On the inpainting of incomplete CMB sky maps

Estudio sobre el *inpainting* de cielos incompletos del FCM

Trabajo de Fin de Máster
para acceder al

MÁSTER INTERUNIVERSITARIO EN FÍSICA DE PARTÍCULAS Y DEL COSMOS

Autor: Diego Álvarez Ortega

Directores: Patricio Vielva Martínez

Elena de la Hoz López-Collado

Septiembre - 2022

Index

Abstract

The aim of this project is to develop a new inpainting technique for masked Cosmic Microwave Background (CMB) sky maps. Due to the presence of foreground emissions, the CMB is usually recovered by using component separation methods, which allow us to separate the CMB signal from this foreground contamination. However, there are some regions (e.g the galactic plane, point sources) in which the signal recovery is so poor that it needs to be masked. Inpainting techniques try to fill in this masked region with random signals that are statistically coherent with the rest of the observed sky.

In order to do so, in this work we have proposed a novel inpainting technique, the wavelet diffuse inpainting. This method tries to extend an existing procedure, called diffuse inpainting, by applying a multiresolution approach using a type of wavelet known as `HEALPix` wavelet. This allows us to improve the range of angular scales at which the inpainting is effective, as well as to reduce computing times. For this study, we have characterized the optimal inpainting parameters of this new method and the diffuse inpainting it is based on. To test the efficacy of the technique, we have compared it with the diffuse inpainting at three levels, visual sky map reconstruction, angular power spectra recovery, and presence of correlations between different angular scales. Furthermore, some preliminary results have been obtained by applying this technique to *Planck* data.

Keywords: Cosmology, Cosmic Microwave Background, Inpainting, Image Processing

El objetivo de este proyecto es desarrollar una nueva técnica de *inpainting* para mapas enmascarados del Fondo Cósmico de Microondas (FCM). Debido a la presencia de emisiones situadas entre el FCM y nosotros, los mapas del FCM suelen obtenerse mediante el uso de métodos de separación de componentes, que permiten separar la señal de la contaminación debida a dichas emisiones. Sin embargo, existen una serie de regiones (p. ej. el plano galáctico, fuentes puntuales) donde la recuperación de la señal del FCM es demasiado pobre y, por tanto, deben ser enmascaradas. Las técnicas de *inpainting* tratan de rellenar esta región enmascarada con señales aleatorias pero estadísticamente coherentes con el resto del cielo observado.

Para ello, en este trabajo se ha propuesto una nueva técnica de *inpainting* denominada *wavelet diffuse inpainting*. Esta trata de extender la aplicación de un método conocido como *inpainting* difusivo añadiendo un enfoque multiresolución con ondículas, en concreto se ha empleado la decomposición en `HEALPix` *wavelets*. Esto nos permite ampliar el rango de escalas angulares a las que el *inpainting* es efectivo, así como reducir su coste computacional. Para este estudio, se han caracterizado los parámetros óptimos tanto de este *inpainting* con ondículas como del *inpainting* difusivo en el que se basa. Para testear la eficacia de este nuevo método se ha realizado una comparación con el *inpainting* difusivo a tres niveles: reconstrucción visual del mapa, recuperación del espectro angular de potencias y presencia de correlaciones entre distintas escalas. Además, se han obtenido unos resultados preliminares a partir de la aplicación de esta técnica a datos de *Planck*.

Palabras clave: Cosmología, Fondo Cósmico de Microondas, Rellenado, Procesamiento de Imágenes

Chapter 1

Introduction

In 1964, while working with a very sensitive horn antenna, Penzias and Wilson detected a strange excess of noise. Their intention was to study possible sources of radiation in the microwave frequency band that could be potentially harmful to satellites. Their detection was consistent with an isotropic white noise with a temperature of $3.5 \pm 1\,\mathrm{K}$, which was compatible with the predicted signal of the cosmic microwave background (CMB) [1]. They published their results in 1965 [2], for which they were finally awarded the Nobel Prize in Physics in 1978[1]. This was the first observational evidence of the existence of such cosmic background.

The CMB had already been predicted some years before Penzias and Wilson's discovery as the remnant of a primordial hot universe. The Big Bang model postulates a very hot and dense initial stage (10^{-6} seconds after the Big Bang) in which the Universe was filled with a hot interacting plasma consisting of photons, electrons, and baryons. Temperatures at this stage were high enough to keep photons confined in the plasma via Thomson scattering with the electrons, rendering thus an opaque universe. As the Universe expanded, its temperature started to drop. When the Universe was around 380 000 years old, its temperature reached 3000 K, rendering the formation of neutral hydrogen energetically favorable. Once neutral atoms began to form, Thomson scattering became strongly suppressed and photons were free to travel through the Universe. This epoch is called recombination, i.e., the epoch when photons decoupled from baryonic matter. At the time of decoupling, radiation followed a black-body distribution with a temperature close to 3000 K. As the Universe expands these photons are redshifted, hence experiencing an energy decrease. This leaves us today with a background radiation characterized by a black-body temperature of $T_0 = 2.72548 \pm 0.00057\,\mathrm{K}$ [3]. This emission lies within the microwave band of the electromagnetic spectrum, hence the name of the CMB. According to measurements so far, this emission is consistent with being homogeneous and isotropic.

Subsequent experiments, observing in different wavelengths, confirmed the black-body shape of the CMB spectrum. Furthermore, theoretical works also predicted the existence of some inhomogeneities of the order of 10^{-4} - 10^{-5} [4]. These inhomogeneities were to be produced by different effects on the last scattering surface (primary anisotropies) as well as by interactions experienced by the photons on their journey to us (secondary anisotropies). Among these primary anisotropies, we can find baryon acoustic oscillations [5] or the non-instantaneous nature of recombination [6]. As examples of secondary anisotropies, we can mention gravitational lensing [7] and the Sunyaev–Zel'dovich effect [8]. NASA's COBE mission (COsmic Background Explorer) was the first to properly observe these fluctuations of the CMB, of the order of 1 part in 10^5, in 1992 [9].

The importance of the anisotropies of the CMB is that they constitute the oldest cosmic probe to which we have access today. Their amplitude and angular dependence are highly sensitive to cosmological models and they can be used to constrain their parameters. Nowadays, the most simple

[1] https://www.nobelprize.org/prizes/physics/1978/summary/

model that is able to explain most of our Universe's properties is the ΛCDM (Lambda-Cold Dark Matter) model, also known as the standard model of Cosmology or concordance model. This specific model depends on six parameters: the Hubble parameter H_0, the physical baryon density Ω_b, the physical matter density Ω_m, the scalar spectral index n_s, the reionization optical depth τ, and the curvature fluctuation amplitude Δ^2_R. The values of these parameters can be inferred from the study of CMB anisotropy data and the large-scale structure in the universe, thus constituting the best data sources currently available in cosmology.

Following the success of COBE in observing the CMB, WMAP (Wilkinson Microwave Anisotropy Probe) [10] and ESA's *Planck* satellites [11] were launched in 2001 and 2009 respectively. These missions focused on measuring and characterizing the temperature (intensity) of the CMB, in the case of *Planck* (Figure 1.1) reaching the cosmic variance limit[2] at most scales. Although it was not designed as a polarimeter, *Planck* was also able to measure CMB's polarized E-mode. *Planck* mission was operative until 2013 and has provided the most accurate estimations of many cosmological parameters to date [12]. Future experiments aim to further characterize CMB polarization and include space missions (JAXA's LiteBIRD mission [13]), ground-based telescopes (CMB-S4 [14], Simons Observatory [15]) and even balloon experiments (SPIDER [16]).

In the following sections, we will study how the CMB anisotropies are characterized via their angular power spectra, both in temperature Section 1.1 and polarization Section 1.2. Then, we will describe other astrophysical signals in the microwave range of the electromagnetic spectrum and how they affect the detection of the CMB Section 1.3. Lastly, we will explain how to mitigate the effects of these contaminants using masks, as well as different proposals on how to recover the angular power spectra on masked maps, some of which have been studied in detail in this project Section 1.4.

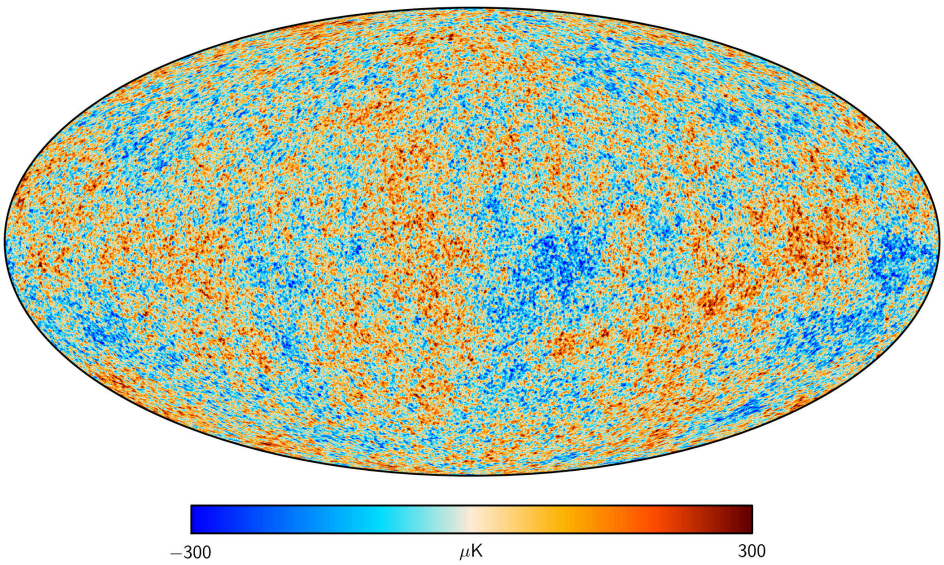

Figure 1.1: CMB temperature anisotropies from SMICA using *Planck* data in 2015. Credit: ESA and the *Planck* Collaboration[3].

[2]The cosmic variance is the inherent statistical uncertainty in CMB observations due to the limited number of observable modes.

1.1 CMB temperature angular power spectrum

As stated above, CMB fluctuations contain a lot of information about cosmology. As with any radiation field, these fluctuations can be studied by looking at their intensity and polarization along the line of sight. Since the CMB radiates as a black-body, we can relate its intensity with its temperature through Planck's law. The temperature of the CMB constitutes then a scalar field, and its fluctuations are defined as:

$$\frac{\Delta T}{T_0}(\hat{n}) = \frac{T(\hat{n}) - T_0}{T_0}, \tag{1.1}$$

where $\hat{n}$ is the unitary vector in the direction of the observation, and T_0 is the average temperature of the CMB at the present day.

Since the CMB signal is distributed all around the celestial sphere, it is useful to express these fluctuations in terms of an expansion in spherical harmonics:

$$\frac{\Delta T}{T_0}(\hat{n}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n}), \tag{1.2}$$

where $Y_{\ell m}(\hat{n})$ are spherical harmonic functions, being ℓ and m the degree and order of the spherical harmonics. ℓ are usually called multipole moments, or simply multipoles. The spherical harmonic coefficients $a_{\ell m}$, satisfying $\langle a_{\ell m} \rangle = 0$[4], are given by:

$$a_{\ell m} = \int Y^*_{\ell m}(\hat{n}) \frac{\Delta T}{T_0}(\hat{n}) d\hat{n}. \tag{1.3}$$

Since we are working under the hypothesis of isotropic fluctuations, the variance of these coefficients does not depend on m. We can therefore write this variance as:

$$\langle a^*_{\ell' m'}, a_{\ell m} \rangle = \tilde{C}_\ell \delta_{\ell \ell'} \delta_{m m'}, \tag{1.4}$$

where $\tilde{C}_\ell$ is the angular power spectrum. From the measured sky we can obtain an approximation of $\tilde{C}_\ell$, labeled as C_ℓ, by summing over the $(2\ell+1)$ independent m modes of each $a_{\ell m}$ coefficient:

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \tag{1.5}$$

Our measurements are limited to our own Universe, so the number of observable modes is also

[3] https://www.cosmos.esa.int/web/planck/picture-gallery

[4] where $\langle . \rangle$ represents the ensemble average.

limited. This limitation is translated into an inherent statistical uncertainty on these coefficients, ΔC_ℓ, known as the cosmic variance. This uncertainty is specially relevant for lower multipoles since fewer m modes are available for them. The cosmic variance can be expressed as:

$$\sigma^2(C_\ell) = \frac{2}{2\ell+1} C_\ell^2. \tag{1.6}$$

Lastly, in order to improve the visualization of the anisotropies in the plots of the angular power spectrum, it is also interesting to define D_ℓ coefficients:

$$D_\ell = \frac{\ell(\ell+1)}{2\pi} C_\ell. \tag{1.7}$$

A representation of the angular power spectra making use of these D_ℓ is shown in Figure 1.2.

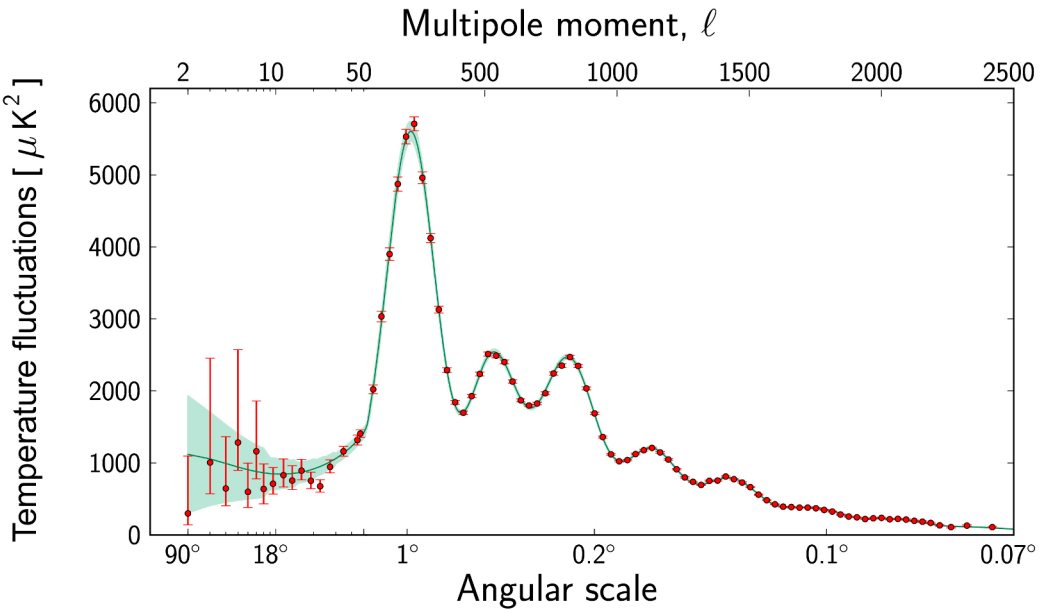

Figure 1.2: D_ℓ representation of the angular power spectrum of CMB temperature fluctuations as a function of angular scale. Red dots represent *Planck*'s measurements with their error bars. The green curve represents the best fit from the standard cosmological model to *Planck* data. The region around the curve stands for the cosmic variance. Credit: ESA and the *Planck* Collaboration[5].

[5] https://sci.esa.int/web/planck/-/51555-planck-power-spectrum-of-temperature-fluctuations-in-the-cosmic-microwave-background

1.2 CMB polarization angular power spectra

The electric field components of an electromagnetic wave with frequency ω_0 propagating along the z directions can be written as:

$$\begin{aligned} E_x &= a_x(t)\cos\left[\omega_0 t - \theta_y(t)\right], \\ E_y &= a_y(t)\cos\left[\omega_0 t - \theta_y(t)\right]. \end{aligned} \tag{1.8}$$

Polarization is often studied in terms of the Stokes parameters, which are defined as:

$$I \equiv \langle a_x^2\rangle + \langle a_y^2\rangle \tag{1.9}$$
$$Q \equiv \langle a_x^2\rangle - \langle a_y^2\rangle \tag{1.10}$$
$$U \equiv \langle 2a_x a_y \cos(\theta_x - \theta_y)\rangle \tag{1.11}$$
$$V \equiv \langle 2a_x a_y \sin(\theta_x - \theta_y)\rangle, \tag{1.12}$$

where I represents intensity (or temperature), Q and U account for linear polarization and V is the parameter related to circular polarization. Thomson scattering does not produce any circular polarization so in our case the V parameter is equal to 0.

Both I and V represent physical observables and thus they do not depend on the coordinate system of the observer. On the other hand, Q and U describe orthogonal modes of linear polarization and are affected by the coordinate system in which the polarization is measured. It is then useful to find quantities that provide an invariant description of linear polarization.

