Una continuación natural y prometedora de este trabajo sería aplicar la misma metodología a los futuros datos de polarización de alta precisión.

Appendix A:
Convergence of the covariance matrix

In Section 5.3, we show the performance of the inpainting for two different scenarios: (1) CMB signal only and (2) CMB signal plus a semi-realistic noise realization. In the second case, the methodology is able to reconstruct accurately the systematics and noise. However, this is only possible if the full anisotropic covariance matrix is well characterised. In a more realistic case, where a limited number of simulations are used to compute the matrix numerically, the non-convergence of the matrix induces a mismatch between the matrix and the pixels outside the mask. This introduces some artifacts in the inpainted realizations. Figure 8.1 shows an example of the input and output T, Q, and U maps using just 20 semi-realistic noise simulations to compute the covariance matrices. Even if the mismatch is mainly for the polarisation field[1], strong cold and hot spots are induced in the inpainted temperature map through the TE correlation.

The mismatch is even more clear in Figure 8.2, where the probability distribution of the z variables (see Eq. 5.12) is plotted for T, Q, and U. For comparison, a Gaussian curve with the same standard deviation is plotted. In all the cases shown in Figure 8.2, the variable z_T, corresponding to the temperature field, follows a Gaussian distribution with zero mean and unit variance, $\mathcal{N}(0, 1)$. This is because there is not a mismatch between the statistical properties of the pixels and those encoded in the matrix. In this case, even if we are not including correctly the noise and systematics on the matrix, the mismatch is simply masked by the regularization noise. However, the mismatch strongly affects the polarisation field. In particular, since Q and then U are computed recursively, this effect is most notable for the z variables associated to U, which is found to be the broadest one. The z_Q and z_U distributions become effectively $\mathcal{N}(0, 1)$ when several thousand of noise simulations are considered to construct the matrix.

---

[1] For temperature the noise is much smaller than the signal and, even if it were not well modelled in the covariance matrix, its effect is negligible as the error in the z variables is masked by the regularization noise.

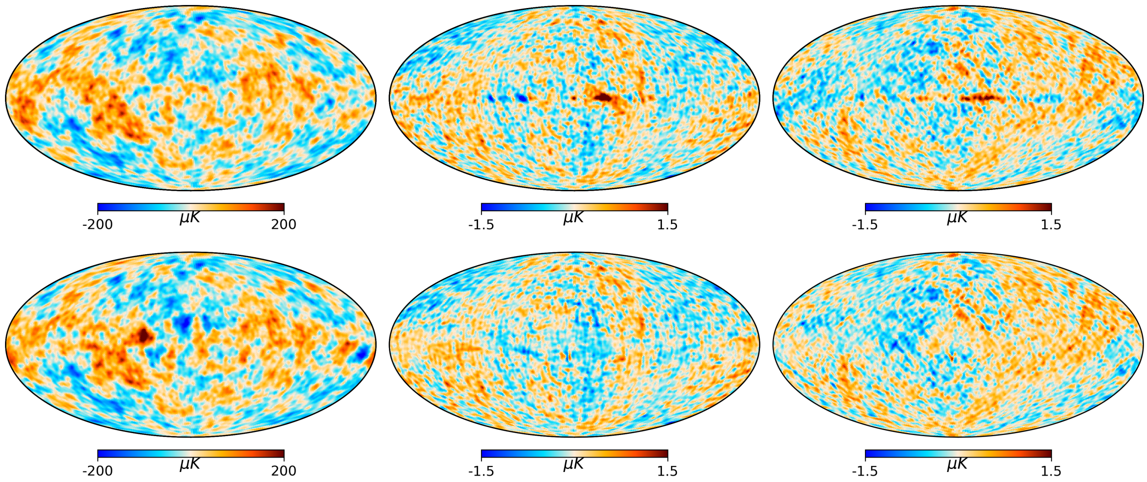

FIGURE 8.1: Example of a T (left), Q (middle), and U (right) inpainted realization in the case where 20 semi-realistic noise simulations are used to characterise the noise covariance matrix. The top and bottom panels correspond to the input and inpainted maps, respectively.

