Improving the Planck sky maps through Bayesian component separation

Trygve Leithe Svalheim

Thesis submitted for the degree of
Master of Science in Astronomy

Institute of Theoretical Astrophysics
University of Oslo

June, 2018

Abstract

The Planck satellite has provided a multitude of data images of the full microwave sky, or *sky maps*, since it saw first light in 2009. Its observations of the Cosmic Microwave Background (CMB) data has been paramount in the estimation of cosmological parameters vital in many branches of physics and astronomy. Although the processing of this information is coming to an end, the data still exhibit significant systematic effects coupled with foreground contamination, which impairs our ability to determine these parameters accurately.

In this thesis, we provide an overview of the Planck data analysis process with emphasis on Bayesian component separation methods for foreground removal. Furthermore, we seek to improve upon a new set of sky maps provided by Reijo Keskitalo at Lawrence Berkeley Laboratories. We do this by applying component separation in order to reveal systematic effects, which are subsequently corrected during map-making. This process is then repeated until systematic errors and foregrounds are suppressed to a satisfactory extent.

Lastly, we present the results of our analysis with the application of the latest generation sky maps. These maps were produced as a result of the efforts described in this thesis. We conclude that the new sky maps exhibit a significant reduction in instrumental errors, in comparison to the current state-of-the-art.

Acknowledgments

Writing a master thesis in astrophysics is a challenging task and one that requires an absurd amount of work. Although I wrote the words, my accomplishment is a product of all of the people who shaped me into the person I am today. First off, I would like to thank my family for making me a person. Second, I would like to thank my excellent supervisor Hans Kristian Kamfjord Eriksen for pushing me to my limits and forcing me to grow up, thereby making me a person with a master degree. Third, I would like to thank His Majesty the Blind Goat for making me an enlightened person, as well as the Great Crab in the South for making me a well-fed person. Last but not least, I would like to thank all my friends in Porsgrunn, Blindern Studenterhjem, and everywhere else for making me a happy person. You are all collaterally responsible for the outcome of this thesis.

Finally, I would like to explicitly thank the following people for their moral support:

Mathias Knudsen Sollie, Mathias Tangenes, Bjørn Ingeberg Fesche, Eskil Ravn Liane, Andreas Walmann, Even Pettersen, Liban Mahamed Omar, Vegard Andersen, Maja Regine Marrable, Eivind Ulstein Standal, Elias Kristian Zahl-Pettersen, Victoria Myhr Skjetne, Johan Niklas Ellegård Høiseth, Sebastian Gregorius Winther-Larsen, Kristina Totland Carm, Helene Brygmann, Sindre Skotaam, Erik Hernholm Haatveit, Oda Kibsgaard-Petersen, Simen Hunding Strømme, Siri Skotvedt, Gøran Svaland, Karoline Eriksen Pedersen, Hanna Totland Carm, Jonas Mjeltevik Halvorsen, Snorre Damm Tollefsen, Pål Kjellevold, Håvard Wilsher and Frida Amanda Fimland.

Your support is deeply appreciated.