We know that the Stokes parameters Q and U transform as if they were a two-dimensional symmetric and traceless rank tensor. These tensors can be decomposed in a scalar (curl-free) and a pseudo-scalar (divergence-free) part [17]. We call these parts E and B modes in analogy with the electric field of a point charge and the magnetic field, which are also curl-free and divergence-free respectively. E modes show a radial or tangent polarization pattern while B modes are tilted 45° and change direction depending on the sign, as can be seen in Figure 1.3. E and B modes are defined as linear combinations of Q and U as follows:

$$(Q \pm iU)(\hat{n}) = \sum_{\ell=2}^{\infty}\sum_{m=-\ell}^{\ell} a_{\ell m}^{\pm 2}\,{}_{\pm 2}Y_{\ell m}(\hat{n}) = \sum_{\ell=2}^{\infty}\sum_{m=-\ell}^{\ell} (a_{\ell m}^{E} \pm i a_{\ell m}^{B})\,{}_{\pm 2}Y_{\ell m}(\hat{n}), \tag{1.13}$$

where ${}_{\pm 2}Y_{\ell m}$ are the spin-weighted spherical harmonics [17] and $a_{\ell m}^{\pm 2}$ their respective coefficients. Spin-weighted spherical harmonics are generalizations of the standard spherical harmonics and they fulfill the same orthogonality and completeness relations. This means that the $a_{\ell m}^{\pm 2}$ are computed analogously to (1.3). We can then write $a_{\ell m}^{E}$ and $a_{\ell m}^{B}$ as their linear combinations:

$$a_{\ell m}^{E} = \frac{1}{2}(a_{\ell m}^{+2} + a_{\ell m}^{-2}), \tag{1.14}$$
$$a_{\ell m}^{B} = \frac{-i}{2}(a_{\ell m}^{+2} - a_{\ell m}^{-2}). \tag{1.15}$$

These new coefficients describe two quantities, the E and B polarization modes. They can be defined in the real space in the same way we did with the temperature fluctuations in (1.2):

$$E(\hat{n}) = \sum_{\ell} E_{\ell}(\hat{n}) = \sum_{\ell} \sum_{m} a^{E}_{\ell m} Y_{\ell m}(\hat{n}), \tag{1.16}$$

$$B(\hat{n}) = \sum_{\ell} B_{\ell}(\hat{n}) = \sum_{\ell} \sum_{m} a^{B}_{\ell m} Y_{\ell m}(\hat{n}). \tag{1.17}$$

Each one of the E modes (E_ℓ) has even parity (-1)$^\ell$ while the B modes (B_ℓ) have odd parity (-1)$^{\ell+1}$. Temperature modes (T modes) do also have even parity.

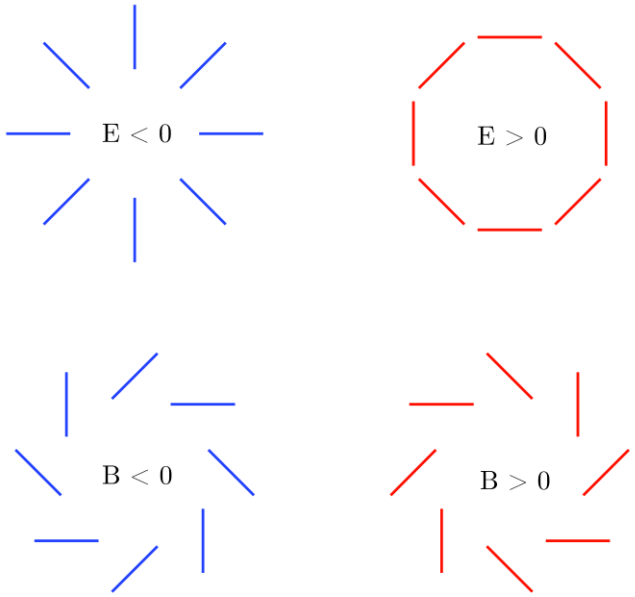

Figure 1.3: E-like and B-like polarization patterns. Figure from [18].

As stated above, it is more interesting to work with E and B since they are scalar fields. We also know that E modes are produced by scalar and tensor perturbations in the primordial Universe, while B modes are solely produced by tensor perturbations. Since primordial tensor perturbations are introduced in the metric by inflation, the detection of primordial B modes would be a smoking gun of an inflationary phase at the beginning of the Universe. The detection of this primordial signal constitutes the main goal of future CMB experiments. This signal is characterized by the tensor-to-scalar ratio, r, which describes the relative intensity of scalar and tensor modes. The existence of an inflationary phase would imply a non-zero r, but the tensor modes might be too faint to be detected. The simplest inflation model up to date [19] predicts a value of $r \approx 0.003$ [20], which could still be measured. Current observations give an upper limit of $r < 0.032$ at 95% confidence [21] so we know that B modes contain much less power than E modes.

There are nonetheless other effects that need to be taken into account regarding the study of B modes. One of the most significant is the polarization transfer between E and B modes due to the distortion produced by gravitational lensing. Astrophysical foregrounds, described in Section 1.3, are also a source of E and B modes that need to be carefully subtracted. In addition, we also need to take care of systematic effects, such as cosmic rays or detector noise, turning primordial B mode detection into an even more challenging mission.

Up to this point, we have defined three scalar fields that describe both the temperature fluctuations (T[6]) and the polarization (E, B). Using the spherical harmonic coefficients that we have derived for each quantity we can define six power spectra (TT, EE, BB, TE, TB, and EB):

$$C_\ell^{XY} = \frac{1}{2\ell+1}\sum_m a_{\ell m}^X \left(a_{\ell m}^Y\right)^*, \qquad X,Y \in (\mathrm{T,E,B}). \tag{1.18}$$

If parity is conserved, then C_ℓ^{TB} and C_ℓ^{EB} power cross spectra are expected to be zero in the ΛCDM model. Thus, in the following, the remaining four spectra (C_ℓ^{TT}, C_ℓ^{EE}, C_ℓ^{BB} and C_ℓ^{EB}) are the ones that we will use to characterize the temperature and polarization fluctuations of the CMB.

These spectra have been measured by different missions, but they can also be obtained analytically using cosmological models. In this project, we have used the Python wrapper[7] of `CAMB` (Code for Anisotropies in the Microwave Background), developed by Anthony Lewis [22], to generate these spectra assuming ΛCDM with the best cosmological coefficients obtained by *Planck* [12]. Given a cosmological model and a set of values of its cosmological parameters, this code obtains the analytical angular power spectra. These spectra can be used later to generate sky simulations of the CMB. An example of the spectra produced by `CAMB` using ΛCDM and the aforementioned parameters is shown in Figure 1.4. In this work we have considered $r = 0$, thus neglecting the existence of primordial B modes.

Although the data observed by the different experiments agrees with the predictions of the ΛCDM model, several anomalies have been detected at large angular scales in temperature. These anomalies include the absence of correlations at large angular scales [23] and the hemispherical power asymmetry [24] among others. Whether these fluctuations have a cosmological origin or are just statistical anomalies is still unknown. Temperature data at large scales obtained by *Planck* are already limited by cosmic variance and thus further study of the anomalies cannot be done by only looking at temperature. Polarization maps, specially the E-mode, provide an independent source to check if these anomalies are really there, and in that case try to get relevant information about their nature. While *Planck* data presents a too low signal-to-noise in polarization at large scales, future experiments are expected to improve these measurements and provide relevant constraints for several of these anomalies [25].

1.3 Foreground emission

One of the main problems regarding the detection of the CMB is the presence of foreground contamination both from galactic and extragalactic sources. Successful removal of these contaminants is fundamental to obtain the CMB angular power spectra and access to all the information about cosmology encoded in the CMB.

The main sources of microwave signal at large scales are diffuse Galactic emissions. Some of the most important phenomena are:

- Synchrotron emission: Synchrotron is the emission that arises from the cosmic ray electrons interacting with magnetic fields and it is the dominant galactic component in polarization at lower frequencies (<70 GHz). A significant fraction of this radiation is linearly polarized and its intensity and spectrum depends both on the intensity of the magnetic field and the energy of the cosmic rays, resulting, unlike the CMB, in a very anisotropic signal.

[6] We will relabel $\Delta T/T$ as T for simplicity.
[7] https://github.com/cmbant/CAMB

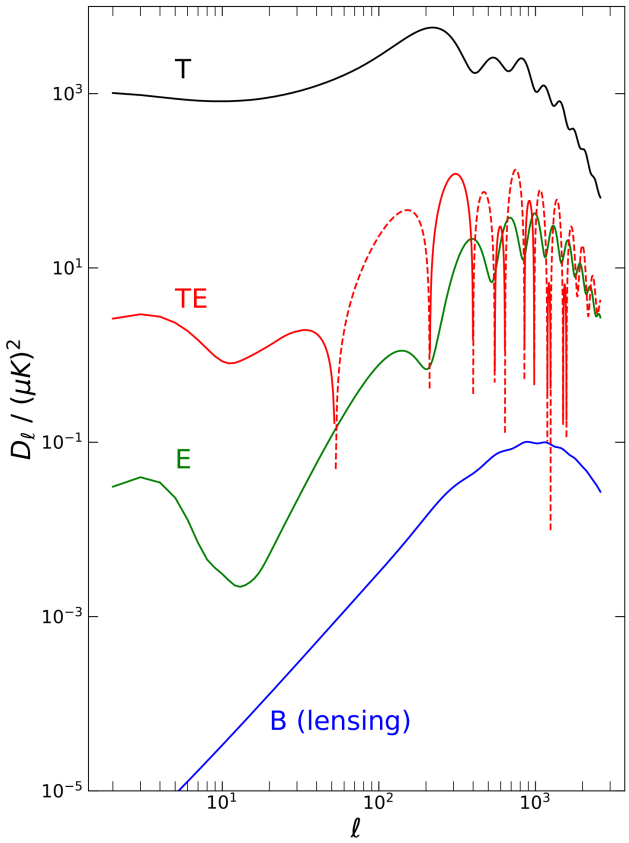

Figure 1.4: CMB angular power spectra produced by `CAMB` using ΛCDM and the parameters from [12]. The tensor-to-scalar ratio r has been set to zero, so B modes correspond solely to lensing B modes. TE spectrum comprises both positive and negative values, so its absolute value is displayed. The positive part of the TE spectrum is represented with solid lines while the negative part corresponds to the dashed lines.

- Free-free emission: Thermal bremsstrahlung, known as free-free emission, is produced by the scattering of free electrons with ions in the interstellar medium. It is also a relevant component in brightness temperature at low frequencies. This emission is intrinsically unpolarized since the interaction is produced isotropically.

- Thermal dust emission: At higher frequencies (>70 GHz) the sky signal is dominated by the thermal dust emission. This contribution can be modeled as a modified black-body with temperatures $T \approx 20$ K modified to account for opacity. Due to the asymmetrical shape of the dust grains, photons are emitted or absorbed along the longest axis. These grains tend to be aligned perpendicularly to magnetic fields, producing then polarized radiation. In this case, the fractions can be up to 20% [26], making this radiation a relevant contribution to polarized microwave sky signal.

- Anomalous Microwave Emission (AME): this emission of unknown origin can be found at lower frequencies. One of the preferred candidates to explain the signal excess observed in the data is spinning dust. It could be produced by tiny dust particles with dipole moment spinning at GHz frequencies. Its spectrum is expected to be highly peaked at the rotation frequency of the smallest particles and its polarization is expected to be low. The most stringent constraints on the polarization fraction, Π, have been provided by [27] with $\Pi < 0.22\%$ at 41 GHz.

The power of these sources, both in intensity and E-mode polarization, is displayed in Figure 1.5. We can see that this contamination leaves only a small frequency window in which the CMB anisotropies

are the dominant microwave source in temperature. In polarization, the situation is worse. At low latitudes, close to the galactic plane, the foregrounds can be up to an order of magnitude greater even at the optimal observation frequencies. This figure highlights the importance of identifying and removing these foregrounds in order to obtain measurements from the CMB.

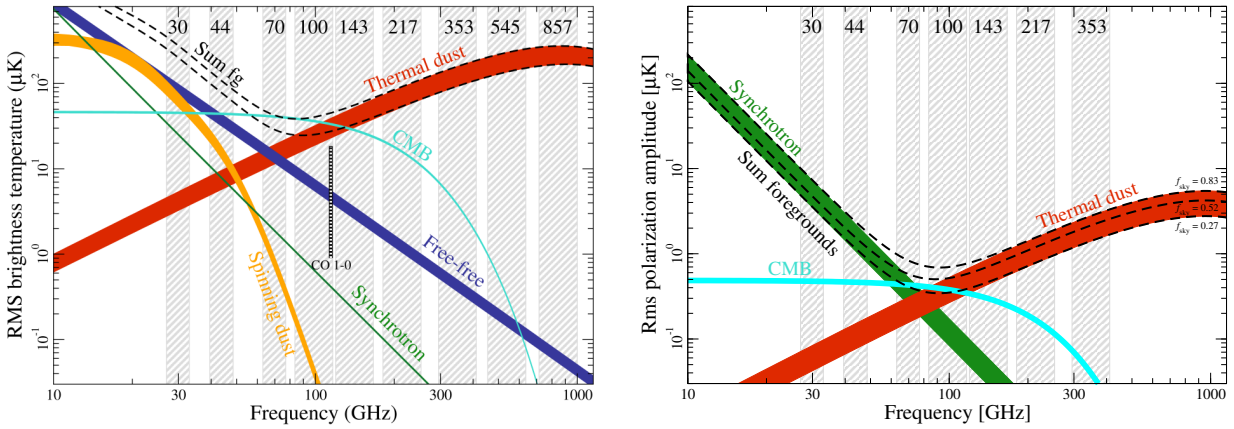

Figure 1.5: Left: Root Mean Square (RMS) brightness temperature of anisotropies in intensity at different frequencies of the CMB and the most important foreground components. Figure from [28]. **Right**: Same plot for the RMS polarization amplitude. Only foreground sources with a significant amount of polarization are displayed. Figure from [29].

These contamination sources present a series of features that allow us to distinguish them from the CMB signal. One of the things that we can take a look at is, as we can see in Figure 1.5, its frequency dependence. The thermodynamic temperature of the CMB is assumed to be constant in its entire frequency range. This property is exploited by some component separation methods in order to isolate and subtract the different foreground components. In the *Planck* mission, four different component separation methods were used:

- Commander: [30] It implements a Bayesian parametric component separation. It uses a parametric model including cosmological, astrophysical, and instrumental parameters, which is then fitted to the data using a MCMC Gibbs sampling algorithm.
- SEVEM: [31] It stands for Spectral Estimation Via Expectation Maximisation. It produces cleaned maps by template fitting in the real space. Templates are obtained by subtracting two smoothed *Planck* maps of two close frequency channels. A linear combination of these templates is latter subtracted to generate a clean CMB map.
- NILC: [32] Needlet Internal Linear Combination. It consists in a linear combination of *Planck*'s different frequency channels such as it minimises variance on a certain wavelet frame called needlets.
- SMICA: [33] Spectral Matching Independent Component Analysis. This method assigns multipole-dependent weights to each *Planck* map by taking into account their spectral covariance matrices among others. A cleaned map is then constructed as a linear combination of the weighted maps.

Another relevant aspect of foreground emission are the non-Gaussian features of the map. The CMB is expected to be Gaussian and current measurements are consistent with this, so studying the statistical properties of the maps can help to extract the CMB, as the foreground emission is highly non-Gaussian. However, this does not unequivocally identify foreground emission since there are also secondary effects that can induce such non-Gaussianities (such as lensing) asides from plausible non-standard inflation (primordial non-Gaussianities).

None of the methods above is able to provide a perfect subtraction of the foreground emission from the CMB. In order to carry out CMB analysis, we need to mask regions where foreground residuals from these methods are too large to be trusted, leaving us with partial CMB sky maps. In the next section we will discuss the issues derived from masking CMB maps as well as some of the methods to recover the angular power spectrum from the missing regions.

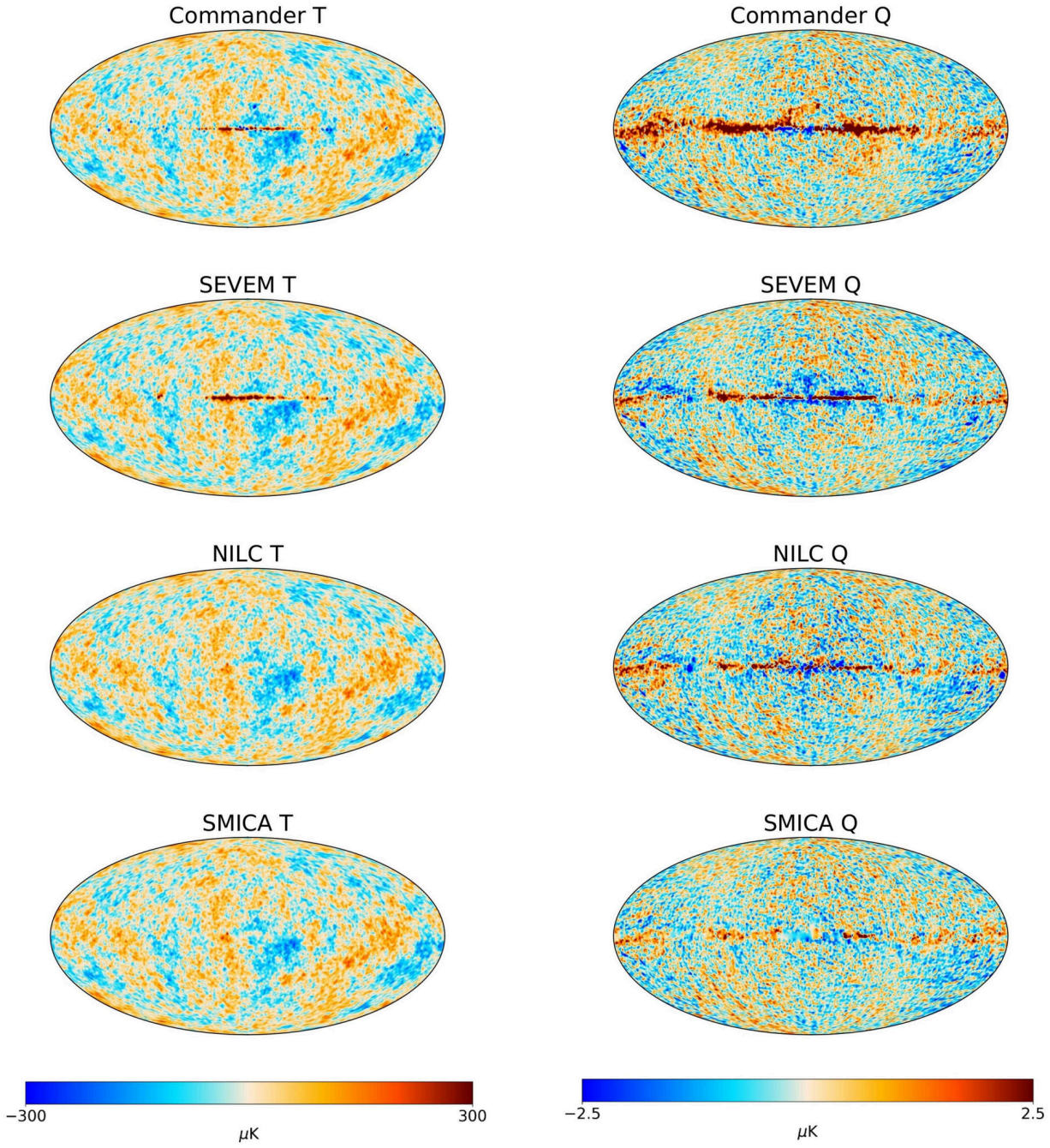

Figure 1.6: T and Q maps from the different component separation methods used in the *Planck* mission, filtered at 80'. Left column shows maps in temperature while right column shows maps in the Stokes parameter Q. Maps obtained from the *Planck* Legacy Archive[8].