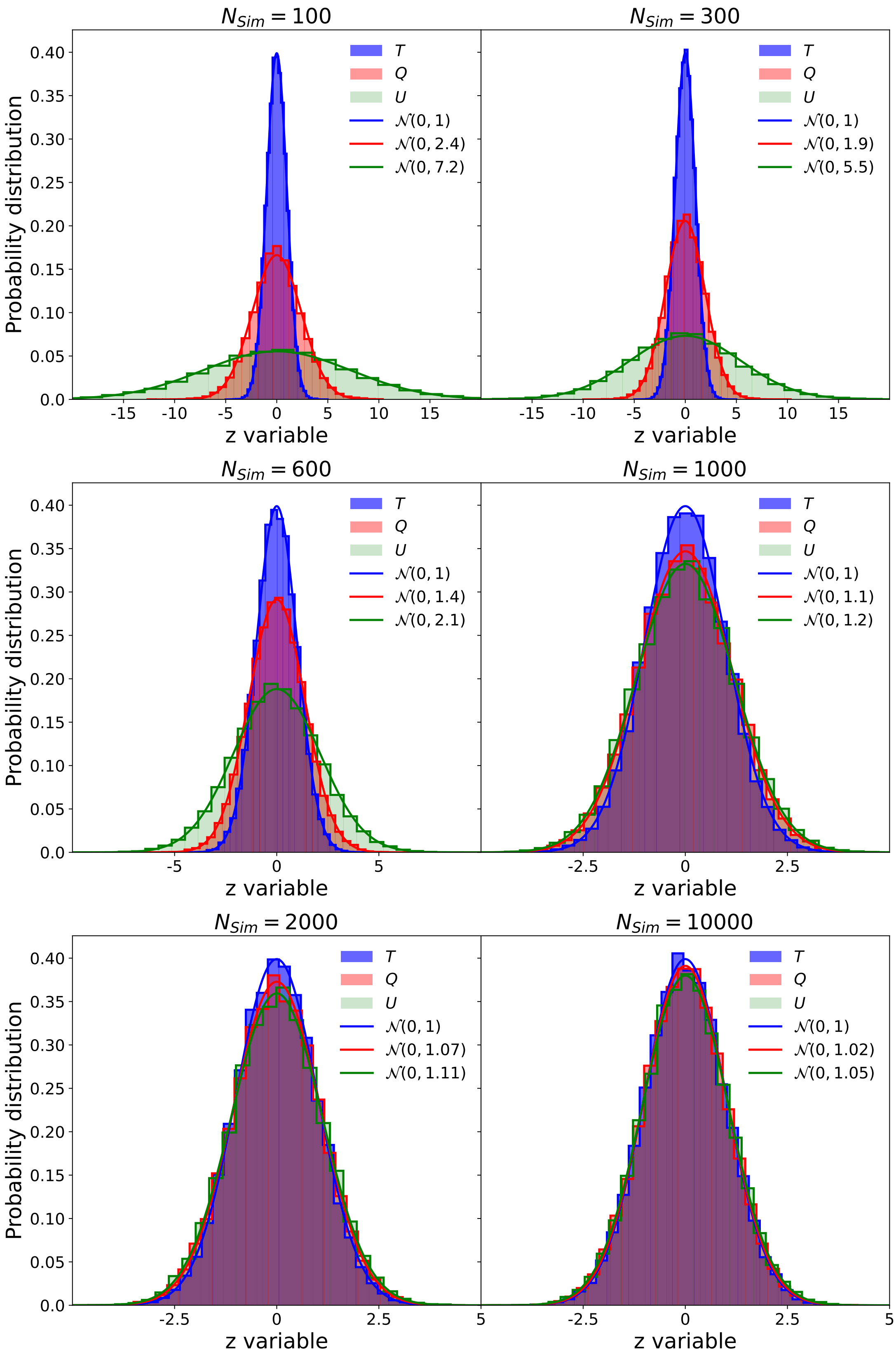

FIGURE 8.2: Distribution of the z variable associated to the T, Q, and U fields for different number of simulations used to estimate the noise covariance matrix. For comparison, a Gaussian with the same standard deviation is also given in each case.

Appendix B:
Robustness of the inpainting against the model

In Section 5.3, we generate a pixel covariance matrix that match perfectly the input simulation. In a real situation this is not possible. Here, we introduce a small mismatch between the model used for the input simulation and the one used for the pixel covariance matrix estimation. As described in Section 5.1, we use the best fit to the ΛCDM model to generate the input CMB simulation. We introduce a small deviation in the input parameters taking into account the correlation matrix of the estimated parameters inferred from the Planck 2018 data. In Table 8.1, we show the best fit parameters and the modified ones. The alternative parameters were obtained as a Gaussian random realisation of the parameters centred in the best-fit model and following the correlation matrix. Therefore, they are also consistent with the Planck data within the estimated errors.

We compute the relative error with respect cosmic variance in the E- and B-mode reconstruction using the 1200 inpainted realizations, as explained in Section 5.3.3, for both cases, the exact and the modified model. Then, we get the differences between previous errors, which are on average at the level of 0.1% (errors are slightly larger in the modified case) as it is shown in Figure 8.3.

	ΛCDM best fit	Modified ΛCDM
$\Omega_b h^2$	0.02238280	0.02244959
$\Omega_c h^2$	0.1201075	0.1194516
H_0	67.32117	67.57953
τ	0.05430842	0.05744905
$\ln(10^{10}A_s)$	3.044784	3.056551
n_s	0.9660499	0.9654136

Table 8.1: Cosmological parameters. *Left*: ΛCDM best fit. Model used for the input simulation. *Right*: Modified model according to the Planck 2018 errorbars and correlations.

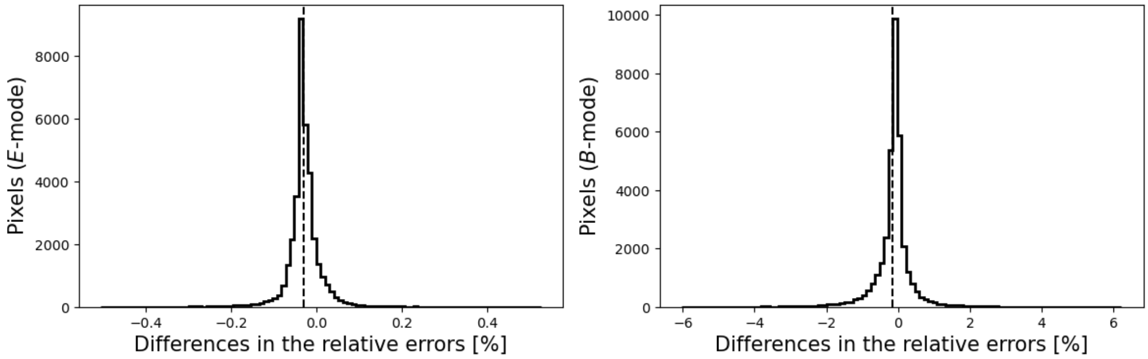

FIGURE 8.3: Distribution of differences of the errors relative to the cosmic variance of the exact minus the modified model for E-mode (*left*) and B-mode (*right*).