Trygve Leithe Svalheim
Blindern, June 1st, 2018

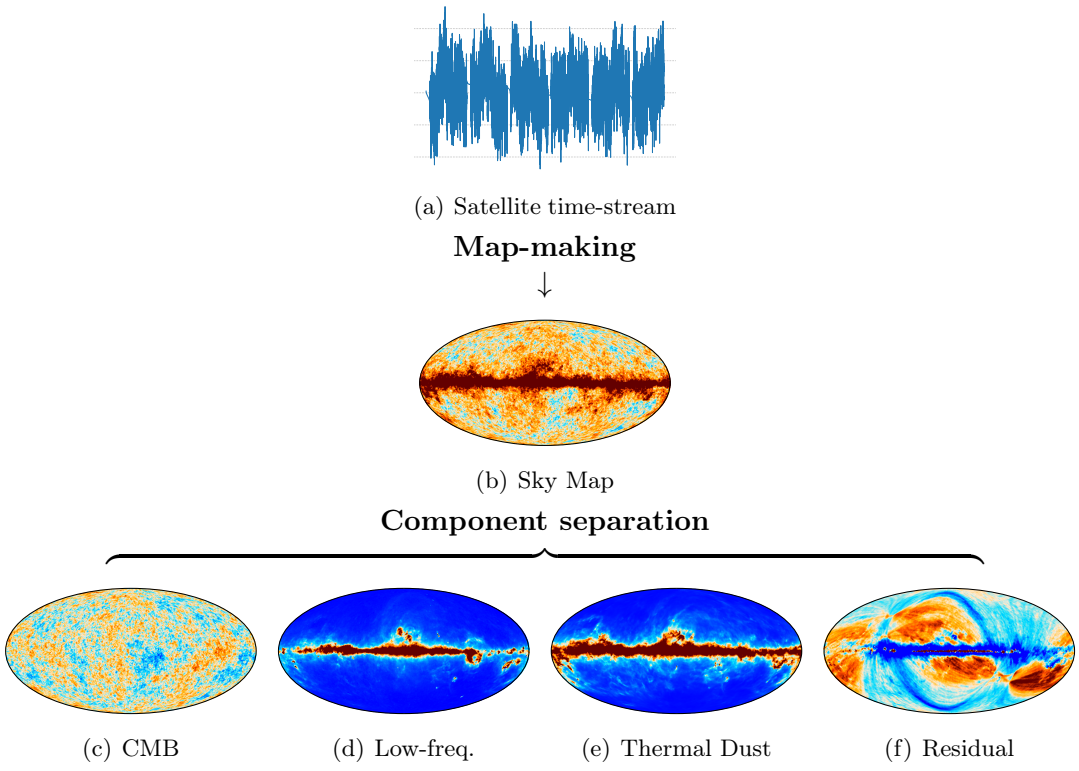
Preface

The topic of this thesis is the study of the *cosmic microwave background* (CMB) radiation. The CMB was the first electromagnetic radiation released following the Big Bang and can be seen today as the backdrop of the night sky. This signal is paramount to the study of the early universe as it serves as the oldest observational data point. Our inability to study the eras prior to this by means of direct observation underscores the importance of measuring the CMB to great precision. Unfortunately, our view of this signal is obstructed by various types of galactic contamination. In order to remove such foregrounds, we must accurately characterize them in a *sky model*. Therefore, we observe their behavior across frequencies in the microwave domain. Our most recent space-borne observation of the microwave sky was done by the satellite known as Planck. Planck's eyes see the sky at nine different frequencies, and delivers the data to us, which we assemble into nine full sky images we call *sky maps*.

The process of removing the foregrounds from these sky maps is called *component separation*, which as the name suggests, decomposes the full sky signal into their respective components. When we observe the behavior and intensity of a component across frequencies, we must do so on sky maps with little instrumental errors. Instrumental errors should in principle be accounted for in the process known as *map-making*. Map-making is the interpretation of instrument data into sky maps. Unfortunately, many instrumental errors are not revealed until after component separation. At the moment, the challenge in CMB research lies in characterizing the interplay between systematic errors and foreground contamination. The goal of this thesis is to deal with this problem and improve upon the sky maps by iterating between component separation and map-making. We attempt to construct a set of maps that leave no instrumental defects and adheres to our model of the sky. For a schematic overview of this analysis see figure 1. Many different methods of map-making have been introduced to the CMB community. In this thesis, we will compare the official Planck collaboration maps produced by the Planck Data Processing Center (DPC), to a new set of maps called NPIPE. NPIPE is however at an early stage of development, and the focus of this thesis will be on the improvement of these. The engine used to fit the model to the sky maps is a Bayesian statistical tool known as Gibbs sampling. In this thesis, we apply this Gibbs sampling through the `Commander` code [1]. `Commander` inputs sky maps, estimates signal parameters, and outputs separated component maps. However, for the code to converge to the correct values of the model parameters, and produce the correct component maps, we must feed it an appropriate model for the microwave sky. The work presented in

this thesis relies on the slow and steady adjustment of this model by ways of educated trial and error. This work is motivated by the fact that the scientific products provided by the Planck mission are essential in determining physical constants vital in many different branches of physics and astronomy.

In order to present our findings in an understandable manner, we provide an overview of the basics of cosmology in chapter 1. In order to gain an understanding of our place in the grand scheme of scientific efforts, the following chapter elaborates on the history of CMB research. The second part of this thesis discusses the implications of external effects on our signal. We categorize the various foreground emissions, and present some of the systematic effects introduced by the Planck instrument, before finally describing the tools used to deal with these. In the final part of the thesis, we present the results of our study. First we discuss the application of our tools on the current state of the art sky maps as provided by the Planck Collaboration, before moving on the main focus of the thesis; the improvement of the NPIPE sky maps. Finally, we present the results of our efforts, which exhibit great improvement in regards to systematic effects. We conclude our study with a discussion of possible improvements and the future of component separation.



(a) Satellite time-stream

(b) Sky Map

(c) CMB (d) Low-freq. (e) Thermal Dust (f) Residual

Figure 1: Schematic overview of map-making and component separation. In this thesis we aim to minimize systematic errors in the residual map by constructing a satisfactory sky-model for use in map-making. The overarching goal is the precise analysis of the CMB.

Contents

List of Figures

Part I

Introduction

Chapter 1

A Cosmology primer

"I would like to uncover the secrets of the universe."

— Megan Fox

Even though the modern human has existed for approximately 200,000 years, it was not until relatively recently that we started systematically solving the mysteries of the cosmos. Despite the universe carrying no obligation to reveal its secrets, the smartest minds of our young race have made their attempts at explaining it all, and considering we only invented the wheel about 7000 years ago, the results are staggering.

1.1 What is cosmology?

The branch of science devoted to understanding the evolution and dynamics of the universe is called *cosmology*. One of the foundations of the field is the assumption that we live in a universe described by the *cosmological principle*, which states that the universe appears the same in every direction. A second foundation is *general relativity*, which describes gravity. One of the more recent discoveries in cosmology is the Cosmic Microwave Background radiation (CMB) observed in 1964 [2]. The CMB is a signal which can be seen on the full sky and represented as a snapshot image of the universe 378,000 years after coming into existence. This picture serves as one of the most critical data points in unraveling the early history of the universe. In conjunction with General Relativity and the cosmological principle, the CMB provides a mean to demystify the elusive adolescence of the universe, and perhaps even its infancy.