1.4 Partial CMB sky maps

The different component separation methods mentioned in (Section 1.3) produce independent CMB maps both in intensity and polarization, shown in Figure 1.6. Each of these maps presents a region where foreground residuals are too large and in consequence have to be masked for cosmological analysis. These leave us with eight different masks (in temperature and polarization) that can be combined into common masks, which are the recommended ones for the analysis of the cleaned CMB maps. These common masks are shown in Figure 1.7. Masks are usually characterized by the fraction of the sky that is left unmasked, also known as f_{sky}. Common masks in Figure 1.7 have f^T_{sky}= 0.780 and f^P_{sky}= 0.782 respectively.

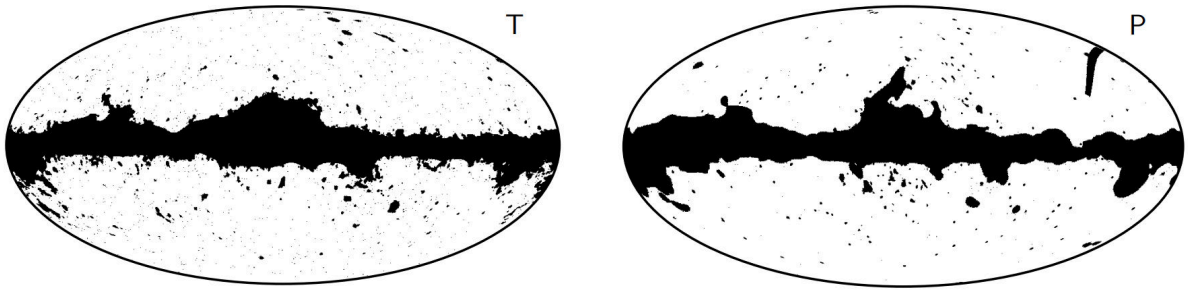

Figure 1.7: *Planck* 2018 common masks for temperature and polarization. Figure from [29].

The use of masks entails some problems, the most evident being the loss of information from the masked regions. This prevents us from obtaining reliable cosmological information on large scales. Besides, trying to reconstruct E and B by projecting an incomplete map of Q and U on the E/B basis gives rise to power leakage between modes. In the CMB, where E modes contain more power than B modes, this power exchange between modes hinders significantly the ability to detect primordial B modes. These effects are also sensitive to the mask choice and it is also important to find a balance between the effects of the mask and the foreground emission we are trying to remove.

There have been proposed several methods to recover CMB's angular power spectra from partial sky coverage. One of them is pseudo-C_ℓ algorithms, as the one described in [34]. These algorithms define a new estimator, $\hat{C}_\ell$, obtained from the observed angular power spectra. These $\hat{C}_\ell$ are then related to the real sky's CMB spectra $\tilde{C}_\ell$ using a coupling matrix that encodes the effects that studying an incomplete sky has on the data. This method has the advantage of being computationally efficient at the cost of producing sub-optimal angular power spectra at large scales. Another proposed method is Quadratic Maximum Likelihood (QML), e.g., the one described in [35]. This method defines a quadratic estimator using models from both the CMB signal and the noise, more specifically using their covariance matrices. QML gives optimal results but it implies storing and inverting large matrices, which translates into a high computation cost.

Another approach to recover information from masked CMB maps are the so-called inpainting techniques. While the pseudo-C_ℓ and QML methods focus on recovering the angular spectra, inpainting aims to reconstruct a complete CMB sky map, which can be useful when looking for large scale anomalies. It consists of filling in the masked region with random signals statistically coherent with the rest of the sky. Some of the different inpainting techniques used in the literature are:

[8] https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/CMB_maps

- Diffuse inpainting: [36] The masked pixels are iteratively filled by averaging over their 8 closest neighbours. This method recovers the largest angular scales and reduces the sharp-edge effects produced in the discontinuity between the mask and the map. It is the simplest and most extended of the techniques. In this work we study and characterize it with more detail in Section 2.2.

- Purified inpainting: [37] The aim of this technique is to reconstruct a full-sky polarization map $\tilde{p}$ from a masked one $M \cdot p$, where the reconstructed map is equal to p in the unmasked region. In the masked region, modes are assigned either to the E and B subspaces or to an "ambiguous" assignation where some mode mixing is allowed. Initially all modes are assigned to this ambiguous mode. Then, non-ambiguous modes are separated one by one using a Krylov subspace [38]. The masked region is finally inpainted by solving $\nabla^2 \mathbf{a} = 0$ inside the mask, where $\mathbf{a}$ corresponds to the ambiguous mode map. This methods obtains results on pair with the ones from maximum-likelihood methods at a much lower computation cost.

- Constrained realizations: [39] This method assumes a certain cosmological model and inpaints the masked region using information from outside the mask. Once defined a probabilistic model based on the chosen cosmology, the inpainting is done by calculating the conditional probability density $p(\hat{\mathbf{d}}|\mathbf{d})$, where $\hat{\mathbf{d}}$ is the inpainted field and $\mathbf{d}$ the observed data. This probability density is a Gaussian distribution, whose mean and variance can be obtained from the pixel covariance matrix of the map, which in turn is obtained from its angular power spectrum C_ℓ. This procedure achieves optimal results at the cost of being computationally expensive, making it only suitable for low resolution maps. It also relies on choosing the correct cosmological model.

- Convolutional Neural Network: [40] Both trained and untrained neural networks (NN) are used in image restoration and can be equally applied to the reconstruction of CMB maps. In the case of trained NN, one possible approach, known as generative adversarial network (GAN), is to use both a generator and a discriminator. The generator is trained to generate fake samples resembling a certain training set, while the discriminator is trained to distinguish between fake maps and the ones from the sample set. Both networks are trained until the generator is able to generate inpainted maps that the discriminator is not able to tell if they are generated or not. In the case of untrained NN, it is assumed that the information necessary to reconstruct the map is contained in the masked map and the architecture of the NN. Given some input the image is processed and filled by fitting some NN parameters. Both of these methods are aimed to inpaint point-sources that affect the small angular scales.

Inpainting is a valuable tool in the study of the CMB at the largest scales, i.e. low ℓ. As pointed in Section 1.2, there exist some anomalies in temperature at these large scales that are in tension with the predictions of the ΛCDM model. It is then important to find a technique that allows us to reconstruct full CMB maps from observations, providing reliable inpaintings while maintaining a reasonable computational cost. The recovery of E maps is of special interest, since they can be used as an independent probe for the existence of these anomalies.

In this project we propose a novel inpainting technique, a combination of the diffuse inpainting with the multiresolution approach of `HEALPix` wavelet scheme. This method allows us to improve the range of angular scales at which the diffuse inpainting is effective, as well as to reduce computing times. In the following chapter we provide a detailed description of both the traditional diffuse inpainting and this new technique. We also describe the methodology used in its characterization, as well as the parameters used to simulate and work with CMB observations.

Chapter 2

Methodology

In this chapter, we describe and implement two different inpainting techniques: the diffuse inpainting and an original multiresolution wavelet inpainting. Regarding the latter, we propose two different approaches: one reconstructed using model-constrained details and another using observation-constrained details.

We work with CMB maps generated fron the angular power spectra obtained with `CAMB` using ΛCDM and the default parameters from *Planck*[1]. These parameters set the tensor-to-scalar parameter r to 0, so the only B-modes are E- to B- lensing modes. The maps are generated at a `HEALPix` resolution (defined in Section 2.1) of $N_{side} = 128$. This choice corresponds to a resolution at which the inpainting from constrained realizations are computationally too expensive. At this N_{side} we are working with maps with $N_{pix} = 12(N_{side})^2 = 196\,608$ and a maximum multipole $\ell_{max} = 383$.

All of the maps have been smoothed with a Gaussian symmetric beam with a Full Width Half Maximum (FWHM) equal to:

$$\text{FWHM} = 2.4 \times \sqrt{\frac{4\pi}{N_{pix}}}, \tag{2.1}$$

which, for $N_{side} = 128$, is equal to $\sim 1°$.

As for the mask choice, we apply the common temperature mask shown in Figure (1.7), obtained from *Planck* Legacy Archive[2]. Since we are working with simulated maps with no foreground, this mask is valid for both temperature and polarization maps.

In this project we will obtain inpainted maps using different techniques. In order to compare the goodness-of-fit of these inpainting methods we use a statistical estimator, $\tilde{\chi}$, defined as:

$$\tilde{\chi}(\ell_{max}) = \sum_{2}^{\ell_{max}} w_\ell \Delta_\ell, \tag{2.2}$$

where:

$$\Delta_\ell = \left| C_\ell^{\text{rec}} - C_\ell^{\text{ref}} \right| \qquad \text{and} \qquad w_\ell = (2\ell + 1), \tag{2.3}$$

[1] https://github.com/cmbant/CAMB/blob/master/inifiles/planck_2018.ini
[2] http://pla.esac.esa.int/pla/#maps

where C_ℓ^{rec} is the recovered angular power spectrum and C_ℓ^{ref} is some reference power spectrum. This estimator is just the sum up to some ℓ_{max} of the residuals with a linear weighting. We omit from this sum the monopole and the dipole. The weights used give more importance to higher multipoles, where a greater number of m modes are available. In fact, for a given ℓ there are $(2\ell + 1)$ m modes, so these weights can also be defined as the number of available m modes in each multipole.

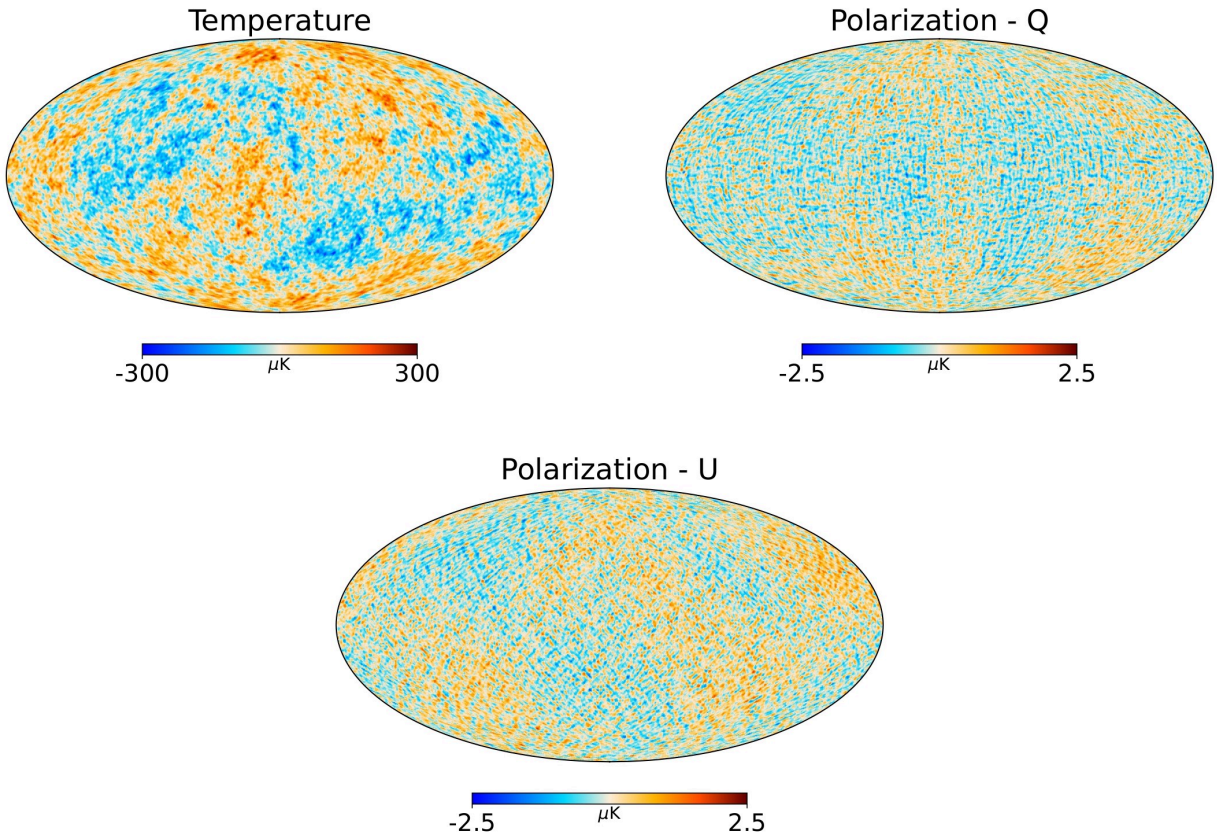

Figure 2.1: T, Q, and U CMB maps produced by `CAMB` using ΛCDM and the best cosmological parameters inferred from *Planck* data. They are generated at a resolution $N_{side} = 128$ and filtered using a Gaussian beam with FWHM = 66'.

2.1 `HEALPix`

Throughout this work we will make use of `HEALPix` (Hierarchical Equal Area isoLatitude Pixelization) [41]. This pixelization scheme aims to provide a tool for handling pixelated data on the sphere. To do so, the sky is divided into pixels of equal area. At the lowest resolution this division consists of 12 pixels, called parent pixels. Each subsequent increase in resolution splits each pixel into 4 pixels, giving a total number of pixels $N_{pix} = 12 \cdot N_{side}^2$. The N_{side} parameter is a power of 2 that characterizes the resolution of the map. Figure 2.2 illustrates this scheme. In order to work with this pixelization we use `healpy` [42], the Python wrapper for `HEALPix`.

Given a certain N_{side} value, the effective angular resolution of a pixel side can be approximated by:

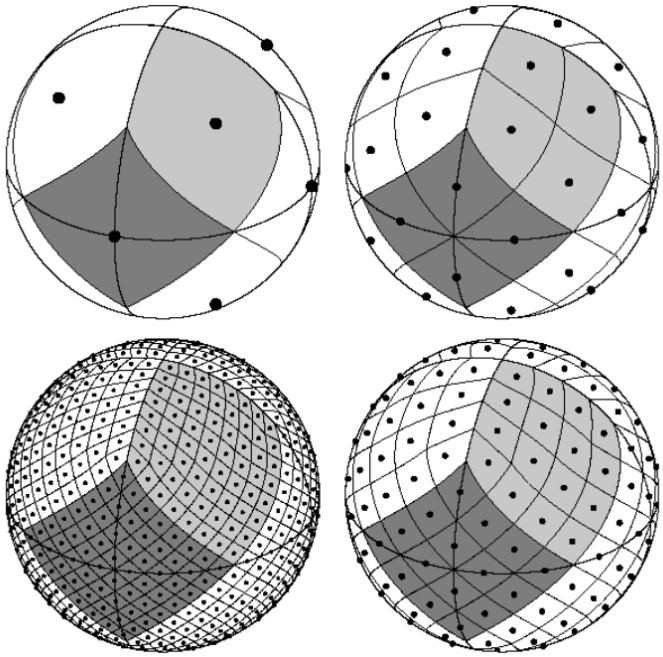

Figure 2.2: Clockwise from top left: `HEALPix` pixelation of the sphere for resolutions $N_{side} = 1$, 2, 4, and 8. The total number of pixels in each sphere are 12, 48, 192 and 768. Pixels in each color are originated from the same parent pixel. Figure from [41].

$$\theta \approx \sqrt{\frac{4\pi}{N_{pix}}}. \tag{2.4}$$

Due to aliasing, the maximum multipole that can be resolved in a map is given by $\ell_{max} = 3N_{side} - 1$.

These CMB maps are stored as arrays where each value corresponds to one position in the sky. These values can be indexed following one of two ordering schemes. The RING ordering enumerates pixels in horizontal rings starting from the North pole. On the other hand, the NESTED ordering enumerates pixels by grouping them inside big parent pixels, as it is shown in Figure 2.3. This NESTED ordering is specially useful due to its synergy with scaling operations to the map.