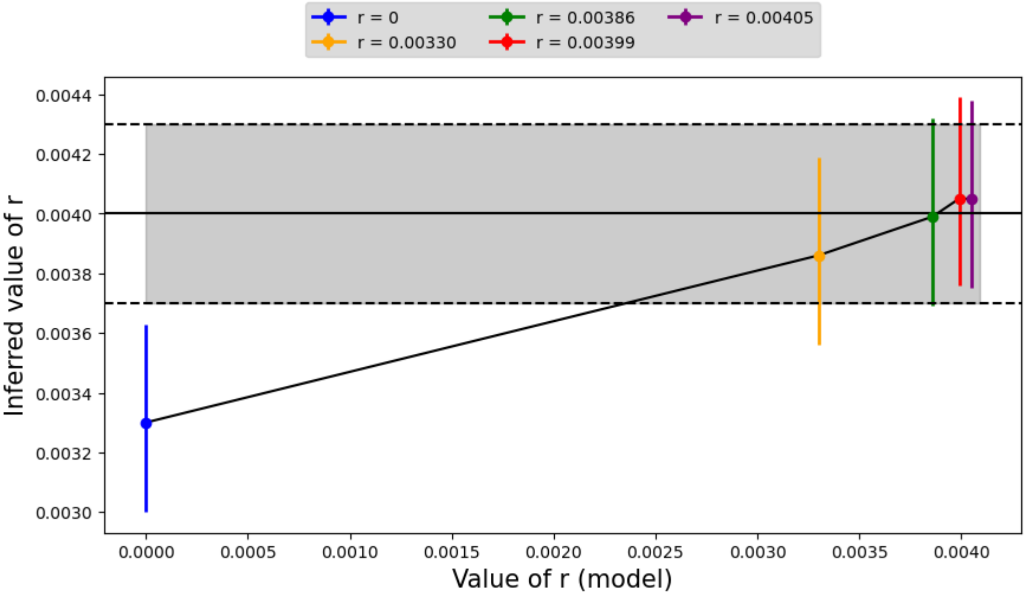

FIGURE 8.4: Value of r assumed in the fiducial model versus the inferred value. Grey shaded region corresponds to the 1σ interval of the posterior for the input simulation, which is generated assuming a tensorial modes with r = 0.004. Solid black line corresponds to the value where the posterior peaks for the input simulation, which by chance corresponds to the same value as the model. Dots corresponds to the values of r where the mean posterior peaks for each assumed model in the iterative process. Error bars are the 1σ interval of the mean posterior.

Similar results are found at the power spectra level. This is expected taking into account that the differences between models in this multipole range ($\ell = 2 - 192$) is much smaller than the cosmic variance, so smaller than the differences between different realizations of the same model.

As an extra test, we consider a special scenario where tensor-to-scalar ratio (r) is equal to 0.004. In this case, most of the constraining power is encoded in the B-mode large scales, so differences in the model can impact and bias the r estimation. We demonstrate that an iterative process can be a good approach to this issue. Taking into account that tensorial modes have not been detected yet, it is reasonable to start with a model where $r = 0$. We generate 1200 inpainted realizations assuming $r = 0$ model and we estimate for each of them the posterior (using a flat prior) of r using the following exact likelihood in harmonic space:

$$-\log \mathcal{L} = \sum_{\ell} \left[\frac{\hat{\mathcal{C}}_\ell}{\mathcal{C}_\ell} + \log \mathcal{C}_\ell - \frac{2\ell - 1}{2\ell + 1} \log \hat{\mathcal{C}}_\ell \right], \tag{8.1}$$

where $\hat{\mathcal{C}}_\ell$ is the B-mode power spectrum of the realization, $\mathcal{C}_\ell$ is the theoretical spectrum, and the sum is done up to $\ell = 2N_{\mathrm{side}}$. We generate the mean posterior by averaging the -$\log \mathcal{L}$, and then, we use the value of r where the mean posterior peaks to generate the model for the next iteration. Results for the different iterations are shown in Figure 8.4, where the blue contour corresponds to the 1σ interval obtained from the posterior distribution of the input simulation (a uniform prior is used for r). Red error bars corresponds to 1σ interval of the mean posterior. It becomes apparent that after a few iterations the correct value of r is recovered. Therefore, when using the inpainting technique, a comparison between the assumed model and the one recovered from the inpainted maps is recommended in order to test the consistency of the results.

We finally show a comparison between the B-mode power spectrum obtained using different models. Left panel of Figure 8.5 shows D_ℓ^{BB} of the input simulation (red), the theoretical curve of the input model (black), and the estimated $\tilde{D}_\ell^{BB}$ for the different fiducial models (corresponding to different iterations), which are described by the median of the distribution of the 1200 spectra, obtained from each of the inpainted realizations, and the 68% C.L (error bars). In the right panel we show β_ℓ defined as,

$$\beta_\ell = \frac{x_\ell}{\sigma_\ell}, \tag{8.2}$$

where x_ℓ is the difference between the input spectrum and the median, and σ_ℓ is the 68% two sided C.L, thus allowing asymmetric error bars for low ℓ. It is clear that a bad choice of the model has a bigger impact on the first multipoles, corresponding to the reionization bump, than in the lensing dominated range.

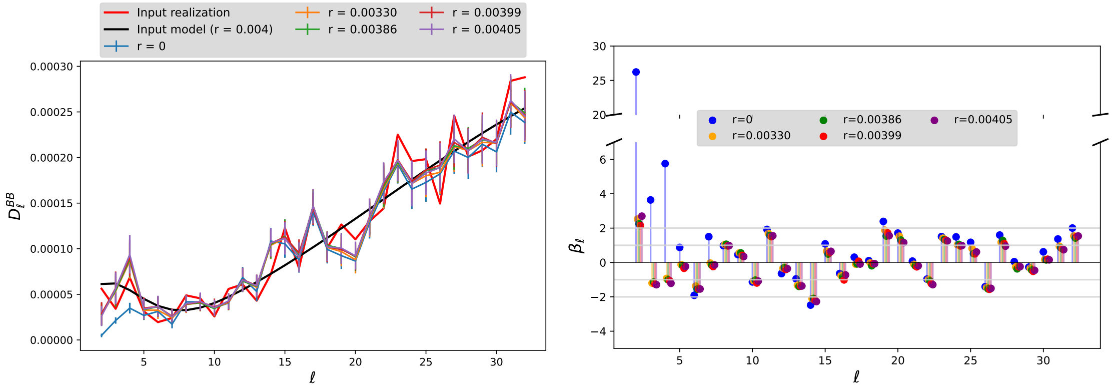

FIGURE 8.5: *Left panel*: D_ℓ^{BB} for the input simulation (in red), theoretical curve for r = 0.004 (in black), and $\tilde{D}_\ell^{BB}$ for different fiducial models, obtained from the distribution of 1200 inpainted realizations. Error bars are also derived from the 68% C.L of that distribution. *Right panel*: β_ℓ for different models.