We endeavor to understand the universe by constructing cosmological models. These models are mathematical descriptions of how the universe evolves through time and encapsulates its physical properties. We may model our observed universe in a myriad of ways, but the majority of cosmologists subscribe to the ΛCDM model as our best-educated guess and hence refer to it as *the standard model*. So far, this model seems to predict our observations of the large-scale properties of the universe in a satisfactory manner. In contrast to some previous models, the ΛCDM model describes an expanding universe dominated by Dark Energy (Λ) and Cold Dark Matter (CDM). These

two components alone make up 69 and 26 percent, respectively, of the contents of the universe today [3]. The significance of these percentages is great, as they determine how the universe has evolved through time. A more in-depth description of how the contents of the universe shape its evolution is presented in section 1.3.1. However, now, let us delve deeper into the fundamental concepts upon which the study of the universe rests.

If not otherwise stated, the information in this chapter is based upon the book "Modern Cosmology" by Scott Dodelson as well as "Space-time and Geometry" chapter 8 by Sean Carroll [4, 5].

1.2 A Cosmological Principle and Space Time

One of the cardinal concepts of cosmology is known as the Cosmological Principle, and it states that the universe is homogeneous and isotropic on large scales, the implication is that on large scales, it looks the same in all directions. Had this not been the case, we would be unable to infer statistical properties about the universe as a whole, which is imperative in cosmological research. Sir Isaac Newton was the first person to formulate its modern understanding through a series of mathematical proofs that predicted the motions of the planets (almost) perfectly according to a law of "universal gravitation". He postulated that because the planets obey the same laws of gravitation within our solar system, so must the stars we see on the night sky [6].

Over two hundred years later, Albert Einstein refined this theory by introducing what he called *space-time* [7]. As the name implies, it describes the fabric of the universe, by elegantly weaving together our three spatial dimensions with time. This theory creates a framework in which to place physical events, or in other words; the chessboard of the universe. This framework has birthed many fundamental theories in modern physics. Mathematically, it is represented by the *metric*, in the form of the tensor $g_{\mu\nu}$[1]. One of the most groundbreaking properties of space-time is the fact that it bends when matter is present. Much like rocks on a stretched bed sheet, shapes space, and thereby the way objects move while traversing it. Einsteins description replaces the Newtonian theory of gravity, which describes gravity as a force acting between two objects. This characterization of the fabric of space is known as Einstein's theory of general relativity (GR). Within this framework, the relationship between bed sheets and rocks, or matter and curvature, is mathematically described through the Einstein equations,

$$E_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{1.1}$$

On the left-hand side, we find the Einstein tensor, $E_{\mu\nu}$ which is expressed using the metric tensor and represents the geometry of space-time. On the right-hand side, $T_{\mu\nu}$ is the energy-momentum tensor, and describes the distribution of matter and energy, while G is the gravitational constant. The μ's and ν's is the Einstein summation notation, a

[1] A tensor is a mathematical object that generalizes vectors and matrices. A vector, for example, is a rank one tensor while a scalar is a rank 0 tensor. In this case, we introduce the tensor notation as a way to include the time dimension.

useful alternative to matrix notation when working with tensors. By applying the cosmological principle, we can derive the metric representative of our universe. This metric is an ideal solution to equation 1.1, and is known as the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric, which we may summarize through the "line element",

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

Just like Cartesian coordinates do with $dl^2 = dx^2 + dy^2$, the line element incorporates the shape of the universe and time and is often referred to as *proper time*. In equation 1.2, dt is the change in time, and dr, $d\theta$ and $d\phi$ are the three spatial spherical coordinates. Additionally, we introduce k, being the shape of space-time itself. The k parameter can be -1, 1 or 0, depending on whether the shape of the universe is open, closed or flat, respectively. For example, if our universe were projected onto a sphere, and we sent two particles in parallel out into space, we would observe them slowly converging. Conversely, in an open universe, they would diverge. If the universe is flat, however, we would observe these two particles continuing in parallel, which becomes evident when we insert $k = 0$ into the metric and recognize the equation for ordinary Euclidean space.

Another interesting property of this metric is the factor $a(t)$; the *scale factor*. The scale factor compensates for the fact that the size of the universe may change over time; to have a general way to measure distances in a fluctuating universe, we must have a fluctuating ruler. This is necessary because we require a definite measure of distances within the universe regardless of its expansion. This way of factoring out the expansion of the universe is known as measuring *comoving distance* and is an important tool in cosmology. The necessity for comoving distances came with the observation that distant stars are generally moving away from us in all directions, made by Edwin Hubble in 1929 [8], which led to the conclusion that the space between us is expanding. This major discovery is formulated through the Hubble law,

$$v = H_0 d. \tag{1.3}$$

Where we see that the velocity, v, of an object relative to us depends linearly on the distance, d, between us. H_0 is the Hubble constant or the present value of the *Hubble factor*. The Hubble factor is expressed by the scale factor, and changes in time through the relation,

$$H(t) \equiv \frac{da/dt}{a} = \frac{\dot{a}}{a}. \tag{1.4}$$

Exactly how the scale factor changes over time is determined by the Friedmann equations, but observations and careful calculations show that the universe has always expanded. The implication of this is that at the universe at one point must have been infinitely small.