The `HEALPix` tessellation makes working with spherical harmonics a much easier task. In addition to the representation features of the package, `healpy` also includes some useful functions to perform spherical harmonic analysis of the maps. Given a temperature or polarization map, the *anafast* function is able to compute its angular power spectra up to a maximum multipole ℓ_{max}, with its operation count scaling as $O(N_{pix}^{1/2} l_{max}^2)$. The inverse process can be done by making use of the *synfast* function. Using the C_ℓ's as an input, *synfast* computes a set of $a_{\ell m}$ using a random seed. These $a_{\ell m}$ are then used to create CMB maps. The temporal complexity is again $O(N_{pix}^{1/2} l_{max}^2)$. These functions are able to work both with 3-tuples of arrays (representing T, Q, and U maps) and with individual maps (T, E, or B maps), and we will make extended use of them in the following sections.

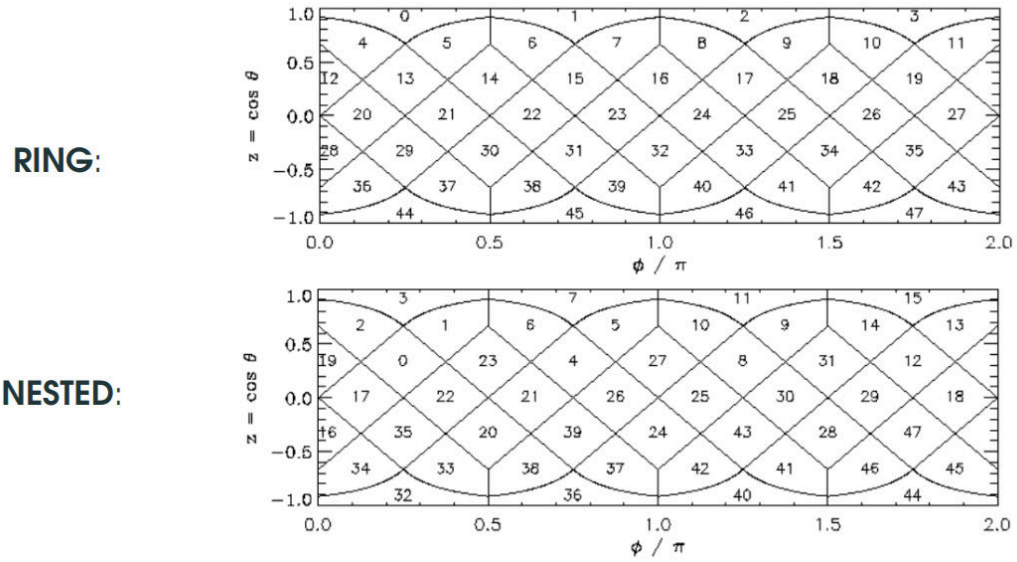

Figure 2.3: RING and NESTED pixel ordering schemes from `HEALPix`. Figure from [41].

2.1.1 `HEALPix` wavelet

Wavelets are localized wave functions that allow for a multi-resolution treatment of data. In this work we will make use of the `HEALPix` wavelet, described in [43]. This wavelet takes advantage of the `HEALPix` scheme described above to carry out this multi-resolution analysis while reducing computational time. Using this method we can decompose a CMB map in the wavelet coefficient space in a series of maps with different resolutions j, from the original map's resolution J to a desired lower limit j_0, where the N_{side} of each map is given by $N_{side} = 2^j$.

`HEALPix` wavelet functions are defined as:

$$\begin{aligned}
\Psi_{0,j,k}(x) &= \phi_{j+1,k_0}(x) - \frac{\phi_{j,k}(x)}{4}, \\
\Psi_{1,j,k}(x) &= \phi_{j+1,k_1}(x) - \frac{\phi_{j,k}(x)}{4}, \\
\Psi_{2,j,k}(x) &= \phi_{j+1,k_2}(x) - \frac{\phi_{j,k}(x)}{4}, \\
\Psi_{3,j,k}(x) &= \phi_{j+1,k_3}(x) - \frac{\phi_{j,k}(x)}{4},
\end{aligned} \tag{2.5}$$

where

$$\phi_{j,k}(x) = \begin{cases} 1 \text{ if } x \in P_{j,k} \\ 0 \text{ otherwise,} \end{cases} \tag{2.6}$$

and $P_{j,k}$ refers to the pixel at position k at resolution j.

Wavelet decomposition of a temperature map (the same would apply for polarization) can be

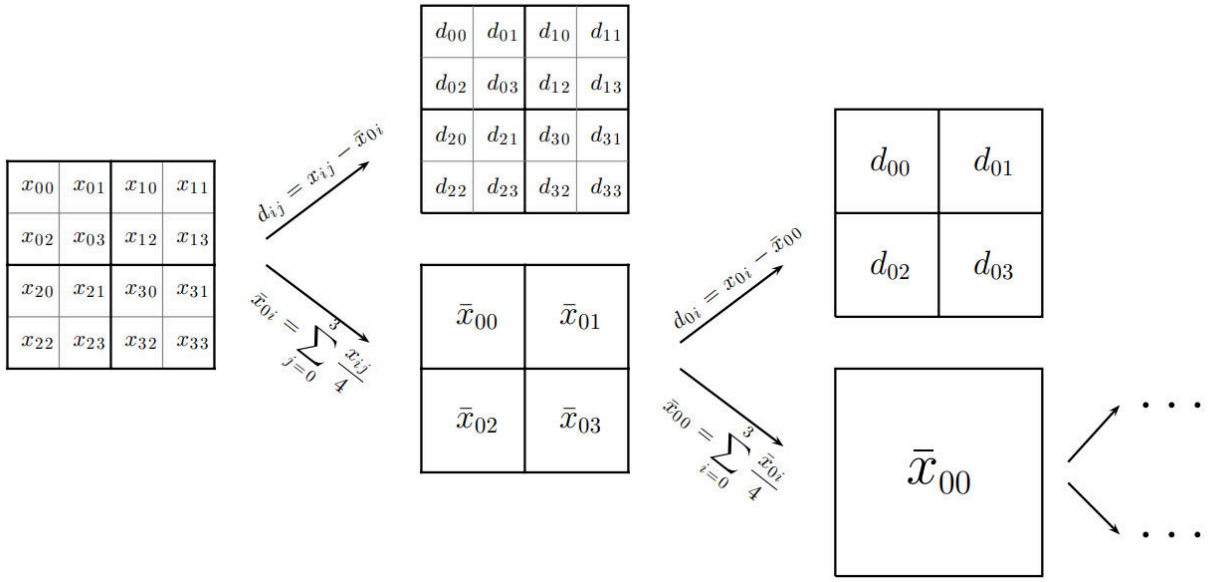

Figure 2.4: Construction of the approximation and detail coefficients using the `HEALPix` wavelet. Approximation coefficients x are obtained as the average of their 4 children pixels. Detail coefficients d are obtained as the difference between the original pixels and that average. Figure from [43].

written in terms of these functions as:

$$\frac{\Delta T}{T}(x_i) = \sum_{k=0}^{N_{j_0}-1} \lambda_{j_0,k}\phi_{j_0,k}(x_i) + \sum_{j=j_0}^{J}\sum_{m=0}^{3}\sum_{k=0}^{N_j-1} \gamma_{m,j,k}\Psi_{m,j,k}(x_i), \tag{2.7}$$

where N_j is the number of pixels at a resolution j, and $\lambda_{j,k}$ and $\gamma_{m,j,k}$ are the approximation and detail coefficients.

Approximation coefficients at any resolution are obtained just by downgrading the map at the resolution immediately below. In the `HEALPix` scheme this corresponds to averaging the four daughter pixels that correspond to each coefficient:

$$\lambda_{j,k} = \frac{1}{4}\sum_{i=0}^{3} \lambda_{j+1,k_i}. \tag{2.8}$$

On the other hand, detail coefficients at a resolution j are obtained by subtracting the approximation at resolution j from the resolution at $j+1$:

$$\begin{aligned}
\gamma_{0,j,k} &= \lambda_{j+1,k_0} - 4\lambda_{j,k} \\
\gamma_{1,j,k} &= \lambda_{j+1,k_1} - 4\lambda_{j,k} \\
\gamma_{2,j,k} &= \lambda_{j+1,k_2} - 4\lambda_{j,k} \\
\gamma_{3,j,k} &= \lambda_{j+1,k_3} - 4\lambda_{j,k}
\end{aligned} \tag{2.9}$$

Figure 2.4 illustrates this decomposition process. What we finally obtain is an approximation map with resolution j_0 and $(J - j_0)$ detail maps at all the intermediate resolutions. These detail maps store the information that would be otherwise lost with each resolution downgrade of the map. An example of the products of this decomposition can be seen in Appendix A for the case of initial $N_{side} = 128$ $(J = 7)$ and final $N_{side} = 8$ $(j_0 = 3)$.

This wavelet decomposition can be also applied to masked maps. In that case, we perform also the wavelet decomposition of the mask. Since pixels in the mask can only be masked or unmasked (0 or 1), we need to set a threshold t such as pixels with values larger than t will be considered unmasked, while the rest will be treated as being masked. In this work we will set this threshold to 0.5. These mask approximations can then be applied to the decomposition at each resolution. An example of how the masks changes at each resolution can also be seen in Appendix A.

2.2 Diffuse inpainting

We implement a slightly different variation of the simple diffuse inpainting using the Jacobi method [44]. The diffuse inpainting consists in substituting each of the pixels in the mask with the mean value of its 8 closest pixels. In the Jacobi method, values of these closest pixels are taken from the previous iteration. This process is repeated over several iterations until the mask is completely filled. 2000 iterations have been proposed in the literature ([36],[45]) as an acceptable number to ensure convergence. In our study, however, we have set a convergence criteria in order to reduce the number of iterations in each inpainting. We consider the inpainting as completed when the mean relative change in the inpainted pixels between iterations is less than 10^{-3}:

$$\frac{1}{N_{pix}} \sum_{i=0}^{N_{pix}} \frac{x_i(j) - x_i(j-1)}{x_i(j-1)} < 10^{-3}, \tag{2.10}$$

where $x_i(j)$ is the value the i-th inpainted pixel at iteration j. This criteria helps us to reduce the number of iterations and thus the time it takes to run the code.

This method proves to be effective when working on temperature maps. However, averaging the 8 closest pixels is an arbitrary choice, and it turns out to be less effective in polarization maps. To deal with this issue, we propose averaging over the pixels that lie inside a disk centered in each of the pixels of the mask, instead of using just the eight closest neighbours. The optimal radius of this inpainting disk is studied in Section 3.1.

2.3 Wavelet diffuse inpainting

We propose now a new inpainting technique: a combination of the simple diffuse inpainting described above with the `HEALPix` wavelet decomposition introduced in Section 2.1.1. Starting from a masked CMB map at $N_{side} = 2^J$, in our case $N_{side} = 128$, we perform its wavelet decomposition up to a resolution $N_{side} = 2^{j_0}$, obtaining one approximation at j_0 and (J-j_0) detail maps. The same diffuse inpainting technique described above is applied to the approximation map, while we fill the masked details using two different approaches. The map is then reconstructed using the inpainted approximation and the filled details, thus recovering a complete CMB sky map.

The motivation behind this method is to combine the large scale reconstruction of the angular power spectra, attained with the diffuse inpainting, with the small scale information, provided by the details in the `HEALPix` wavelet scheme. The problem now is to find a suitable way of filling the masked details. In this work, we propose two diffferent methods, which are explained below.

2.3.1 Model-constrained details

Our first approach is to fill in the masked details with the details of a simulated sky realization. The general process is as follows:

- We start from a masked CMB map, which we will label as the observed sky. Given this sky, we generate a CMB realization using the best-fitting cosmological parameters for the observed sky. We will label this realization as the theoretical sky. In this work, since we are using simulations as the observed skies, both the observed and the theoretical sky are generated using same parameters.
- We then decompose both maps in wavelets up to a minimum N_{side}. We do the same with the mask and use it to mask the observed sky, both the approximation and its details. An example of the products of this decomposition is shown Appendix A.
- We fill in the masked region in the approximation by performing the diffuse inpainting described in Section 2.2.
- The masked region in the detail maps are filled by substituting the corresponding pixels from the theoretical sky detail maps. We then reconstruct a complete CMB sky summing over the inpainted approximation and the filled details.

In our case, these simulated sky realizations are obtained from the same angular power spectra using `CAMB` and therefore are constrained by the model used to create them, namely ΛCDM with the parameters from [12].

Since the inpainting is done at a low resolution, the number of pixels to be inpainted is greatly reduced with respect to the simple diffuse inpainting. The computational time needed for `HEALPix` wavelet decomposition scales as $O(N_{pix})$, rendering this method much faster and more efficient than using constrained realizations. However, the cosmological model dependence is still present since we need some theoretical details to fill in the masked detail maps.

In this method, two parameters need to be set: the minimum N_{side} in the decomposition and the inpainting disk used. As in the case of the diffuse inpainting, we study the $\tilde{\chi}$ value for a great sample of inpainted maps in order to find the optimal values. These results are shown in Section 3.2 and Section 3.3

2.3.2 Observation-constrained details

In this second approach, we try to remove the cosmological model dependence by filling in the masked details with random Gaussian fields of the same variance as the details. As in the previous case, we simulate an observed map and perform a wavelet decomposition both of the map and the mask, applying the corresponding mask approximation at each resolution. The observed map approximation is then inpainted using the same diffuse inpainting as in the previous sections.

The idea now is to replace the masked region of each detail map with random Gaussian numbers of zero mean and variance equal to the variance of the unmasked details. However, this random

noise does not reproduce the correlations that can be found in the details. It is shown in Figure 2.5 how the angular power spectrum of the details does not correspond with the one of a map filled with random Gaussian noise. In order to account for this feature, we will need to rescale the power spectrum of the Gaussian field. We will call this approach observation-constrained details, since we fill in the masked detail maps by looking at the mean, variance and spectra of the observed detail maps.

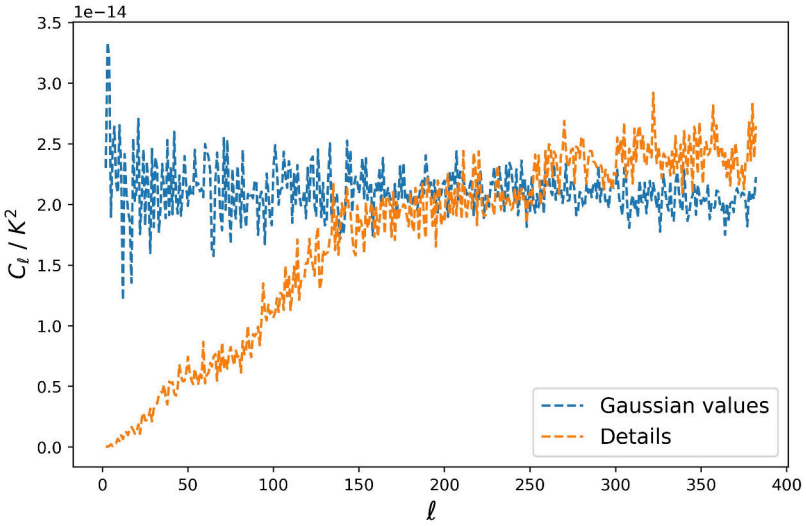

Figure 2.5: TT angular power spectrum of a detail map at $N_{side} = 128$ vs a detail map made out entirely of Gaussian random numbers with the same mean and variance as the details.

The process is as follows:

- To avoid correlations between different scales, we generate a vector of $(N^{128}_{m,pix} + N^{64}_{m,pix} + N^{32}_{m,pix} + N^{16}_{m,pix})$ random Gaussian values with $\mu = 0$ and $\sigma = 1$, where $N^{i}_{m,pix}$ is the number of masked pixels at $N_{side} = i$.
- For each N_{side}, we calculate the deviation of the unmasked pixels in the detail map, σ_i. We then multiply the first $N^{128}_{m,pix}$ random numbers by σ_{128}, the next $N^{64}_{m,pix}$ random numbers by σ_{64}, and so on. Doing so gives us an ordered vector with four groups of random numbers, each group with the same variance as the detail map we want to fill. We now substitute the masked pixels at each resolution with the corresponding group of random numbers.
- The next step is to correct the filled map spectrum. Using `healpy`'s *anafast*, we obtain the angular power spectrum of the masked detail maps C^{obs}_ℓ and of the filled details C^{fill}_ℓ. We also obtain the $a_{\ell m}$ coefficients of the filled details, $a^{fill}_{\ell m}$.
- We now apply the following correction to those coefficients:

$$a^{correct.}_{\ell m} = a^{fill}_{\ell m} \cdot \sqrt{\frac{C^{obs}_\ell / f_{sky}}{C^{fill}_\ell}}, \tag{2.11}$$

where we account for the masked region of the details by multiplying its angular power spectra with f^{-1}_{sky}. If we recall how we obtained the angular power spectrum from the $a_{\ell m}$ coefficients

(1.5), we can see that this correction modifies the angular spectrum of the filled details to make it match exactly with the spectrum of the observed details.

- Finally, we generate a map of corrected details using the $a_{\ell m}^{correct.}$. We make our final details by using the observed details outside of the mask and the corrected details inside the mask.

Multiplying the observed C_ℓ^{obs} by f_{sky}^{-1} can be interpreted as a zeroth-order correction towards recovering the full sky power spectrum from partial skies. The pseudo-C_ℓ and the QML methods mentioned in Section 1.4 are some examples of how to obtain better approximations of the full sky spectrum.

Following this method, we are able to fill in the masked details with Gaussian values of the same mean and deviation, which also possess an angular power spectrum similar to the one from the original details.The computational time is still much shorter than using constrained realizations, and we also got rid of the cosmological model dependence.