Appendix C:

Robustness of results on cosmological parameters

In order to test the robustness of the results on the anomalous amplitude of A_s, we consider a set of pipeline runs while modifying some parameters. We run the following cases:

- **Debiased case ($TTTEEE$):** Same pipeline as in the main analysis but subtracting the unknown transfer function from TT, TE, and EE. Remember that these transfer functions are calibrated with the simulations. The main goal of this run is to validate that the transfer function is not introducing any anisotropy in the cosmological parameters.
- **Debiased case (only T):** Same pipeline but only for the temperature data.
- **Different τ:** In this case, we modify the value to which τ is fixed. We fix τ to 0.0544, which was the best-fit value inferred from Planck Data Release 3 (Planck Collaboration et al., 2020c), instead to the input value used for the E2E simulations. In this sense, we can check if the anomaly persists when the fixed value is not the correct one, which may be the case for the data.
- **$\ell_{\mathrm{max}} = 1000$:** In this case we remove some multipole bins for TE and EE. In particular, we consider $\ell_{\mathrm{max}}^{TE} \sim 1500$ and $\ell_{\mathrm{max}}^{EE} \sim 1000$.
- **$\ell_{\mathrm{min}} = 62$:** In this final case, we remove the large angular scale bins, and start at $\ell = 62$ for TT, TE, and EE.

The values of the PTEs for the five cosmological parameters are shown in Table 8.2, together with the values obtained in the main analysis. Notably, both main and debiased cases produce very similar results, finding again that the PTE for A_s is below 1% if polarisation is included. The results also remain very stable when using an alternative value of τ. Furthermore, it appears to be robust against different choices of multipole cuts. In particular,

Scenario	PTE (H_0)	PTE ($\Omega_b h^2$)	PTE ($\Omega_c h^2$)	PTE (A_s)	PTE (n_s)
Main (Inc. Pol)	13.3%	81.0%	10.2%	0.83%	5.8%
Main (Only T)	85.5%	60.7%	81.0%	33.0%	38.8%
Debiased (Inc. Pol)	14.3%	81.3%	10.8%	0.83%	6.0%
Debiased (Only T)	86.8%	62.7%	81.7%	33.5%	38.5%
$\tau = 0.0544$	13.5%	80.7%	10.3%	0.83%	5.8%
$\ell_{\mathrm{max}} = 1000$	21.2%	86.3%	11.7%	1.5%	10.2%
$\ell_{\mathrm{min}} = 62$	22.0%	79.2%	18.8%	2.7%	12.8%

Table 8.2: PTEs of the five cosmological parameters for seven different scenarios as a test of robustness.

if the last 500 multipoles of the E modes are removed, and only the first 1000 multipoles are considered, the PTE remains below 2%. If the first bin, which contains multipoles between $\ell = 32$ and $\ell = 61$, is removed, the PTE increases to 2.7%. Although this decrease in significance may be attributed to the loss of statistical information when excluding some multipoles, it may also suggest that the anomaly has a stronger contribution from large scales. Notably, it is precisely at large angular scales where the known CMB anomalies are observed in the temperature data, particularly the HPA.

We pay particular attention to the case where τ is fixed to a different value. In this scenario, the other cosmological parameters remain largely unaffected except for A_s, which is expected given their correlation — a smaller τ produces a smaller A_s, and vice versa. However, because τ is fixed to the same value in all patches, A_s is shifted uniformly, leaving the dipole amplitude unchanged. In other words, fixing τ to the same values for all patches allows the same amount of fluctuations in A_s independently of the fixed value. Consequently, we obtain a similar PTE.

Figure 8.6 shows results for the five cosmological parameters with two fixed values of τ, computed for patch number 4. It is clear that, except for A_s, the other parameters remain unaffected, while A_s is reduced. However, the input value is still within 1σ. Similarly, the left panel of Figure 8.7 shows the shift in A_s. As previously mentioned, this shift does not affect the dipole amplitudes, as shown in the right panel.

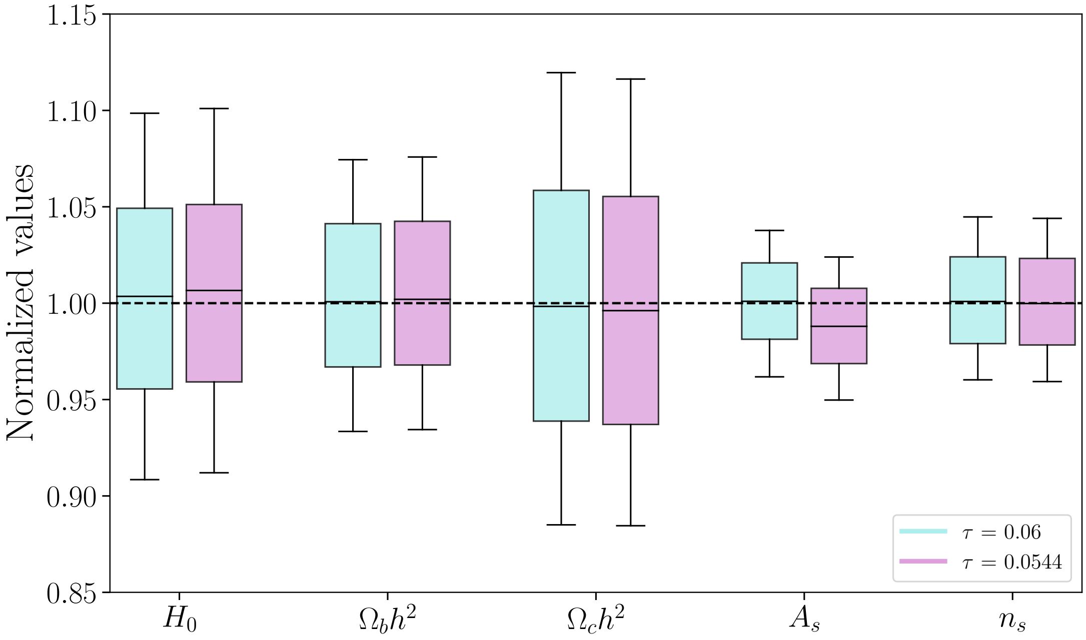

FIGURE 8.6: Distribution of the five cosmological parameters for the 600 E2E simulations computed on patch number 4 ($f_{\rm sky} \approx$ 2%) and fixing τ to two different values. Distributions are normalized to the input values. The boxes represent 68% of the probability, while the whiskers include 95%.