1.3 The dynamics of the universe

When describing the evolution of the geometry of our universe, we look back to Einstein's relationship between geometry and energy. This time we take a step back, and

use the energy-momentum tensor $T_{\mu\nu}$ to express the contents of the universe as a perfect, isotropic fluid; a fluid with no preferred direction. This significantly reduces the number of equations given by the Einstein relation, and leaves us with the two *Friedmann equations*,

$$\text{1st Friedmann equation:} \qquad \dot{a}^2 + kc^2 = \frac{8\pi G}{3}\rho a^2, \tag{1.5}$$

$$\text{2nd Friedmann equation:} \qquad \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right). \tag{1.6}$$

Here, dots denote the derivative with respect to time, ρ is the energy density, and p is the pressure. Hence, these equations describe the evolution of the expansion of the universe as a function of its contents. While ρ and p in this equation describe the full energy contents of the universe, the reality is that they are a sum of many different contributing types of energy.

1.3.1 Contents of the universe

When we look around us, we see a world composed of ordinary matter; protons, neutrons, and electrons bound together to form atoms. However, it was not until very recently that scientists discovered that such "baryonic" matter does not, in fact, contribute the vast majority of the energy contents of the universe. This fact is imperative in our ability to model the universe, as their associated properties in large part determine the dynamics of the world in which we live. The four main components of the universe included in the modern ΛCDM-model are baryonic matter and radiation, as well as dark matter and dark energy.

Baryonic Matter is ordinary matter, it is what makes up all of us and is the stuff of which atoms are made.

Radiation is simply photons, waves or mass-less particles traversing space.

Non-baryonic Matter (Dark Matter) is "invisible" matter that only interacts with gravity. Dark matter has not yet been directly detected, but its effect on its surroundings are inevitable from astrophysical observations. A few different types of dark matter has been proposed, but in the standard model, we subscribe to the type called cold dark matter (CDM).

Dark Energy is a mysterious force driving the accelerated expansion of the universe.

Time, distance, and expansion

Of special interest to us is the time evolution of the expansion in regards to their relative density to each other. We often express these quantities in the form of fractional densities, where they all add up to one in a flat universe. In such a description Ω_{b0} is the fractional density of baryonic matter, Ω_{c0} is dark matter, Ω_{r0} is radiation, Ω_{k0}

is curvature and Ω_Λ is dark energy. With our components in order, we may employ the Friedmann equations to construct a simplified and intuitive relation to how these fractional densities shape the expansion through time, with

$$\frac{H^2}{H_0^2} = (\Omega_{b0} + \Omega_{c0})a^{-3} + \Omega_{r0}a^{-4} + \Omega_{k0}a^{-2} + \Omega_\Lambda. \tag{1.7}$$

However, when conducting observational astronomy, one effectively looks back in time. By virtue of this relationship, the scale factor a may also be attributed to cosmological distance. In astronomy, the distance of an object is often described by its redshift[2], z, which is related to the scale factor by $\mathrm{a}(t) = (1 + \mathrm{z})^{-1}$. Hence, we have a cosmological relationship between distance, expansion and time. Looking back at equation 1.7, we thus note that as we move towards the earliest of times $t = 0$, or, $a \ll 1$, the second term grows, describing a universe dominated by radiation. By measuring these fractional densities today, we may learn crucial information about the evolution of the early universe. In the next section, we will discuss the earliest physical processes responsible for shaping the world in which we live.

1.4 The infant universe

In this section, we will outline the early moments of the universe. When studying the grand scheme of things, as one does in cosmology, these are critical moments that shaped the world as we see it. We split these moments into four; the beginning, inflation, Big Bang nucleosynthesis, and recombination. We are unable to directly study the epochs prior to recombination by means of electromagnetic radiation, as they all took place in the period before the first light was allowed to travel freely. Our cosmological models in conjunction with observations such as the CMB, however, allows us to make educated guesses.

1.4.1 The beginning, $t < 10^{-43}$ s

Intuitively, we begin the story of the universe at $\mathrm{t} = 0$. It is hard to say exactly what that time looked like, but a popular conjecture is the idea of an initial singularity; a point in space and time with infinite density and temperature where the meaning of time becomes ill-defined. The following 10^{-43} seconds is known as the *Planck era*. At this point, the universe was so vastly different than our current understanding of physics no longer applied. It is theorized that during this period, all four fundamental forces, gravity, weak- and strong nuclear force and the electromagnetic force, were one, and only became separate in the period following, known as the *Grand Unification Epoch*.

[2]The redshift of an object an effect attributed to the relative velocity between us and it. This velocity induces a Doppler-shift of observed electromagnetic radiation. Considering the expansion of the universe objects a certain distance away from us, are red-shifted a certain amount according to Hubble's law.