Chapter 3

Results

In this chapter we present the results obtained from the different methods described above. We first characterize the parameters of both inpainting methods, i.e. the optimal values for the inpainting disk radius in the diffuse inpainting, and for the minimum N_{side} and disk radius in the wavelet diffuse inpainting Section 3.1. Once we have obtained the parameters for each of the techniques we proceed to compare them in Section 3.2. This comparison is done at three levels. First we perform a visual comparison at a map reconstruction level by showing T, Q, U, E and B maps for the diffuse inpainting and the two approaches of the wavelet diffuse inpainting. We then analyze the recovered spectra by using the $\tilde{\chi}$ estimator defined in eq. (2.2), in the previous chapter. Lastly, we check the existence of correlations between the multipoles in the different techniques. In Section 3.3 we apply the diffuse inpainting and the wavelet diffuse inpainting to *Planck* temperature data, and compare it to the *Planck* Collaboration inpainted map.

3.1 Parameters

In this section we study the optimal inpainting parameters for the diffuse inpainting and the wavelet diffuse inpainting. We make use of the $\tilde{\chi}$ estimator defined previously to compare the quality of the inpainting in different maps as a function of the inpainting disk radius in the diffuse inpainting, and the minimum N_{side} and disk radius in the wavelet diffuse inpainting.

3.1.1 Diffuse inpainting

The classical diffuse inpainting averages over the 8 closest neighbours of each pixel in the mask. This methodology proves to be non-optimal when working with polarization maps. The proposed alternative is to average instead over pixels located inside a disk of a certain radius centered in the masked pixel.

In Figure 3.1 we show the ratio of χ/χ_m for C_ℓ^{TT}, C_ℓ^{EE}, and C_ℓ^{BB}. χ corresponds to the $\tilde{\chi}$ value of the maps recovered using different disk sizes in the inpainting. χ_m refers to the $\tilde{\chi}$ value for the angular power spectra of the masked map, corrected by f_{sky}^{-1}[1]. In both cases the reference power spectrum used to compute the χ value was the one of the complete CMB map. By plotting the ratio χ/χ_m we can see whether the inpainted map outperforms the simple f_{sky}^{-1} correction ($\chi/\chi_m < 1$) or not.

Since this method only tries to reconstruct the angular power spectra at the largest scales, this χ

[1] As mentioned in the previous chapter, multiplying the angular power spectrum by f_{sky}^{-1} constitues a zeroth-order correction towards recovering the full sky power spectrum.

value is calculated up to $\ell_{max} = 25$. We will see in the following sections how diffuse inpainting gets outperformed by the other methods when studying smaller scales. The number of inpainted maps with each disk is N=100 and the map resolution is $N_{side} = 128$. In this figure, the optimal disk size is painted in red and corresponds to the lowest χ/χ_m value, which translates into the smallest residuals when compared to the original map. The temperature angular power spectrum is best recovered at large scales when taking only a small radius of 1°, close to the pixel resolution at this N_{side}. For polarization, larger disks produce better outcomes. Since the inpainting is produced in the Q and U maps, where E is the dominant component, we will apply the diffuse inpainting using a disk of radius equal to 3° to the maps in polarization. In Figure 3.1 we can see how the optimal disk radius does not converge for the BB spectrum. As we will see in the next section, the recovered spectra for BB are rather poor, so we will instead focus on the results for EE. Besides, going into larger inpainting disks becomes increasingly time-consuming, as larger disks require more pixels to be averaged in each iteration. As we can see in the figure, the ratios χ/χ_m are always below 1, meaning that, at large scales, applying diffuse inpainting is always better than just leaving the pixels masked and then applying a f_{sky}^{-1} correction.

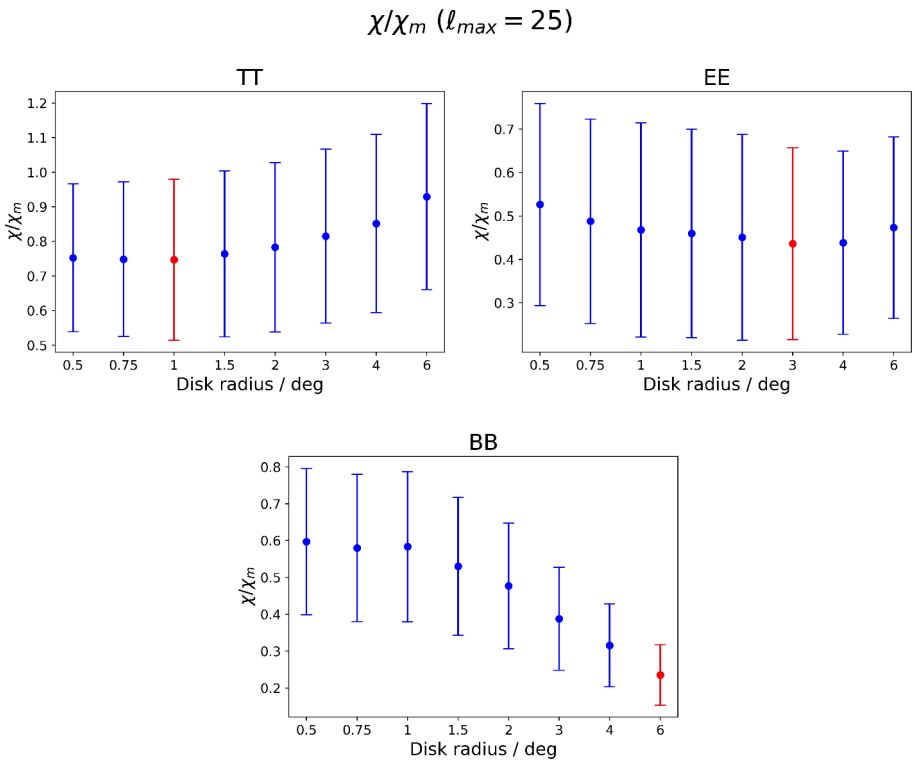

Figure 3.1: Mean χ/χ_m values and their deviation of diffuse inpainted CMB maps as a function of the disk size used in the inpainting. Red dots represent the optimal inpainting disk in order to recover each of the power spectra.

3.1.2 Wavelet diffuse inpainting

First we want to find the optimal value at which to stop the decomposition during the wavelet diffuse inpainting. This value is the N_{side} of the approximation and, therefore, the resolution at which the inpainting is done.

To obtain this value we calculate the $\tilde{\chi}$ value of spectra reconstructed using the diffuse wavelet inpainting at several minimum N_{side} values. As in the diffuse inpainting, we plot the ratio χ/χ_m in order to study the performance of the technique against the masked spectra. Results are similar for the approach using model-constrained details Section 2.3.1 and for the observation-constrained details Section 2.3.2, so we will only show results using model-constrained details for the sake of simplicity.

We study a total of N=100 maps for each approximation resolution. We set the 8 nearest neighbours as a common inpainting disk for all the resolutions. In this case, the value of χ has been computed up to the maximum multipole available for the original map resolution. Results are shown in Figure 3.2. The optimal minimum N_{side} is 8 both for temperature and polarization maps. The obtained χ/χ_m ratios are not as good as the ones shown in Figure 3.1 because here we are reaching up to $\ell = 383$, instead of limiting ourselves to the best-performing scales as in the previous analysis.

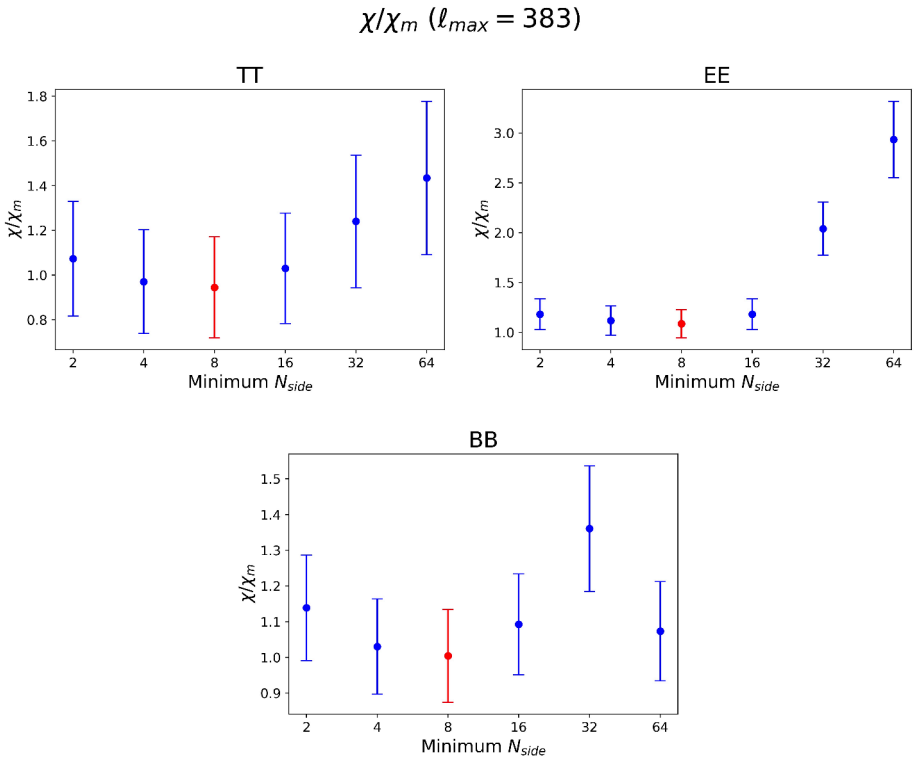

Figure 3.2: Mean χ/χ_m values and their deviation of N=100 inpainted maps using wavelet decomposition as a function of the minimum N_{side} of the decomposition. Red dots indicate the minimum χ value, and thus the best inpainting.

Decomposing up to that $N_{side} = 8$ makes sense if one thinks about how diffuse inpainting and wavelet decomposition combine. As we will see in Section 3.2, diffuse inpainting is only able to recover the largest scales of the angular power spectrum (ℓ<25). On the other hand, details at resolution j store the information lost when downgrading from $N_{side} = 2^j$ to $N_{side} = 2^{j-1}$. If we get up to $N_{side} = 8$, the first details that we will have, are the details at $N_{side} = 16$. These details store the information of the low scales that $N_{side} = 16$ has but not $N_{side} = 8$, i.e. multipoles from $\ell = 23$ to $\ell = 47$. Details at $N_{side} = 32$ cover from $\ell = 47$ to $\ell = 95$ and so on. This means that going up to $N_{side} = 8$ allows us to recover information in $\ell < 25$ thanks to the diffuse inpainting, and

information in $\ell > 23$ using the details from the wavelet, thus covering the entire angular spectrum. Stopping at a larger N_{side} would leave an interval of multipoles unrecovered, while stopping at a smaller resolution might leave out too much information for the diffuse inpainting to be as effective.

Once the optimal N_{side} of the approximation in the wavelet diffuse inpainting has been set to $N_{side} = 8$, we need to perform the same study as in Section 3.1.1 about the inpainting disk size.

We study the χ/χ_m value of N=100 maps painted using the wavelet diffuse inpainting, where different inpainting disks have been used in the approximation. The results are shown in Figure 3.3. In this case, the inpainting disks are much larger than in the simple diffuse inpainting. We need to take into account that, as we are inpainting at $N_{side} = 8$, the effective angular resolution of each pixel is $\theta \approx 7°$ and therefore larger disks are needed. It is important to note that working at such a low N_{side} implies dealing with much less masked pixels. This makes the inpainting process much faster at $N_{side} = 8$ than at $N_{side} = 128$, making the wavelet inpainting an improvement over the diffuse inpainting in terms of computing time.

Since we are working with larger pixels than in $N_{side} = 128$, taking only the closest neighbours is

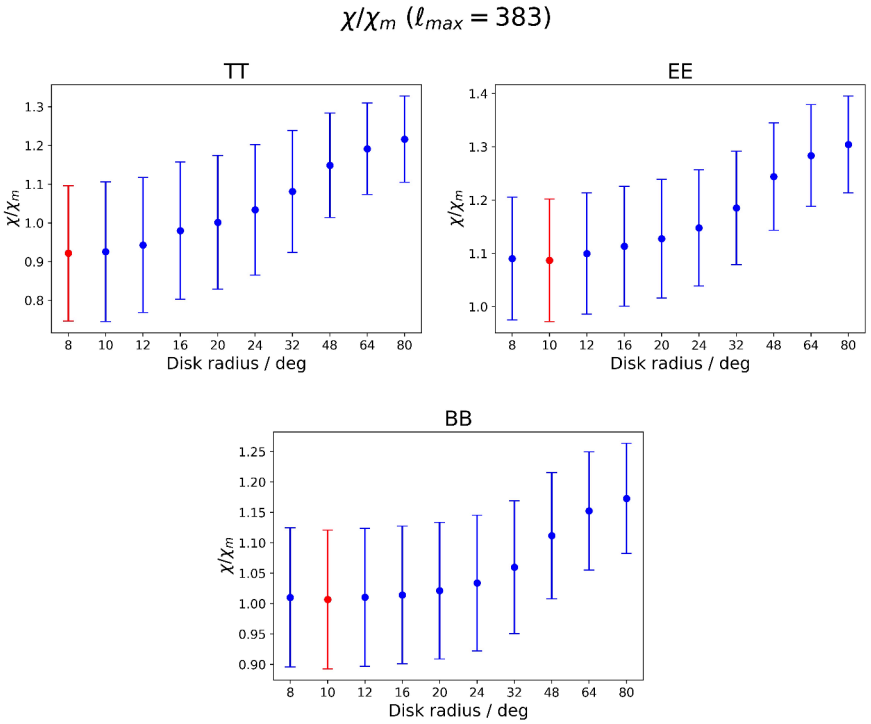

Figure 3.3: Mean χ/χ_m values and their deviation of N=100 inpainted maps using wavelet decomposition as a function of the disk size used in the diffuse inpainting of the approximation map at $N_{side} = 8$. Red dots indicate the minimum χ value, and thus the best inpainting.

already enough to capture the information at great scales. The differences between disk sizes are minimal so we will inpaint using only the 8 closest neighbours.

In contrast to the diffuse inpainting, in this case the ratio χ/χ_m is close to 1 or even above. The main difference is that here we are considering all the available multipoles, up to $\ell = 383$. We will see in Figure 3.9 how both EE and BB are best recovered at lower multipoles. This means that this technique improves the f_{sky}^{-1} correction of the masked angular power spectra but only up to a certain scale.

3.2 Method comparison

In this section we perform a comparison between the simple diffuse inpainting and our new method, the wavelet diffuse inpainting. This comparison consists of three parts: visual reconstruction of the CMB maps, angular power spectra recovery and study of correlations bewteen multipoles.

3.2.1 Inpainted maps

The maps obtained after applying diffuse inpainting and wavelet inpainting with the two detail approaches are shown in Figure 3.4. The inpaintings have been done in T, Q and U maps using the optimal parameters determined in the previous section. E and B polarization maps are constructed using Q and U maps by obtaining a_{lm}^E and a_{lm}^B and converting them back to E and B maps as described in equations (1.16) and (1.17).

The left column of Figure 3.4 corresponds to maps inpainted using the simple diffuse inpainting. In these maps the inpainted region is easily distinguished and features a blurry reconstruction of the map at larger scales. Looking at the E-mode map we can see how the blurry region from Q and U carries on. However, the B-mode map exhibits a saturated region in the boundaries of the galactic mask and in the masked point sources. This saturated region also seems to contaminate the inner part of the mask. This saturation is caused by the power leakage between E- and B-modes due to the anisotropies present in the reconstructed Q and U maps. Overall, the map is roughly recovered at large scales for T and E, while the inpainted B map is affected by E-to-B power leakage.

The center column of Figure 3.4 displays the results for the wavelet diffuse inpainting when using model-constrained details (Section 2.3.1). In this case, the inpainted region cannot be distinguished from the overall map and the reconstructed T, Q, U, and E maps are visually identical to a full CMB sky map. B-mode map is still affected by anisotropies in Q and U, giving rise to the same excess of power as in the simple diffuse case.

The right column corresponds to the results for the wavelet diffuse inpainting when using observation-constrained details (Section 2.3.2). The maps obtained with this approach, although more detailed than the diffuse ones, still show a distinguishable inpainted region in temperature and in polarization. Power leakage in the B map is even worse than in previous cases. From this visual comparison, this approach of the wavelet diffuse inpainting gives more detailed maps than the simple diffuse inpainting, but falls behind with respect to the model-constrained details approach.

3.2.2 Recovered power spectra

In this subsection we compare the mean recovered spectra of $N = 100$ maps inpainted with each of the techniques. In Figures 3.5 - 3.8 we represent the different mean spectra with their standard deviation, as well as the theoretical spectra used by `CAMB` to generate the observed maps. We also show the power spectra obtained from the masked map and corrected by f_{sky}^{-1} for comparison.

Figure 3.5 corresponds to the TT mean spectra. The performance at lower ℓ of the three methodologies is similar. This is due to the fact that they are all diffuse inpainting, applied to N_{side}=128 in the simple diffuse inpainting, and to N_{side}=8 in the wavelet diffuse inpainting. Going into smaller scales we can see how the diffuse inpainting fails to recover the power at larger multipoles. The observation-constrained wavelet inpainting gives a slightly better reconstruction, but it presents sudden power drops at the multipoles at which the `HEALPix` wavelet decomposition is done. The best result is obtained using the model-constrained wavelet inpainting. This approach achieves a good performance at multipoles up to ℓ=150, while for larger multipoles it exhibits a small power excess.