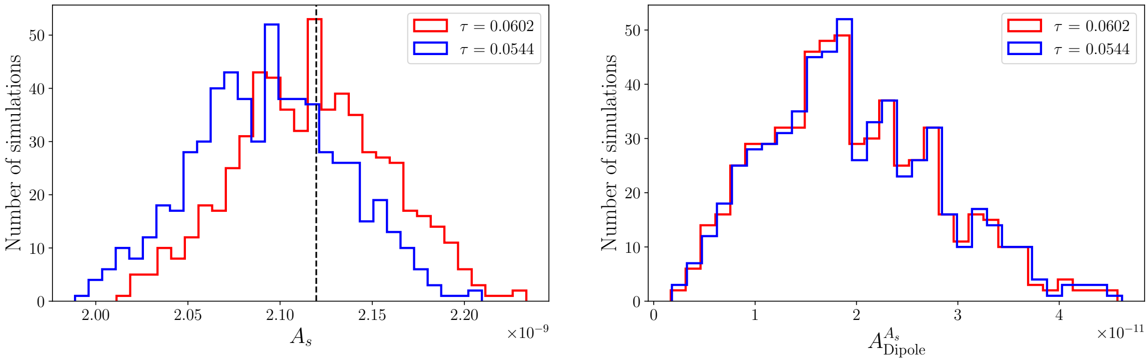

FIGURE 8.7: *Left panel:* Distribution of the A_s parameter for the 600 E2E simulations fixing τ to 0.0602 (in red) and 0.0544 (in blue). The black dashed line corresponds to the input A_s value for the simulations. *Right panel:* Distribution of the dipole amplitudes of A_s for the 600 E2E PR4 simulations. Again the red colour is for τ = 0.0602, while blue is used for τ = 0.0544.

Bibliography

Abazajian K. N., et al., 2016, arXiv e-prints, p. arXiv:1610.02743

Abbott B. P., et al., 2017, Nature, 551, 85

Abbott B. P., et al., 2021, ApJ, 909, 218

Abrial P., Moudden Y., Starck J. L., Fadili J., Delabrouille J., Nguyen M. K., 2008, Statistical Methodology, 5, 289

Adachi S., et al., 2022, ApJ, 931, 101

Adame A. G., et al., 2025, J. Cosmology Astropart. Phys., 2025, 021

Ade P. A. R., et al., 2014, A&A, 571, A23

Ade P. A. R., et al., 2016, A&A, 594, A16

Ade P., et al., 2019, J. Cosmology Astropart. Phys., 2019, 056

Adler R. J., Casey B., Jacob O. C., 1995, American Journal of Physics, 63, 620

Aghanim N., et al., 2020, A&A, 641, A5

Akrami Y., Fantaye Y., Shafieloo A., Eriksen H. K., Hansen F. K., Banday A. J., Górski K. M., 2014, ApJ, 784, L42

Alonso D., Sanchez J., Slosar A., 2019, Mon. Not. Roy. Astron. Soc., 484, 4127

Alpher R. A., Bethe H., Gamow G., 1948, Phys. Rev., 73, 803

Aluri P. K., Shafieloo A., 2017, arXiv e-prints, p. arXiv:1710.00580

Aricò G., Angulo R. E., Zennaro M., 2021, arXiv e-prints, p. arXiv:2104.14568

Axelsson M., Fantaye Y., Hansen F. K., Banday A. J., Eriksen H. K., Gorski K. M., 2013, ApJ, 773, L3

BICEP2 Collaboration et al., 2014, Phys. Rev. Lett., 112, 241101

BICEP2/Keck Collaboration et al., 2015, Phys. Rev. Lett., 114, 101301

Banday A. J., et al., 2025, arXiv e-prints, p. arXiv:2508.16451

Barkana R., Loeb A., 2001, Physics Reports, 349, 125

Barreiro R. B., Vielva P., Hernandez-Monteagudo C., Martinez-Gonzalez E., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 747

Basak S., Delabrouille J., 2012, MNRAS, 419, 1163

Baumann D., 2009, arXiv e-prints, p. arXiv:0907.5424

Baumann D., 2022, Cosmology. Cambridge University Press, doi:10.1017/9781108937092

Benitez N., et al., 2014, arXiv e-prints, p. arXiv:1403.5237

Bennett C. L., et al., 2003, ApJS, 148, 1

Bennett C., et al., 2011, ApJS, 192, 17

Bennett C. L., et al., 2013, ApJS, 208, 20

Benoit-Lévy A., Déchelette T., Benabed K., Cardoso J. F., Hanson D., Prunet S., 2013, A&A, 555, A37

Bilbao-Ahedo J. D., Barreiro R. B., Herranz D., Vielva P., Martínez-González E., 2017, J.~Cosmology Astropart. Phys., 2017, 022

Bilbao-Ahedo J. D., Barreiro R. B., Vielva P., Martínez-González E., Herranz D., 2021, JCAP, 07, 034

Billi M., Barreiro R. B., Martínez-González E., 2024, J. Cosmology Astropart. Phys., 2024, 080

Blas D., Lesgourgues J., Tram T., 2011, J. Cosmology Astropart. Phys., 2011, 034

Bucher M., Louis T., 2012, MNRAS, 424, 1694

Cabella P., Kamionkowski M., 2004, arXiv e-prints, pp astro–ph/0403392

Cardoso J.-F., Martin M., Delabrouille J., Betoule M., Patanchon G., 2008, arXiv e-prints, p. arXiv:0803.1814