1.4.2 Inflation, $t \approx 10^{-35}$ s

With the end of the Grand Unification Epoch, some unknown effect triggered an event known as *cosmic inflation*. Inflation was a period of exponential expansion around the time when the universe was 10^{-35} seconds old. As a result of this, quantum fluctuations in the cosmic soup were scaled up to irregularities of cosmic size, which would later allow for the formation of structure. Furthermore, a period of rapid expansion explains among other things why the universe appears isotropic and homogeneous, because it left enough time for matter to equilibrate, before being blasted so far apart each corner of the cosmos could no longer exchange properties, such as temperature. This period in the history of the universe is particularly interesting to cosmologists today, and a validation of the theory by indirect observation would almost certainly lead to a Nobel prize. As inflation ended, the theoretical explanation is that the "inflaton field" decayed into ordinary particles in a process called reheating. Moreover, as baryons and anti-baryons formed, they annihilated to create photons, leaving a universe dominated by radiation, with only a fraction of baryons surviving due to an initial asymmetry in baryon-anti-baryon numbers.

1.4.3 Big Bang Nucleosynthesis (BBN), $t \approx 1$ s

As the universe continued to expand, it cooled, and allowed for the formation of new elements. At this point, the temperature of the universe had been reduced to $T \approx 1$ MeV, and weak interactions had been "frozen out", meaning that their rate of interaction was not fast enough to fight the rate of expansion and therefore were unable to remain in thermal equilibrium. This significantly affected the particle abundances in the young universe, and as neutrinos reached the end of their short life-span, they too started to decay, cooling the universe further down to $T \approx 1$ eV. The reduced temperature of the primordial plasma sparked the formation of the very first atoms, through the process known as *Big Bang Nucleosynthesis* (BBN). This process is responsible for the creation of all neutral hydrogen atoms in the universe today. It could only occur because the temperature of the primordial plasma had dropped below that of the atoms binding energy, allowing particles to combine without being torn apart by their surroundings. The temperature continued to drop, allowing heavier elements to form. The abundances of these heavier elements today serve as essential measurements in constraining models of the early universe.

1.4.4 Recombination, $t \approx 378,000$ years

We jump ahead in the history of the universe to the epoch most important for this thesis, known as Recombination. When the universe was approximately 378,000 years old, the formation of new elements had caused a dramatic drop in the abundance of free electrons. Up until this point, the cosmic soup was so dense that photons were unable to travel freely without scattering off an electron and changing direction in an endless loop, essentially not getting anywhere. Another way of saying it is that the photons had a very short *mean free path*, which left the universe opaque. However, as the universe

cooled to below the 13.6 eV (3000 K) binding energy of neutral hydrogen, protons and electrons were consumed in the formation of the first atoms. This subsequently allowed photons to travel freely without being scattered, rendering the universe transparent. These photons can be seen as microwave radiation on the sky today, and constitute the Cosmic Microwave Background Radiation, or CMB. In order to fully appreciate the significance of this signal, we introduce the term *last scattering surface*. The last scattering surface of a photon is the surface of which it was last reflected, the connotation of which is that the energy information it carries, is directly proportional to the medium of whence it came. When we look at the sky, we often see clouds, these clouds did not produce the photons that hit our eyes, but is rather the last scattering surface of the photons emitted from the sun. The idea of the last scattering surface is essential in this thesis, as it allows us to see a perfect picture of the universe at the time of recombination, but not the time of their creation. While the photons did exist prior to recombination, they were unable to pass on the information they carried at that time. Therefore, in order to understand the physics of the universe prior to recombination, we must use the CMB as a starting point and extrapolate backward in time.

Chapter 2

The Cosmic Microwave history

"I feel like I'm too busy writing history to read it."

— Kanye West

In this chapter, we look at efforts that led to the discovery of the Cosmic Microwave Background and the subsequent hunt for its finer details. With the analysis of the actual CMB presented in the last chapters, it is useful to know what has been done before, as well as get an idea of our place in the big picture. We rely on the book "The cosmic microwave background" by Rhodri Evans [9].

2.1 The birth of modern cosmology - Model building

We trace the roots of modern cosmology back to Albert Einstein's groundbreaking theory presented in 1916. This theory, known as the theory of General Relativity (GR), laid out the framework of the universe and established space-time as a way of parameterizing space and time [7]. With his theory, he meticulously explained how matter and energy bend space through a set of elegant equations. With the presentation of these revolutionary new equations, the race was on to solve them for a solution that best describes the universe as we see it today. The solutions of these equations would serve as models for our universe, resting upon the general understanding of how the universe is stitched together as described by Einstein. Many such models have been presented throughout the years. Some models describe expanding universes, and others describe static universes, some describe universes with dark matter and others without. In the decade following Einstein's discovery, a set of non-static, expanding, homogeneous and isotropic solutions were published by the Russian cosmologist and mathematician Alexander Friedmann. Up until this point, the consensus in the scientific community was that the universe was neither expanding nor contracting. After Friedmann's premature death, however, the solution was expanded upon by a Belgian by the name of Georges Lemaitre who proposed that red-shifts of nebulae may be signs of an expanding universe [10]. Furthermore, using available data he calculated the rate of the expansion to be

625km/s/Mpc[1], which is now what we call the "Hubble constant" but with a vastly different value. However, Einstein was not convinced, and therefore neither was the rest of the scientific community at the time. It was not until he published a letter in the prestigious scientific journal Nature, where he described the beginning of the universe as a super dense and small "primordial atom" that he managed to persuade the general public and Einstein himself [11].