The same discussion made in the TT spectra can be done for Figure 3.6, which shows the mean EE

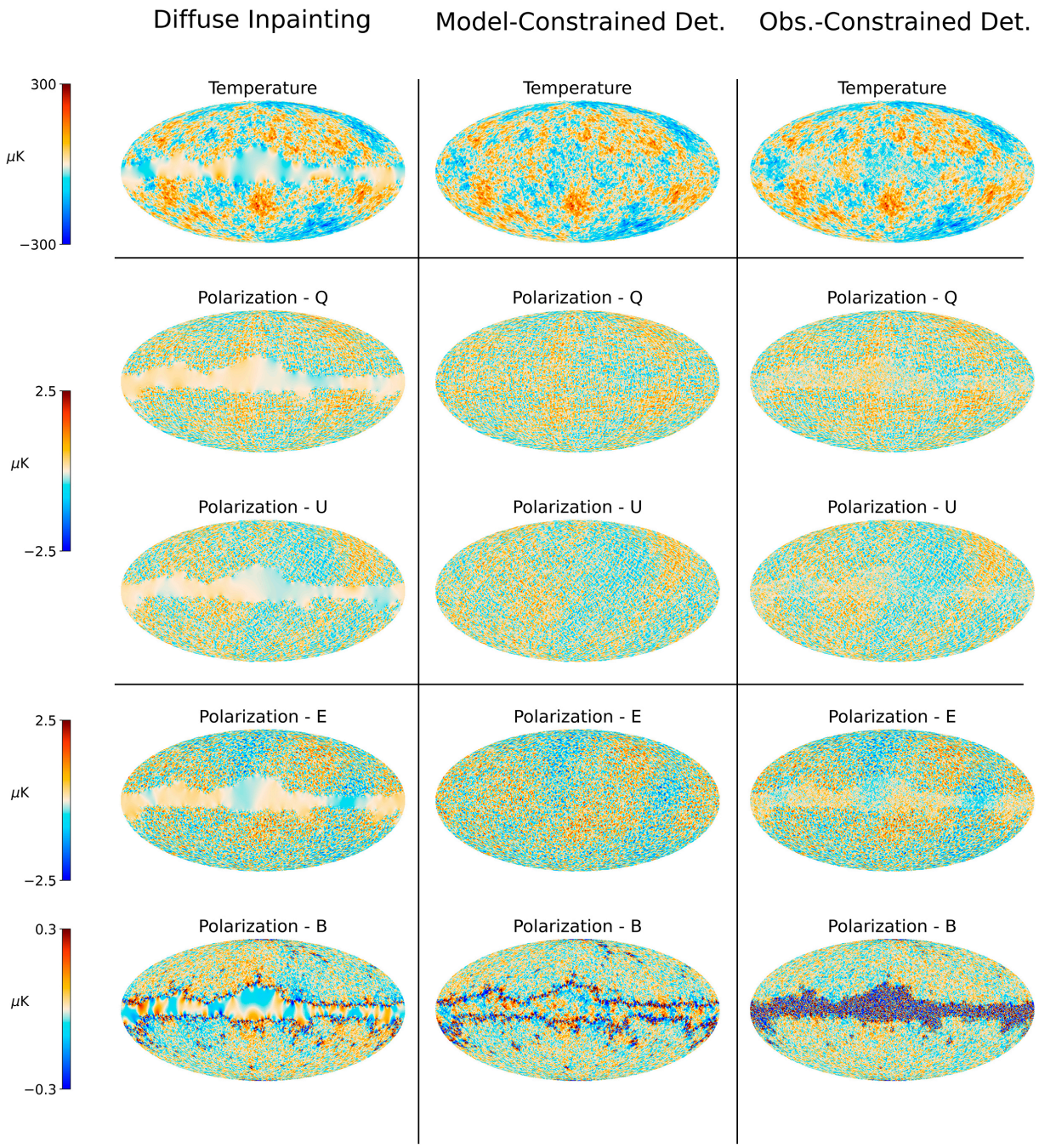

Figure 3.4: Inpainted T, Q, U, E, and B CMB maps using diffuse inpainting (left column) and wavelet diffuse inpainting with model-constrained details (center column) and observation-constrained details (right column). Resolution is $N_{side} = 128$ and the masked region corresponds to the common temperature mask shown in Figure 1.7. Inpainting has been done using the parameters obtained in Section 3.1.

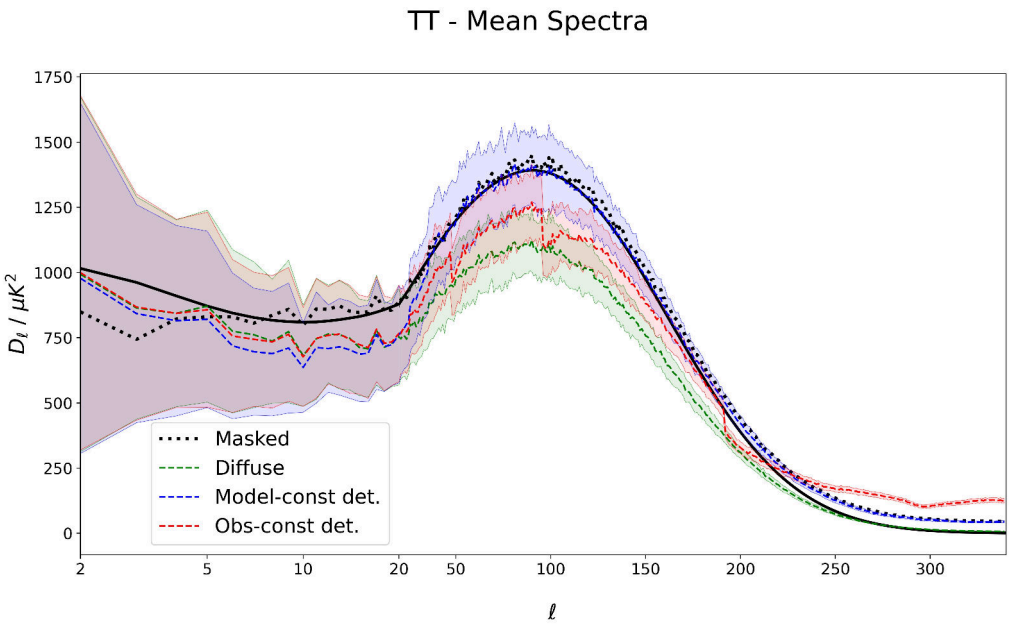

Figure 3.5: Mean recovered TT power spectra and standard deviation of $N = 100$ CMB maps inpainted using the three different methods described in this work. The black solid line represents the spectrum obtained with `CAMB` to simulate the observed CMB maps. Black dotted line represents the masked map poweer spectrum corrected by f_{sky}^{-1}. Scale in the x-axis is logarithmic up to $\ell = 20$, then linear.

power spectra. In this case, due to the power range that the EE spectrum encompasses, the plot only displays a clear view of the mutipoles $50 < \ell < 200$. The results are the same as in temperature: the diffuse inpainting fails to recover the power in this range, the observation-constrained wavelet inpainting presents sudden drops in power, and the model-constrained wavelet inpainting gives the best results.

Figure 3.7 shows the results of the BB angular power spectra. As expected from the inpainted maps, the recovered spectra are far from the theoretical spectrum. Results from the observation-constrained wavelet inpainting are remarkably bad, not being able to recover even the shape of the original spectrum. Results for the TE cross spectrum in Figure 3.8 display the same patterns as in the T and E spectra.

We can also make use of the $\tilde{\chi}$ estimator to compare the different techniques. In Figure 3.9 we plot the mean value of the ratio χ/χ_m as a function of ℓ_{max}, where χ is the estimator value of the inpainted maps and χ_m the estimator for the f_{sky}^{-1} corrected masked power spectra. These plots show the performance of each method as a function of the maximum ℓ that we are looking at. This allows us to see the best performance region for each technique, as well as whether the recovered spectra represent an improvement over the corrected power spectra from the masked map ($\chi/\chi_m < 1$).

Using this plots we can see how the three techniques are able to recover the TT, EE and TE spectra better than the masked spectra at the larger scales. For higher multipoles, the only method that is able obtain better results than the corrected masked spectrum is the model-constrained wavelet inpainting. In TT, the standard deviation regions of the techniques are overlapped, so a conclusive statement cannot be made with the use of this specific estimator. For BB, the best results are obtained with the simple diffuse inpainting, but as we have seen in Figure 3.7, none of the

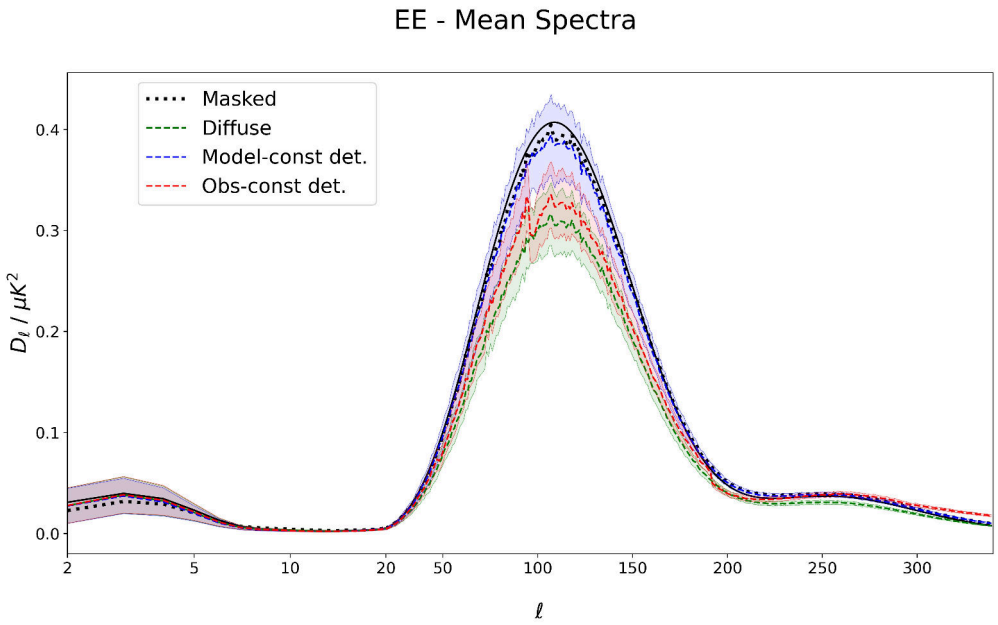

Figure 3.6: Mean recovered EE power spectra and standard deviation of $N = 100$ CMB maps inpainted using the three different methods described in this work. The black solid line represents the spectrum obtained with `CAMB` to simulate the observed CMB maps. Black dotted line represents the masked map poweer spectrum corrected by f_{sky}^{-1}. Scale in the x-axis is logarithmic up to $\ell = 20$, then linear.

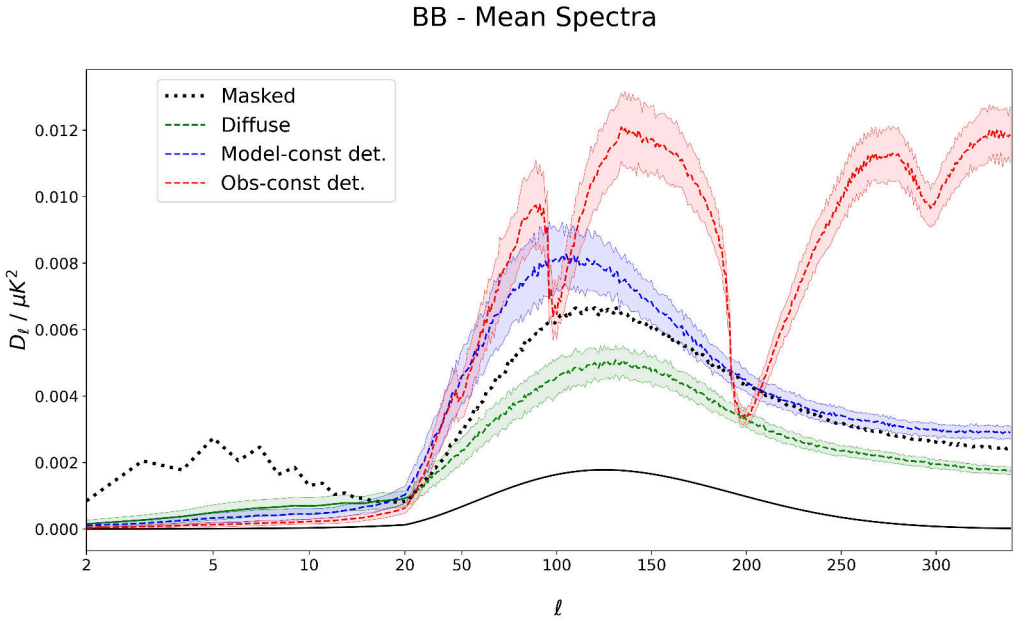

Figure 3.7: Mean recovered BB power spectra and standard deviation of $N = 100$ CMB maps inpainted using the three different methods described in this work. The black solid line represents the spectrum obtained with `CAMB` to simulate the observed CMB maps. Black dotted line represents the masked map poweer spectrum corrected by f_{sky}^{-1}. Scale in the x-axis is logarithmic up to $\ell = 20$, then linear.

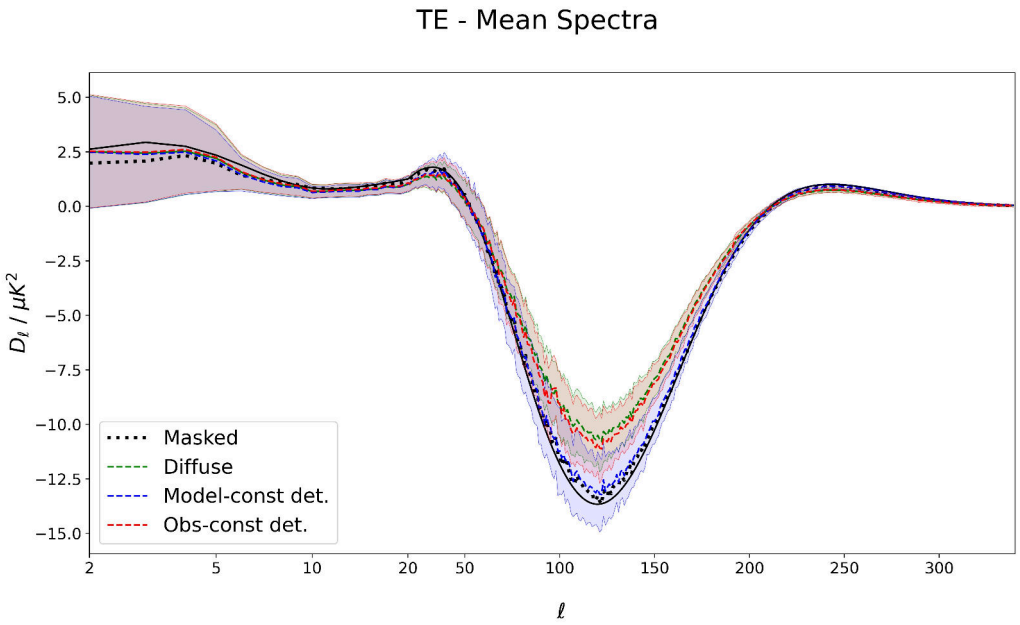

Figure 3.8: Mean recovered TE power spectra and standard deviation of $N = 100$ CMB maps inpainted using the three different methods described in this work. The black solid line represents the spectrum obtained with `CAMB` to simulate the observed CMB maps. Black dotted line represents the masked map poweer spectrum corrected by f_{sky}^{-1}. Scale in the x-axis is logarithmic up to $\ell = 20$, then linear.

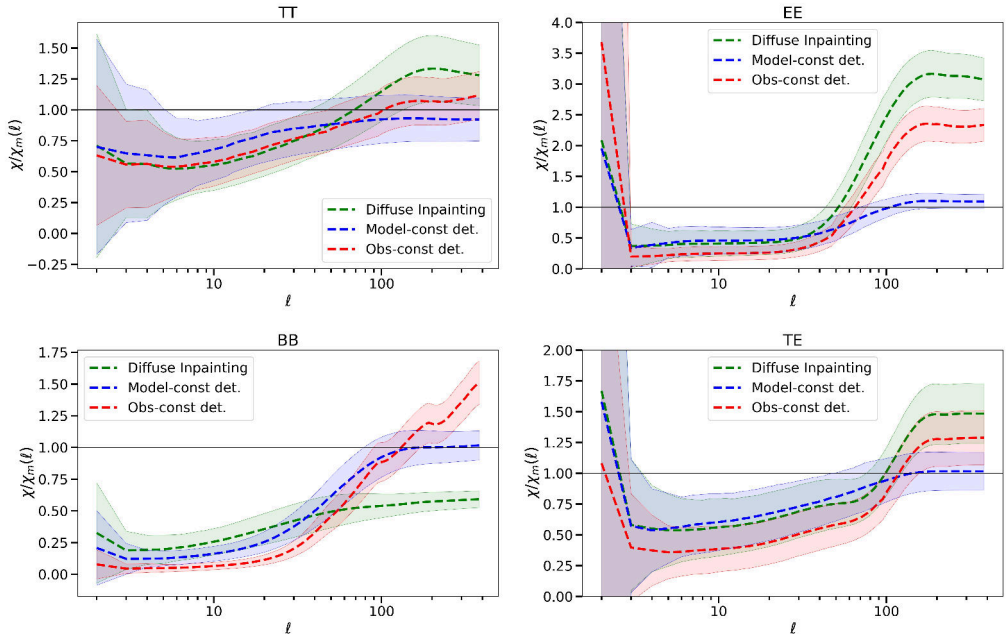

Figure 3.9: Mean χ/χ_m values and standard deviation as a function of ℓ of the spectra of each inpainting method for $N = 100$ maps. In order to calculate the $\tilde{\chi}$ estimator, we take as reference the spectra of the unmasked maps of each realization.

recovered spectra can be trusted. Figure 3.9 is useful to compare how each method performs against each other, and against the f_{sky}^{-1} corrected masked power spectra. However, it does not give us an absolute measure of the goodness of each technique, and it needs to be interpreted alongside the recovered power spectra displayed in Figures 3.5 - 3.8.