Cayón L., Jin J., Treaster A., 2005, MNRAS, 362, 826

Challinor A., van Leeuwen F., 2002, Phys. Rev. D, 65, 103001

Chevallier M., Polarski D., 2001, International Journal of Modern Physics D, 10, 213

Cirelli M., Strumia A., Zupan J., 2024, arXiv e-prints, p. arXiv:2406.01705

Colless M., et al., 2001, MNRAS, 328, 1039

Contaldi C. R., Peloso M., Kofman L., Linde A., 2003, J. Cosmology Astropart. Phys., 2003, 002

Contreras D., Zibin J. P., Scott D., Banday A. J., Górski K. M., 2017, Phys. Rev. D, 96, 123522

Contreras D., Hutchinson J., Moss A., Scott D., Zibin J. P., 2018, Phys. Rev. D, 97, 063504

Cooke R., 2024, arXiv e-prints, p. arXiv:2409.06015

Cooke R. J., Pettini M., Steidel C. C., 2018, The Astrophysical Journal, 855, 102

Copi C. J., Huterer D., Starkman G. D., 2004, Phys. Rev. D, 70, 043515

Copi C. J., Huterer D., Schwarz D. J., Starkman G. D., 2007, Phys. Rev. D, 75, 023507

Copi C. J., Huterer D., Schwarz D. J., Starkman G. D., 2009, MNRAS, 399, 295

Copi C. J., Huterer D., Schwarz D. J., Starkman G. D., 2013, MNRAS, 434, 3590

Copi C. J., Huterer D., Schwarz D. J., Starkman G. D., 2015a, MNRAS, 449, 3458

Copi C. J., Huterer D., Schwarz D. J., Starkman G. D., 2015b, MNRAS, 451, 2978

Cruz M., Martínez-González E., Vielva P., Cayón L., 2005, MNRAS, 356, 29

Cruz M., Tucci M., Martínez-González E., Vielva P., 2006, MNRAS, 369, 57

Cruz M., Turok N., Vielva P., Martínez-González E., Hobson M., 2007a, Science, 318, 1612

Cruz M., Cayón L., Martínez-González E., Vielva P., Jin J., 2007b, ApJ, 655, 11

Cruz M., Martínez-González E., Vielva P., Diego J. M., Hobson M., Turok N., 2008, MNRAS, 390, 913

Cruz M., Vielva P., Martínez-González E., Barreiro R. B., 2011, MNRAS, 412, 2383

Cruz M., Martínez-González E., Gimeno-Amo C., Kavanagh B. J., Tucci M., 2025, J. Cosmology Astropart. Phys., 2025, 079

DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036

DESI Collaboration et al., 2025, arXiv e-prints, p. arXiv:2503.14738

Dai L., Jeong D., Kamionkowski M., Chluba J., 2013, Phys. Rev. D, 87, 123005

Dark Energy Survey Collaboration et al., 2016, MNRAS, 460, 1270

Dawson K. S., et al., 2013, AJ, 145, 10

Dayal P., Ferrara A., 2018, Phys. Rep., 780, 1

Di Valentino E., et al., 2025, arXiv e-prints, p. arXiv:2504.01669

Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, ApJ, 142, 414

Dodelson S., 2003, Modern Cosmology. Academic Press, Amsterdam

Eimer J. R., et al., 2024, ApJ, 963, 92

Einstein A., 1905, Annalen der Physik, 322, 891

Einstein A., 1916, Annalen der Physik, 354, 769

Erickcek A. L., Kamionkowski M., Carroll S. M., 2008, Phys. Rev. D, 78, 123520

Erickcek A. L., Hirata C. M., Kamionkowski M., 2009, Phys. Rev. D, 80, 083507

Eriksen H. K., Hansen F. K., Banday A. J., Górski K. M., Lilje P. B., 2004a, ApJ, 605, 14

Eriksen H. K., Hansen F. K., Banday A. J., Górski K. M., Lilje P. B., 2004b, ApJ, 609, 1198

Eriksen H. K., Banday A. J., Górski K. M., Hansen F. K., Lilje P. B., 2007, ApJ, 660, L81

Eriksen H. K., Jewell J. B., Dickinson C., Banday A. J., Górski K. M., Lawrence C. R., 2008, ApJ, 676, 10

Euclid Collaboration et al., 2025, A&A, 697, A1

Fernández-Cobos R., Vielva P., Barreiro R. B., Martínez-González E., 2012, MNRAS, 420, 2162

Finelli F., García-Bellido J., Kovács A., Paci F., Szapudi I., 2016, MNRAS, 455, 1246

Fixsen D. J., 2009, ApJ, 707, 916

Fosalba P., Gaztañaga E., 2021, MNRAS, 504, 5840

Freedman W. L., Madore B. F., Jang I. S., Hoyt T. J., Lee A. J., Owens K. A., 2024, arXiv e-prints, p. arXiv:2408.06153

Friedman A., 1922, Z. Physik, 10, 377

Fukuda Y., et al., 1998, Phys. Rev. Lett., 81, 1562

Gamow G., 1948, Nature, 162, 680

Génova-Santos R., et al., 2017, MNRAS, 464, 4107

Gimeno-Amo C., Barreiro R. B., Martínez-González E., Marcos-Caballero A., 2023, J. Cosmology Astropart. Phys., 2023, 029

Gimeno-Amo C., Martínez-González E., Barreiro R. B., 2024, J. Cosmology Astropart. Phys., 2024, 038

Gimeno-Amo C., Hansen F. K., Martínez-González E., Barreiro R. B., Banday A. J., 2025, arXiv e-prints, p. arXiv:2504.05597

González-González R., et al., 2025, A&A, 695, A245

Gordon C., 2007, ApJ, 656, 636

Gordon C., Hu W., Huterer D., Crawford T., 2005, Phys. Rev.~D, 72, 103002

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Gruppuso A., 2014, MNRAS, 437, 2076

Gruppuso A., Finelli F., Natoli P., Paci F., Cabella P., de Rosa A., Mandolesi N., 2011, MNRAS, 411, 1445