Regardless of the success of the primordial atom some scientists still favored other theories. One such theory is known as the "steady-state theory" first proposed in 1948, which argues that in order to extrapolate the laws of nature to the outer edges of the universe, the universe cannot be changing with time [12]. The steady-state theory became the main rival of the primordial atom, and the most prominent advocate of which was Sir Fred Hoyle, one of the most influential cosmologists of the twentieth century. With his paper titled "A New Model for the Expanding Universe" he mathematically expanded upon the steady state theory, adding that for the universe to maintain a constant matter density, it must continuously create matter as it expands [13]. He continued to battle the "primordial atom" theory to his death and even coined the term "the Big Bang theory" in a failed attempt to discredit Lemaitre's "primordial atom" theory.

2.2 The age of observational cosmology

As the Second World War reached its end, many problems with the Big Bang theory remained. At the time, the Hubble constant was calculated to be 500 km/s/MPc which describes a 2 billion-year-old universe, half the age of the earth. However, with the rapid improvement of radio technology during the war, the field of radio astronomy was emerging.

Around the same time, a Russian physicist, cosmologist, and biologist George Gamow, along with his Ph.D. student Ralph Alpher, theorized that hydrogen and helium were formed within the first 17 seconds after the big bang. Together they published an article laying out the foundation for the field of nucleosynthesis [14] [2]. Continuing the research on the early universe, Gamow teamed up with fellow John Hopkins researcher Robert Herman. Together they published several papers on nucleosynthesis, where they made the first prediction of relic radiation from the early universe. From their calculations, they predicted that this "background radiation" would have a temperature of "about 5 K" [15].

As the radio astronomy community continued to grow, Alpher and Herman saw the opportunity to look for their predicted relic radiation. Unfortunately, they were told that the signal was too weak so that it could not be observed. However, throughout the 1950s several astronomers reported isotropic background radiation but were unable to identify the significance of their find (such as Andrew Mckellar had described already in 1941 [16]).

[1] Mpc, or megaparsec is a conventional unit used in measuring astronomical distances.

[2] Hans Bethe was only added to the author list to complete the pun "Alpher, Bethe, Gamov".

Then came Arno Penzias and Robert Wilson, two radio astronomers working together at Bell Research Labs in the 1960s. After finishing his Ph.D., Wilson started working at Bell labs, having developed an interest in the 6-m Holmdel "Horn antenna". The Holmdel antenna was a decommissioned instrument used to detect radio signals bouncing off a balloon satellite, making it a perfect tool for the study of weak radio sources. As part of Penzias and Wilsons free research time, they started preparing the antenna for new observations. However, in order to examine faint radio sources, they needed to thoroughly map out the noise properties of the instrument such as the background noise of its environment. They aimed the antenna at a blank spot on the sky with no known radio sources. To their surprise, however, the signal was stronger than expected. Puzzled by their findings, they pointed the antenna in different directions and found the "noise" to be constant across the sky. After an exhaustive examination of the sky, they turned their attention to the instrument itself, checking its electronics and wires to find explanations for the "noise", but without any luck. Being determined scientists, they even scrubbed away all the pigeon droppings on the inside of the horn and proceeded to kill the pigeons.

A whole year went by before they would find the source of the radiation. In 1964 Penzias attended a conference in Montreal where he told cosmologist Bernard Burke about their findings. Burke would later receive a draft of a paper by James Peebles and Robert Dicke on nucleosynthesis where they too, predicted the background radiation. Burke put two and two together and informed Penzias about what he had discovered. After several other experiments confirmed their discovery, Penzias and Wilson went on to win the Nobel Prize in 1978. However, the theoretical explanation for the cause of the radiation was not fully accepted until the 90s with the Cosmic Background Explorer Satellite (COBE). When it did, it disproved the steady-state theory, which had neither predicted, nor could account for the CMB.

2.3 The intricacies of the CMB

In the years following the discovery of the CMB, many experiments were carried out to investigate the phenomenon in greater detail. Some speculated that due to the Earth's motion through space, the signal would be Doppler-shifted and therefore appear to be stronger in one direction, and weaker in the other. This shift would result in a "dipole contribution" in the observed temperature [17]. The speed at which Earth orbits the Sun, and the Sun rotates around the center of the galaxy was an effect that would be measurable, and several cosmologists undertook the challenge of doing it. George Smoot, a particle physicist from MIT, was one of them. After a failed attempt to find anti-matter in balloon-borne experiments, he moved on to the "aether drift" problem; measuring the dipole in the CMB. His team carried out several air-borne experiments which after nearly two years yielded puzzling results. What they had discovered was indeed a dipole, but not in the direction of the rotation of the galaxy, as expected, but rather in the direction of the constellation Leo. By measuring the strength of this Doppler shift, they calculated that the whole of the Milky Way galaxy must be moving

in the direction of Leo with a speed of more than a million kilometers per hour [18]!