3.2.3 Multipole correlations

Lastly, we check for correlations between multipoles in the different methods. Figure 3.10 displays the correlation coefficients between multipoles of N=100 temperature maps using the different methodologies, as well as the coefficients for the masked map.

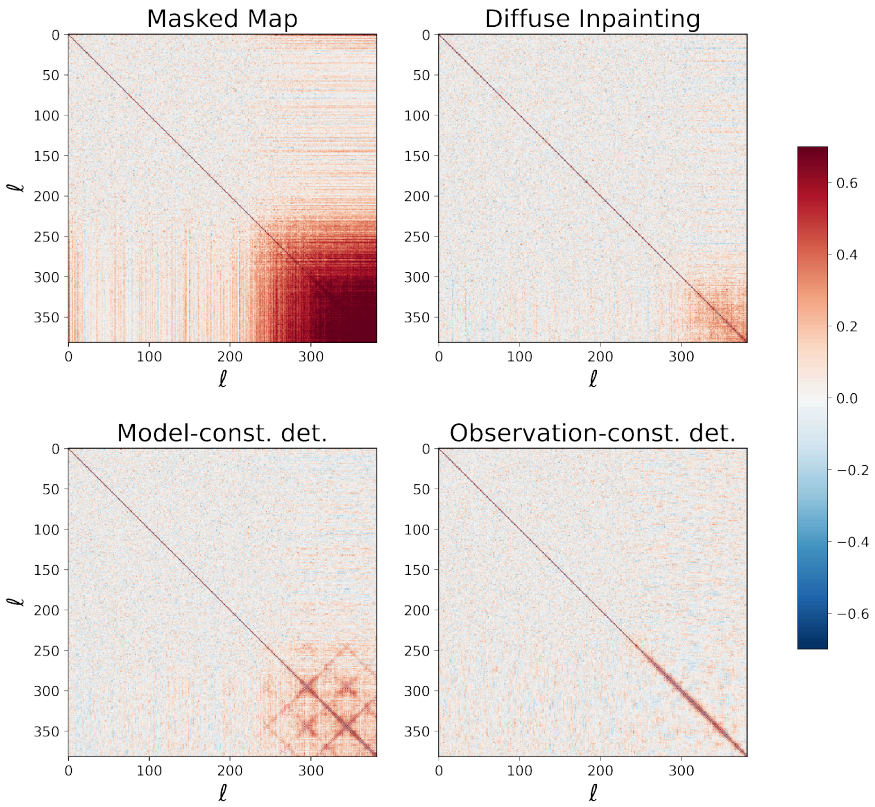

Figure 3.10: Correlation coefficients between multipoles for the masked temperature maps, as well as for the inpainted temperature maps using different techniques. A total of $N = 100$ different maps have been studied.

The masked map presents large correlations at higher multipoles ($\ell > 200$), as well as a linear pattern showing correlations between small-medium scales and large scales. The simple diffuse inpainting shows a similar picture but is able to mitigate these correlations. In the wavelet diffuse inpainting we find two different cases. The model-constrained details approach also reduces the correlations, but still exhibits a geometric pattern at $\ell > 200$ due to the anisotropic nature of the `HEALPix` wavelet decomposition. On the other hand, the observation-constrained details approach does not present such pattern, and only shows correlations between multipoles close to each other. This is so because

the observation-constrained approach fills in the details by taking care of decorrelating the different scales before doing so, as explained in the first step of the method in Section 2.3.2. The presence of strong correlations at $\ell > 200$ in the masked map and model-constrained wavelet inpaintings could be responsible for the power excess shown at large multipoles, especially visible in the TT spectra (Figure 3.5).

In order to fully assess the significance of multipole correlations we need to take into account the results from previous sections. Diffuse inpainting and the observation-constrained approach of wavelet inpainting are able to reduce significantly the correlations present in the masked spectrum. However, as seen in Figure 3.5, they fail to reconstruct the spectrum at most scales. The f_{sky}^{-1} corrected masked spectrum (dotted line in Figure 3.5) is closer to the original one, but it suffers from these strong correlations, especially at higher multipoles. It is in the model-constrained wavelet inpainting that we find a trade-off between recovering the angular power spectrum and mitigating multipole correlations. The presence of this anisotropic pattern in this approach's correlations opens the door to the use of an isotropic wavelet, such as the needlets used in [32], as a possible way to improve the technique.

3.3 Application to *Planck* data

In this section, we apply the diffuse inpainting and the two approaches of the wavelet diffuse inpainting to *Planck* data. We have chosen to apply them to the temperature component separation map obtained by SMICA.

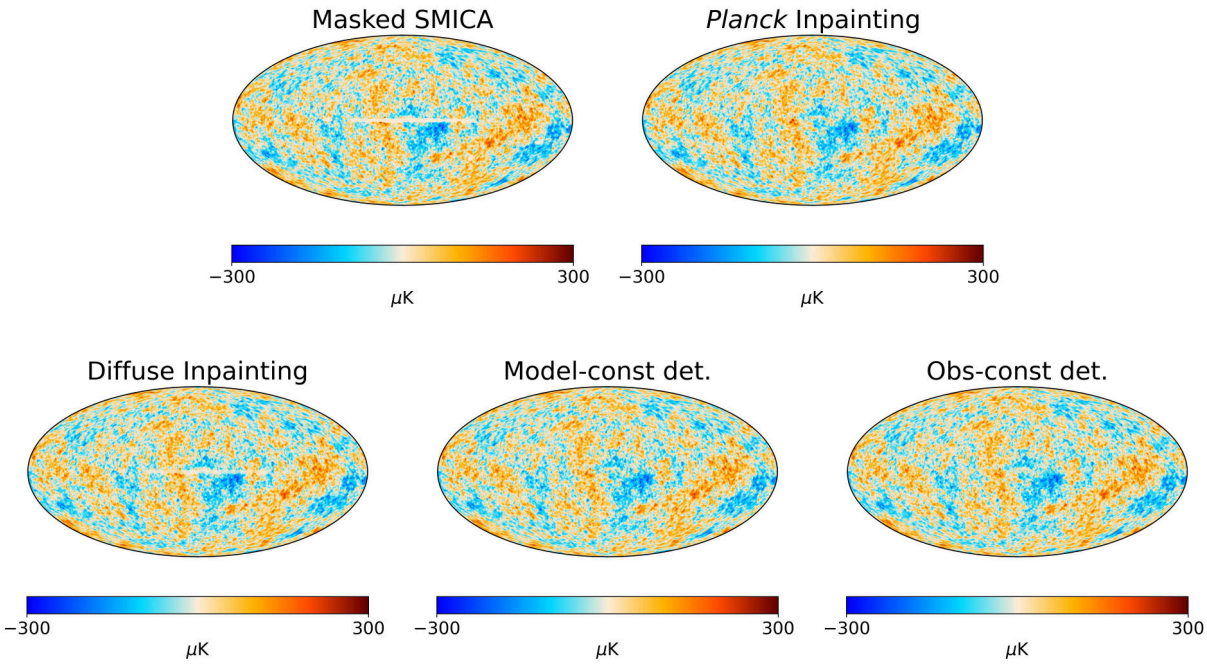

Figure 3.11: Top left: masked temperature map obtained with SMICA downgraded to N_{side} = 128. **Top right**: Inpainted SMICA temperature map obtained by *Planck*. **Bottom left**: map inpainted using diffuse inpainting. **Bottom middle**: map inpainted using wavelet inpainting with simulated detail maps. **Bottom right**: map inpainted using wavelet inpainting with gaussian detail maps

The mask used by SMICA is much smaller than the one we have used in our analysis. In this case, SMICA's mask is characterized by a f_{sky} value of $f_{sky} = 0.98$, while the mask used before corresponded to $f_{sky} = 0.79$. All of the study carried out in the previous sections was done and optimized for the common temperature mask. As such, the results obtained in this section should be taken with a grain of salt. It is likely that different inpaiting parameters from the ones used here could provide better results when applied to smaller masks, e.g. by decomposing up to a higher N_{side} resolution in the wavelet diffuse inpainting.

Figure 3.11 shows the masked SMICA map, the inpainted map obtained by *Planck*, and the inpainted maps using the methods described in Chapter 2. Due to the size of the mask, the wavelet inpainted maps cannot be visually distinguished from the one obtained by *Planck*. The map reconstructed with the diffuse inpainting still exhibits its characteristic blurry region inside the mask.

In order to compare the different methods, Figure 3.12 shows the ratio χ/χ_m for each method, taking as a reference the map inpainted by *Planck*. This reference is not the actual angular power spectrum but the best that *Planck* has been able to obtain. In this case, the diffuse inpainting does not produce a better spectrum than the f_{sky}^{-1} corrected masked SMICA map. It performs better at larger multipoles, but since this method does not inpaint small scales we can assume this effect to be produced by the mask. Both wavelet inpainting approaches stand below the $\chi/\chi_m = 1$, thus improving the spectra obtained from corrected masked spectrum. Best results are obtained by the observation-constrained wavelet inpainting at all ranges, with the other approach performing slighlty worse. As stated above, these are just preliminary results and an analogue study to the one performed in Section 3.1 should be done for this mask in particular.

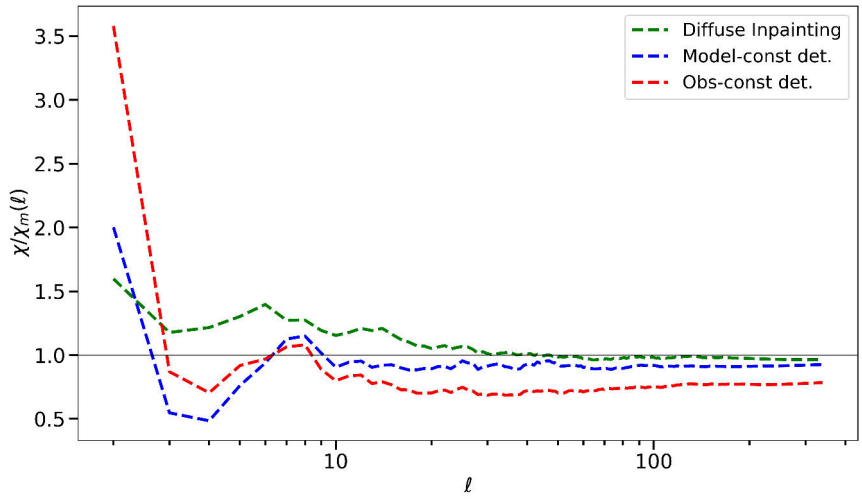

Figure 3.12: χ/χ_m values as a function of ℓ of the SMICA maps inpainted using the different inpainting techniques, taking as a reference the inpainted map obtained by the *Planck* Collaboration.

Chapter 4

Conclusions and future work

In this project we propose a novel inpainting technique for masked CMB sky maps, the wavelet diffuse inpainting. This method combines the large scale reconstruction of the diffuse inpainting with the multiresolution framework provided by the `HEALPix` wavelet scheme. This allows us to recover information at medium scales not covered by the simple diffuse inpainting. It also reduces the computing time by performing a diffusive inpainting at a much lower resolution. Maps at higher resolution are then obtained by means of the detail maps obtained in the `HEALPix` wavelet decomposition. We provide two different approaches to fill the masked details: using model-constrained detail maps and using observational-constrained detail maps.

In order to assess the performance of this new inpainting method, we have made a comparison between the simple diffuse inpainting and this technique. We have evaluated the different inpaintings at three levels: visual map reconstruction, recovered angular power spectra, and presence of correlations between multipoles. From the proposed approaches to the wavelet inpainting, the one using model-constrained detail maps is able to find a trade-off between visual quality, angular power spectra and reduced multipole correlations.

Maps inpainted using this new technique are visually indistinguishable from the original unmasked map. At a power spectra level, it is able to recover the TT, EE and TE spectra up to a much higher ℓ than the diffuse inpainting. However, multipole correlations are still present in this method due to the anisotropic nature of the `HEALPix` wavelet deconstruction employed. These correlations occur mainly between higher multipoles, and might be responsible for the power excess obtained by this method at the smallest scales. There is also some room for improvement in terms of power spectra recovery, especially at the largest scales.

This study opens the door for future works to expand on this technique and its capabilities. Combining the wavelet scheme with other inpainting techniques, such as constrained realizations [46], could be of special interest. Since the inpainting is performed at a low resolution, using more optimal techniques could improve the results at larger scales while keeping a reasonable computing time. In addition, substituting the `HEALPix` wavelet scheme for an isotropic wavelet could be potentially effective in terms of reducing correlations between multipoles, due to the anisotropic nature of the former. It is also interesting to improve on this work by studying a wider range of masks. In this project we have limited ourselves to the common mask provided by *Planck*. However, the preliminary results shown in Section 3.3 suggest that different maks affect the performance of this technique. Exploring and characterizing the inpainting for different masks would help to better identify the strengths of this method, and whether it can be applied to real data.

In conclusion, one of the proposed approaches clearly improves the results obtained with the diffuse inpainting. However, the recovered power spectra are still sub-optimal and further research needs to be done in order to confirm if this method performs good enough to be applied to real data.

Bibliography

[1] R.H. Dicke, P.J.E. Peebles, P.G. Roll and D.T. Wilkinson, *Cosmic Black-Body Radiation.*, *Astrophys. J.* **142** (1965) 414.

[2] A.A. Penzias and R.W. Wilson, *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*, *Astrophys. J.* **142** (1965) 419.

[3] D.J. Fixsen, *The Temperature of the Cosmic Microwave Background*, *Astrophys. J.* **707** (2009) 916 [0911.1955].

[4] E.R. Harrison, *Fluctuations at the Threshold of Classical Cosmology*, *Phys. Rev. D* **1** (1970) 2726.

[5] P.J.E. Peebles and J.T. Yu, *Primeval Adiabatic Perturbation in an Expanding Universe*, *Astrophys. J.* **162** (1970) 815.

[6] S. Seager, D.D. Sasselov and D. Scott, *How Exactly Did the Universe Become Neutral?*, *Astrophys. J., Suppl. Ser.* **128** (2000) 407 [astro-ph/9912182].

[7] A. Lewis and A. Challinor, *Weak gravitational lensing of the CMB*, *Phys. Rep.* **429** (2006) 1 [astro-ph/0601594].

[8] R.A. Sunyaev and Y.B. Zeldovich, *The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies*, *Comments on Astrophysics and Space Physics* **4** (1972) 173.

[9] G.F. Smoot, C.L. Bennett, A. Kogut, E.L. Wright, J. Aymon, N.W. Boggess et al., *Structure in the COBE Differential Microwave Radiometer First-Year Maps*, *Astrophys. J. Lett.* **396** (1992) L1.

[10] D.N. Spergel, L. Verde, H.V. Peiris, E. Komatsu, M.R. Nolta, C.L. Bennett et al., *First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters*, *Astrophys. J., Suppl. Ser.* **148** (2003) 175 [astro-ph/0302209].

[11] Planck Collaboration, P.A.R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont et al., *Planck early results. I. The Planck mission*, *Astron. Astrophys.* **536** (2011) A1 [1101.2022].

[12] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [1807.06209].

[13] M. Hazumi, P.A.R. Ade, A. Adler, E. Allys, K. Arnold, D. Auguste et al., *LiteBIRD satellite: JAXA's new strategic L-class mission for all-sky surveys of cosmic microwave background polarization*, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 11443 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 114432F, Dec., 2020, DOI [2101.12449].

[14] K. Abazajian, G.E. Addison, P. Adshead, Z. Ahmed, D. Akerib, A. Ali et al., *CMB-S4: Forecasting Constraints on Primordial Gravitational Waves*, *Astrophys. J.* **926** (2022) 54 [2008.12619].

[15] P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali, D. Alonso et al., *The Simons Observatory: science goals and forecasts*, *J. Cosmol. Astropart. Phys.* **2019** (2019) 056 [1808.07445].

[16] P.A.R. Ade, M. Amiri, S.J. Benton, A.S. Bergman, R. Bihary, J.J. Bock et al., *A Constraint on Primordial B-modes from the First Flight of the SPIDER Balloon-borne Telescope*, *Astrophys. J.* **927** (2022) 174.

[17] M. Zaldarriaga and U. Seljak, *All-sky analysis of polarization in the microwave background*, *Phys. Rev. D* **55** (1997) 1830 [astro-ph/9609170].

[18] D. Baumann, M.G. Jackson, P. Adshead, A. Amblard, A. Ashoorioon, N. Bartolo et al., *Probing Inflation with CMB Polarization*, in *CMB Polarization Workshop: Theory and Foregrounds: CMBPol Mission Concept Study*, S. Dodelson, D. Baumann, A. Cooray, J. Dunkley, A. Fraisse, M.G. Jackson et al., eds., vol. 1141 of *American Institute of Physics Conference Series*, pp. 10–120, June, 2009, DOI [0811.3919].

[19] A.A. Starobinsky, *A new type of isotropic cosmological models without singularity*, *Physics Letters B* **91** (1980) 99.

[20] M. Kamionkowski and E.D. Kovetz, *The Quest for B Modes from Inflationary Gravitational Waves*, *Annu. Rev. Astron. Astrophys.* **54** (2016) 227 [1510.06042].