Gruppuso A., Natoli P., Paci F., Finelli F., Molinari D., De Rosa A., Mandolesi N., 2013, J. Cosmology Astropart. Phys., 2013, 047

Guth A. H., 1981, Phys. Rev. D, 23, 347

Hajian A., Souradeep T., 2003, ApJ, 597, L5

Halverson N. W., et al., 2002, ApJ, 568, 38

Hansen F. K., Banday A. J., Górski K. M., 2004, MNRAS, 354, 641

Hansen F. K., Banday A. J., Górski K. M., Eriksen H. K., Lilje P. B., 2009, ApJ, 704, 1448

Hansen F. K., Boero E. F., Luparello H. E., Lambas D. G., 2023, Astron. Astrophys., 675, L7

Hazumi M., et al., 2020, in Lystrup M., Perrin M. D., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 11443, Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave. p. 114432F (`arXiv:2101.12449`), doi:10.1117/12.2563050

Hinshaw G., Banday A. J., Bennett C. L., Gorski K. M., Kogut A., Smoot G. F., Wright E. L., 1996a, ApJ, 464, L17

Hinshaw G., Branday A. J., Bennett C. L., Gorski K. M., Kogut A., Lineweaver C. H., Smoot G. F., Wright E. L., 1996b, ApJ, 464, L25

Hinshaw G., et al., 2013, ApJS, 208, 19

Hirata C. M., 2009, J. Cosmology Astropart. Phys., 2009, 011

Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2

Hoffman Y., Ribak E., 1991, ApJ, 380, L5

Hoftuft J., Eriksen H. K., Banday A. J., Górski K. M., Hansen F. K., Lilje P. B., 2009, ApJ, 699, 985

Hoyt T. J., Jang I. S., Freedman W. L., Madore B. F., Owens K. A., Lee A. J., 2025, arXiv e-prints, p. arXiv:2503.11769

Hu W., White M., 1997, New A, 2, 323

Inoue K. T., Silk J., 2006, ApJ, 648, 23

Inoue K. T., Cabella P., Komatsu E., 2008, Phys. Rev.~D, 77, 123539

Ivezić Ž., et al., 2019, ApJ, 873, 111

Jones J., Copi C. J., Starkman G. D., Akrami Y., 2023, arXiv e-prints, p. arXiv:2310.12859

Jung G., Aghanim N., Sorce J. G., Seidel B., Dolag K., Douspis M., 2024, A&A, 692, A180

K. Aluri P., et al., 2023, Classical and Quantum Gravity, 40, 094001

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

Kim J., Naselsky P., 2010a, Phys. Rev. D, 82, 063002

Kim J., Naselsky P., 2010b, ApJ, 714, L265

Kim J., Naselsky P., Mandolesi N., 2012, ApJ, 750, L9

Kolb E. W., Turner M. S., 1990, The early universe. Frontiers in physics, Westview Press, Boulder, CO, doi:10.1201/9780429492860, `https://cds.cern.ch/record/206230`

Kothari R., 2018, Gen. Rel. Grav., 54, 2022

Kuo C. L., et al., 2004, ApJ, 600, 32

Kurichin O. A., Kislitsyn P. A., Klimenko V. V., Balashev S. A., Ivanchik A. V., 2021, MNRAS, 502, 3045

Land K., Magueijo J., 2005, Phys. Rev. D, 72, 101302

Lee A. T., et al., 2001, ApJ, 561, L1

Leitch E. M., et al., 2002, Nature, 420, 763

Leitch E. M., Kovac J. M., Halverson N. W., Carlstrom J. E., Pryke C., Smith M. W. E., 2005, The Astrophysical Journal, 624, 10

Leloup C., et al., 2024, J. Cosmology Astropart. Phys., 2024, 011

Lemaître G., 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

Lesgourgues J., 2013, arXiv e-prints, p. arXiv:1302.4640

Lesgourgues J., Mangano G., Miele G., Pastor S., 2013, Neutrino Cosmology. Cambridge University Press

Lewis A., 2003, Phys. Rev.~D, 68, 083509

Lewis A., 2013, Phys. Rev.~D, 87, 103529

Lewis A., Bridle S., 2002, Phys. Rev.~D, 66, 103511

Lewis A., Challinor A., 2006, Phys. Rep., 429, 1

Lewis A., Challinor A., 2011, CAMB: Code for Anisotropies in the Microwave Background, Astrophysics Source Code Library, record ascl:1102.026

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Li S.-S., et al., 2023, A&A, 679, A133

Linde A., 1982, Physics Letters B, 108, 389

Linder E. V., 2003, Phys. Rev. Lett., 90, 091301

LiteBIRD Collaboration et al., 2023, Progress of Theoretical and Experimental Physics, 2023, 042F01

Liu Z.-W., Röpke F. K., Han Z., 2023, Research in Astronomy and Astrophysics, 23, 082001

Louis T., et al., 2025, arXiv e-prints, p. arXiv:2503.14452

Luparello H. E., Boero E. F., Lares M., Sánchez A. G., Garcia Lambas D., 2023, MNRAS, 518, 5643

Ma C.-P., Bertschinger E., 1995, ApJ, 455, 7

Madhavacheril M. S., et al., 2024, ApJ, 962, 113

Marcos-Caballero A., Martínez-González E., 2019, J. Cosmology Astropart. Phys., 2019, 053

Marcos-Caballero A., Fernández-Cobos R., Martínez-González E., Vielva P., 2016, MNRAS, 460, L15

Marcos-Caballero A., Martínez-González E., Vielva P., 2017a, J. Cosmology Astropart. Phys., 2017, 026

Marcos-Caballero A., Martínez-González E., Vielva P., 2017b, J. Cosmology Astropart. Phys., 2017, 023

Martin J., Ringeval C., Vennin V., 2013, arXiv e-prints, p. arXiv:1303.3787

Martínez-González E., Gallegos J. E., Argüeso F., Cayón L., Sanz J. L., 2002, MNRAS, 336, 22