The implication of the discovery of the Milky Way's motion through space was that galaxies were not distributed uniformly, but rather clumped together by some attractive force. Meanwhile, the study of rotating spiral galaxies showed signs of invisible matter being present. The idea of such invisible matter was already theorized by Fritz Zwicky 40 years earlier and dubbed "dark matter" [19]. Furthermore, in 1981, Alan Guth proposed the theory of inflation; a period of rapid expansion in the very early universe [20]. This theory predicted that quantum effects would be scaled up to macroscopic scales causing *anisotropies* (deviations from isotropy) in the energy distribution of the universe, which would, in turn, allow for such clumping of matter. As more and more observed phenomena pointed towards the existence of such anisotropies in the CMB, the race to observe them was on. Even though all of these new theories about the structures of the universe pointed towards anisotropies in the CMB, the only observed structure so far had been the dipole. The consensus at this point was that in order to further improve the CMB-data, one had to clean it of atmospheric contamination; i.e., observe from space. The era of space-based CMB missions had begun.

2.3.1 An eye in the sky - COBE

In the wake of the discovery of the CMB, NASA decided to gather all the groups who had proposed space-based CMB experiments and set up a committee made up of well-known researchers such as David Wilkinson and George Smoot. Together they decided on a satellite consisting of three instruments; the "Diffuse Infrared Background Explorer" (DIRBE), the "Far InfraRed Absolute Spectrometer" (FIRAS) and the "Differential Microwave Radiometer" (DMR) which would all sit onboard the "Cosmic Background Explorer" (COBE). DIRBE was to map dust emission from distant galaxies, the Cosmic Infrared Background (CIB). FIRAS was to measure the CMB power spectrum (more on this in 5.1) and DMR to look for anisotropies in three different frequency bands; 31.5, 53 and 90 GHz, carefully chosen to minimize galactic contamination. The mission was launched in 1989 with over 1000 people involved, costing nearly 160 million dollars and became the first space-based CMB mission.

In 1990, the COBE-team announced the FIRAS result which showed the most perfect black-body radiation[3] spectrum ever measured, with a temperature of 2.725K [21]. This was a strong confirmation that the Big Bang theory was correct. Immediately following the FIRAS news, the DMR team presented their preliminary results, the best quality picture of the dipole yet as can be seen in the second image in figure 2.2.

With a high-quality dipole map, and a monopole temperature contribution pinned down to incredible precision, the next step was to find the anisotropies, which would indicate the early signs of structure formation. After having been observing for over a year, people began suspecting that they had not found any anisotropies. If this was

[3]Black-body radiation is a form of radiation emitted from a body which absorbs all incident radiation. This form of radiation can be only emitted by bodies in thermal equilibrium (such as the early universe), and has a well-defined spectrum determined only by its temperature, also known as "Planck's law".

the case, the Big Bang theory was hanging by a thread. By late 1991 the DMR team was beginning to see signs of anisotropies. However, they needed a detailed map of the emission coming from the Milky Way galaxy, which was contaminating their data. After painstaking foreground removal, through the application of external foreground data, they were left with the first, full-sky map of the CMB anisotropies; one of the greatest discoveries of the century. Such data analysis and foreground mapping is to this day of great importance to CMB research, and also the topic of this thesis.