[21] M. Tristram, A.J. Banday, K.M. Górski, R. Keskitalo, C.R. Lawrence, K.J. Andersen et al., *Improved limits on the tensor-to-scalar ratio using BICEP and P l a n c k data*, *Phys. Rev. D* **105** (2022) 083524 [2112.07961].

[22] A. Lewis, A. Challinor and A. Lasenby, *Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models*, *Astrophys. J.* **538** (2000) 473 [astro-ph/9911177].

[23] C.J. Copi, D. Huterer, D.J. Schwarz and G.D. Starkman, *No large-angle correlations on the non-Galactic microwave sky*, *Mon. Not. R. Astron. Soc.* **399** (2009) 295 [0808.3767].

[24] Y. Akrami, Y. Fantaye, A. Shafieloo, H.K. Eriksen, F.K. Hansen, A.J. Banday et al., *Power Asymmetry in WMAP and Planck Temperature Sky Maps as Measured by a Local Variance Estimator*, *Astrophys. J. Lett.* **784** (2014) L42 [1402.0870].

[25] R. Shi, T.A. Marriage, J.W. Appel, C.L. Bennett, D.T. Chuss, J. Cleary et al., *Testing CMB Anomalies in E-mode Polarization with Current and Future Data*, *arXiv e-prints* (2022) arXiv:2206.05920 [2206.05920].

[26] Planck Collaboration, P.A.R. Ade, N. Aghanim, D. Alina, M.I.R. Alves, C. Armitage-Caplan et al., *Planck intermediate results. XIX. An overview of the polarized thermal emission from Galactic dust*, *Astron. Astrophys.* **576** (2015) A104 [1405.0871].

[27] R. Génova-Santos, J.A. Rubiño-Martín, A. Peláez-Santos, F. Poidevin, R. Rebolo, R. Vignaga et al., *QUIJOTE scientific results - II. Polarisation measurements of the microwave emission in the Galactic molecular complexes W43 and W47 and supernova remnant W44*, *Mon. Not. R. Astron. Soc.* **464** (2017) 4107 [1605.04741].

[28] Planck Collaboration, R. Adam, P.A.R. Ade, N. Aghanim, M.I.R. Alves, M. Arnaud et al., *Planck 2015 results. X. Diffuse component separation: Foreground maps*, *Astron. Astrophys.* **594** (2016) A10 [1502.01588].

[29] Planck Collaboration, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., *Planck 2018 results. IV. Diffuse component separation*, *Astron. Astrophys.* **641** (2020) A4 [1807.06208].

[30] H.K. Eriksen, C. Dickinson, C.R. Lawrence, C. Baccigalupi, A.J. Banday, K.M. Górski et al., *Cosmic Microwave Background Component Separation by Parameter Estimation*, *Astrophys. J.* **641** (2006) 665 [astro-ph/0508268].

[31] E. Martínez-González, J.M. Diego, P. Vielva and J. Silk, *Cosmic microwave background power spectrum estimation and map reconstruction with the expectation-maximization algorithm*, *Mon. Not. R. Astron. Soc.* **345** (2003) 1101 [astro-ph/0302094].

[32] J. Delabrouille, J.F. Cardoso, M. Le Jeune, M. Betoule, G. Fay and F. Guilloux, *A full sky, low foreground, high resolution CMB map from WMAP*, *Astron. Astrophys.* **493** (2009) 835 [0807.0773].

[33] J. Delabrouille, J.F. Cardoso and G. Patanchon, *Multidetector multicomponent spectral matching and applications for cosmic microwave background data analysis*, *Mon. Not. R. Astron. Soc.* **346** (2003) 1089 [astro-ph/0211504].

[34] D. Alonso, J. Sanchez, A. Slosar and LSST Dark Energy Science Collaboration, *A unified pseudo-C_ℓ framework*, *Mon. Not. R. Astron. Soc.* **484** (2019) 4127 [1809.09603].

[35] J.D. Bilbao-Ahedo, R.B. Barreiro, P. Vielva, E. Martínez-González and D. Herranz, *ECLIPSE: a fast Quadratic Maximum Likelihood estimator for CMB intensity and polarization power spectra*, *J. Cosmol. Astropart. Phys.* **2021** (2021) 034 [2104.08528].

[36] Planck Collaboration, P.A.R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont et al., *Planck intermediate results. XXIV. Constraints on variations in fundamental constants*, *Astron. Astrophys.* **580** (2015) A22 [1406.7482].

[37] Planck Collaboration, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., *Planck 2018 results. VII. Isotropy and statistics of the CMB*, *Astron. Astrophys.* **641** (2020) A7 [1906.02552].

[38] A.N. Krylov, *On the numerical solution of the equation by which in technical questions frequencies of small oscillations of material systems are determined*, *Otdel. Mat. i Estest. Nauk* **7** (1931) 491.

[39] A. Marcos-Caballero and E. Martínez-González, *Scale-dependent dipolar modulation and the quadrupole-octopole alignment in the CMB temperature*, *J. Cosmol. Astropart. Phys.* **2019** (2019) 053 [1909.06093].

[40] G. Puglisi and X. Bai, *Inpainting Galactic Foreground Intensity and Polarization Maps Using Convolutional Neural Networks*, *Astrophys. J.* **905** (2020) 143 [2003.13691].

[41] K.M. Górski, E. Hivon, A.J. Banday, B.D. Wandelt, F.K. Hansen, M. Reinecke et al., *HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*, *Astrophys. J.* **622** (2005) 759 [astro-ph/0409513].

[42] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon et al., *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python*, *The Journal of Open Source Software* **4** (2019) 1298.

[43] B. Casaponsa, R.B. Barreiro, A. Curto, E. Martínez-González and P. Vielva, *Wilkinson Microwave Anisotropy Probe 7-yr constraints on f_{NL} with a fast wavelet estimator*, *Mon. Not. R. Astron. Soc.* **411** (2011) 2019 [1009.0632].

[44] C. Clapham and J. Nicholson, *The Concise Oxford Dictionary of Mathematics*, Oxford University Press (2009), 10.1093/acref/9780199235940.001.0001.

[45] M. Bucher, B. Racine and B. van Tent, *The binned bispectrum estimator: template-based and non-parametric CMB non-Gaussianity searches*, *J. Cosmol. Astropart. Phys.* **2016** (2016) 055 [1509.08107].

[46] A. Benoit-Lévy, T. Déchelette, K. Benabed, J.F. Cardoso, D. Hanson and S. Prunet, *Full-sky CMB lensing reconstruction in presence of sky-cuts*, *Astron. Astrophys.* **555** (2013) A37 [1301.4145].

Appendix A

Wavelet decomposition products

In this appendix we show an example of the different products of the `HEALPix` wavelet decomposition performed in Section 2.3.1 and Section 2.3.2.

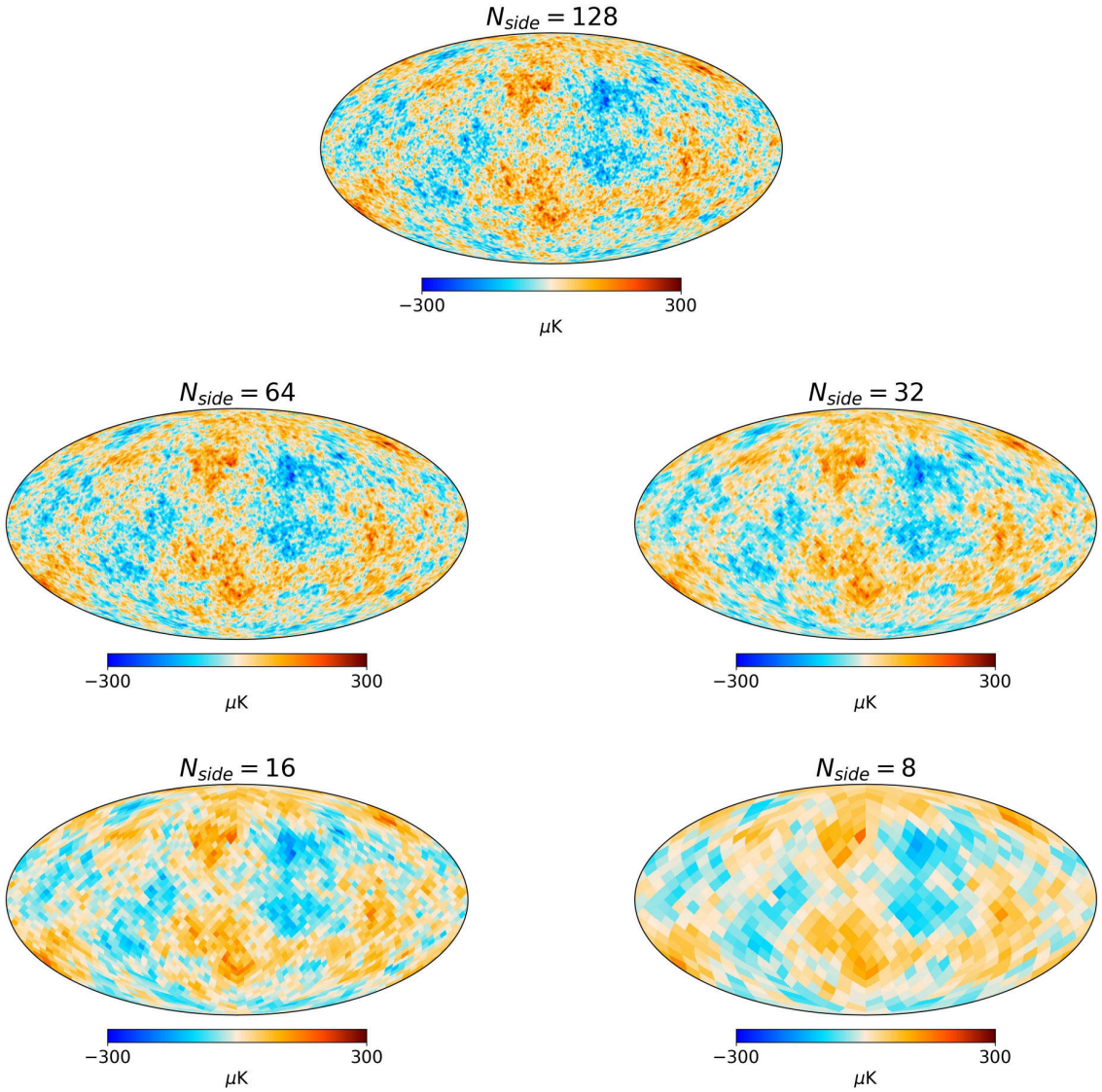

Figure A.1: Approximation maps of a CMB temperature map obtained by performing a wavelet decomposition from $N_{side} = 128$ up to $N_{side} = 8$. The original map has been obtained from the angular power spectra obtained with `CAMB` using ΛCDM and the parameters in 1.

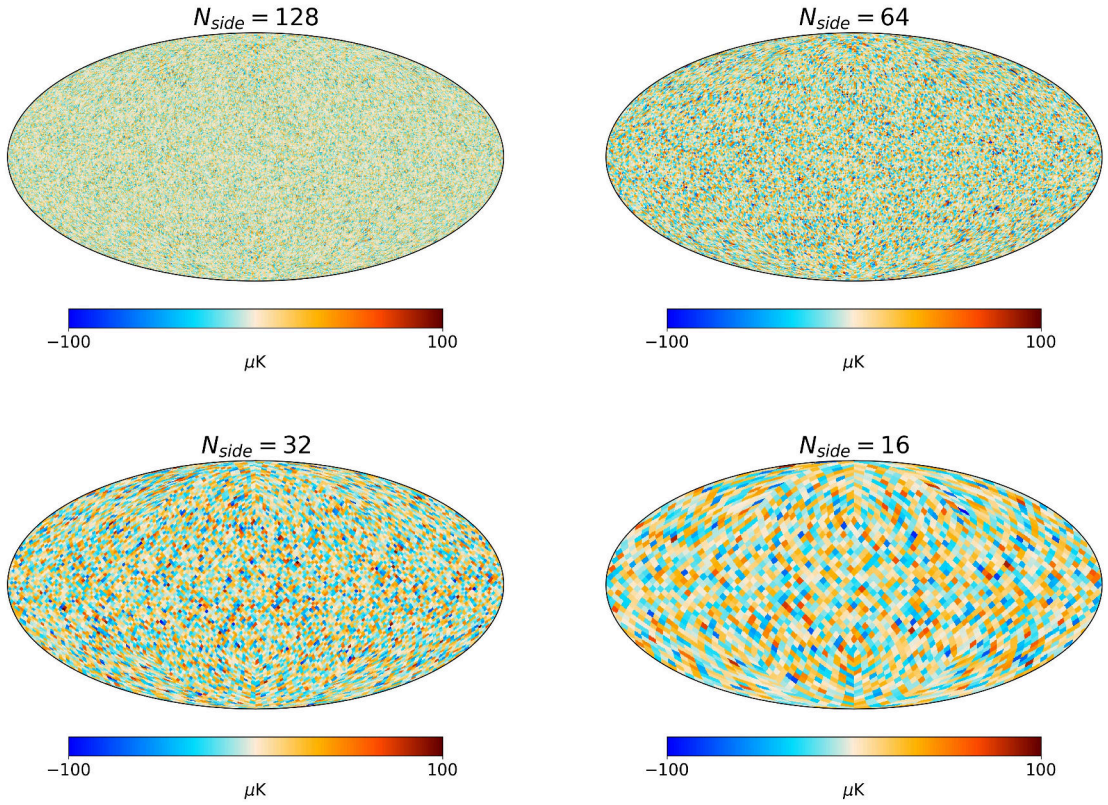

Figure A.2: Detail maps of a CMB temperature map obtained by performing a wavelet decomposition from $N_{side} = 128$ up to $N_{side} = 8$. The original CMB map is the one displayed on the top panel of Figure A.1.

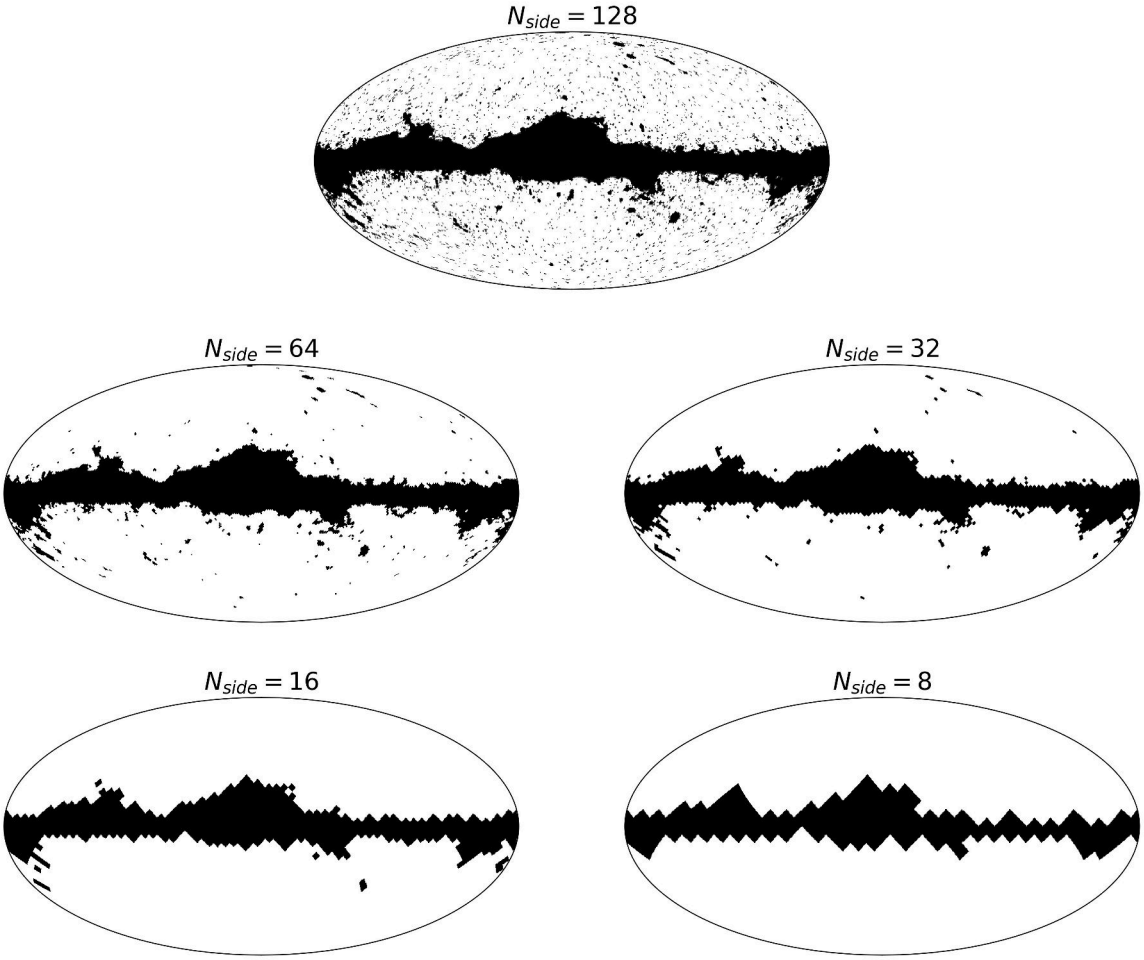

Figure A.3: Approximation maps of a mask obtained by performing a wavelet decomposition from $N_{side} = 128$ up to $N_{side} = 8$. The original mask is the common temperature mask used in *Planck*, obtained from 2.