Masi S., 2002, Progress in Particle and Nuclear Physics, 48, 243

Masnou S., Morel J.-M., 1998, in Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269). pp 259–263 vol.3, doi:10.1109/ICIP.1998.999016

Mather J. C., et al., 1994, ApJ, 420, 439

McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, MNRAS, 465, 2936

Monelli M., Komatsu E., Ghigna T., Matsumura T., Pisano G., Takaku R., 2024, J. Cosmology Astropart. Phys., 2024, 018

Monteserín C., Barreiro R. B., Vielva P., Martínez-González E., Hobson M. P., Lasenby A. N., 2008, MNRAS, 387, 209

Mukherjee S., Aluri P. K., Das S., Shaikh S., Souradeep T., 2016, J. Cosmology Astropart. Phys., 2016, 042

Navas S., et al., 2024, Phys. Rev. D, 110, 030001

Pakmor R., et al., 2023, MNRAS, 524, 2539

Patanchon G., Imada H., Ishino H., Matsumura T., 2024, J. Cosmology Astropart. Phys., 2024, 074

Patel S. K., Aluri P. K., Ralston J. P., 2025, MNRAS, 539, 542

Peebles P. J. E., 1968, ApJ, 153, 1

Peiris H. V., et al., 2003, ApJS, 148, 213

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S., et al., 1999, ApJ, 517, 565

Perotto L., Bobin J., Plaszczynski S., Starck J. L., Lavabre A., 2010, A&A, 519, A4

Philcox O. H. E., Ivanov M. M., 2022, Phys. Rev. D, 105, 043517

Pitrou C., Coc A., Uzan J.-P., Vangioni E., 2018, Phys. Rep., 754, 1

Planck Collaboration et al., 2011, A&A, 536, A1

Planck Collaboration et al., 2016a, A&A, 594, A9

Planck Collaboration et al., 2016b, A&A, 594, A12

Planck Collaboration et al., 2016c, A&A, 594, A20

Planck Collaboration et al., 2020a, A&A, 641, A1

Planck Collaboration et al., 2020b, A&A, 641, A4

Planck Collaboration et al., 2020c, A&A, 641, A6

Planck Collaboration et al., 2020d, A&A, 641, A7

Planck Collaboration et al., 2020e, A&A, 641, A10

Planck Collaboration et al., 2020f, A&A, 643, A42

Pospelov M., Pradler J., 2010, Annual Review of Nuclear and Particle Science, 60, 539

Puglisi G., Bai X., 2020, Astrophys. J., 905, 143

Readhead A. C. S., et al., 2004, Science, 306, 836

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 2022, ApJ, 934, L7

Rubin V. C., Ford Jr. W. K., 1970, ApJ, 159, 379

Ruiz-Granda M., et al., 2025, arXiv e-prints, p. arXiv:2507.22618

Sachs R. K., Wolfe A. M., 1967, ApJ, 147, 73

Sakharov A. D., 1991, Soviet Physics Uspekhi, 34, 392

Schaerer D., 2002, A&A, 382, 28

Schwarz D. J., Starkman G. D., Huterer D., Copi C. J., 2004, Phys. Rev. Lett., 93, 221301

Schwarz D. J., Copi C. J., Huterer D., Starkman G. D., 2016, Classical and Quantum Gravity, 33, 184001

Scott D., Rees M. J., 1990, MNRAS, 247, 510

Scott P. F., et al., 2003, MNRAS, 341, 1076

Secco L. F., et al., 2022, Phys. Rev. D, 105, 023515

Sievers J. L., et al., 2003, ApJ, 591, 599

Silk J., 1968, ApJ, 151, 459

Soares-Santos M., et al., 2019, ApJ, 876, L7

Sotiriou T. P., Faraoni V., 2010, Reviews of Modern Physics, 82, 451

Spergel D. N., et al., 2003, ApJS, 148, 175

Starobinsky A. A., 1980, Physics Letters B, 91, 99

Steigman G., 2007, Annual Review of Nuclear and Particle Science, 57, 463

Stölzner B., et al., 2025, arXiv e-prints, p. arXiv:2503.19442

Strigari L. E., 2013, Phys. Rep., 531, 1

Sunyaev R. A., Zeldovich Y. B., 1970, Ap&SS, 7, 3

Tegmark M., 1997, Phys. Rev. D, 55, 5895

Tegmark M., de Oliveira-Costa A., 2001, Phys. Rev. D, 64, 063001

The BICEP/Keck Collaboration et al., 2024, arXiv e-prints, p. arXiv:2405.19469

Torrado J., Lewis A., 2021, J.~Cosmology Astropart. Phys., 2021, 057

Tristram M., et al., 2022, Phys. Rev. D, 105, 083524

Tristram M., et al., 2024, A&A, 682, A37

Vielva P., 2010, Advances in Astronomy, 2010, 592094

Vielva P., Martínez-González E., Barreiro R. B., Sanz J. L., Cayón L., 2004, ApJ, 609, 22

Vielva P., et al., 2022, J. Cosmology Astropart. Phys., 2022, 029

Weinberg S., 2008, Cosmology

Wright E. L., Bennett C. L., Gorski K., Hinshaw G., Smoot G. F., 1996, ApJ, 464, L21

Yeung S., Chu M. C., 2022, Phys. Rev.~D, 105, 083508

York D. G., et al., 2000, AJ, 120, 1579

Zaldarriaga M., Seljak U., 1997, Phys. Rev. D, 55, 1830

Zebrowski J. A., et al., 2025, arXiv e-prints, p. arXiv:2505.02827

Zeldovich Y. B., Kurt V. G., Syunyaev R. A., 1968, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 55, 278

Zhao G.-B., et al., 2016, MNRAS, 457, 2377

Zibin J. P., Contreras D., 2017, Phys. Rev. D, 95, 063011

Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Gorski K., 2019, Journal of Open Source Software, 4, 1298

Zwicky F., 1933, Helv. Phys. Acta, 6, 110

de Oliveira-Costa A., Tegmark M., Zaldarriaga M., Hamilton A., 2004, Phys. Rev. D, 69, 063516