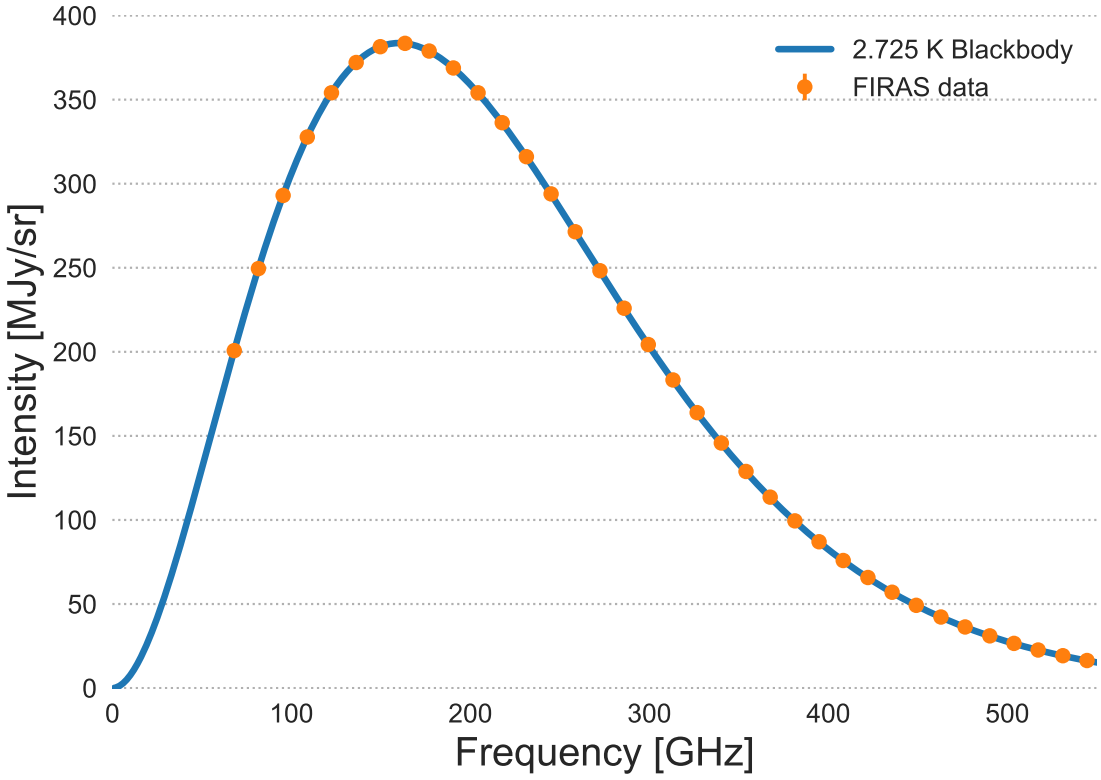


Figure 2.1: A theoretical 2.725 K black-body curve with the FIRAS data overplotted with error bars so small they are invisible. Data courtesy of the NASA CMB legacy archive (LAMDA) [22].

2.4 Improving upon COBE

In the wake of COBE, the primary focus of experimental cosmology shifted towards studying the finer details in the anisotropies. Many smaller experiments were carried out following in its footsteps, such as QMAP which made CMB observations of patches in the sky in 1996 [24]. Such experiments were relevant for pinpointing the first acoustic peak, which corresponds to the first peak in the CMB *power spectrum*. The CMB power

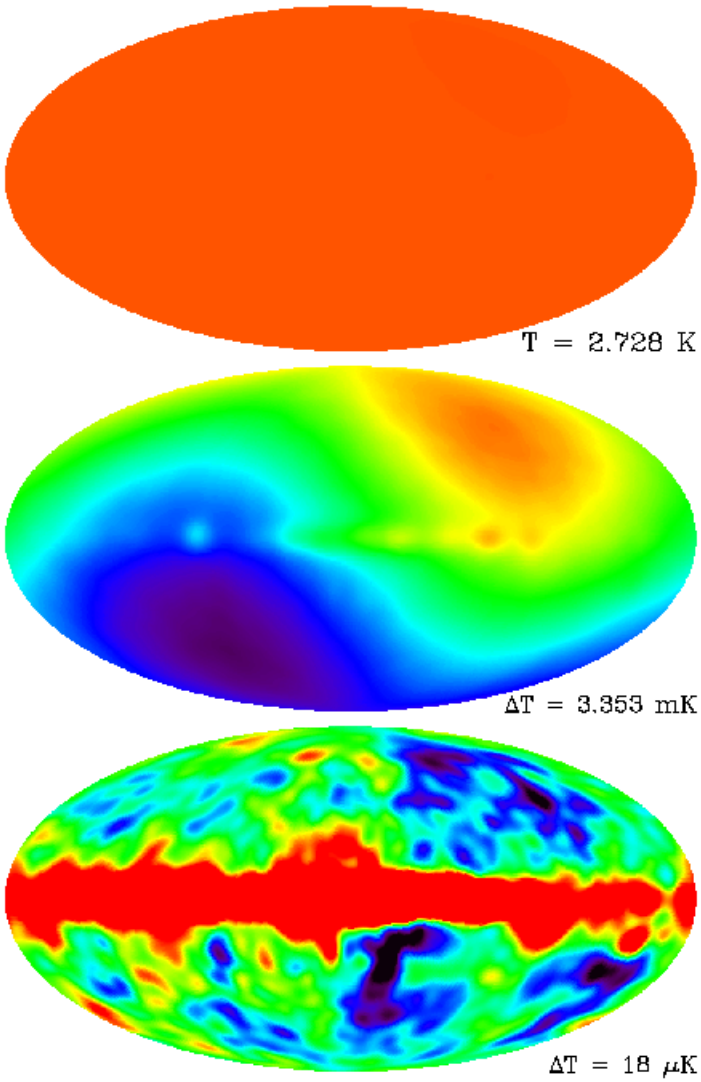


Figure 2.2: From top to bottom: 1) the measured 2.728 K monopole seen uniformly across the sky observed by Pensias and Wilson. 2) The galactic dipole as observed by the DMR instrument on COBE. 3) The COBE DMR full-sky map at 53 GHz smoothed to 7°, showing anisotropies on high latitudes and foreground contamination along the galactic disc. Image courtesy of Lawrence Berkeley Laboratories and NASA [23].

spectrum is perhaps the most essential tool for studying the properties of the universe as a whole. We will get back to this later, but in a nutshell, it describes the relative dominance of different physical scales in the universe. The position of the first peak in the power spectrum, for example, is a probe for the geometry of space itself and determines if the universe is flat or spherical. Therefore, surveys were carried out by experiments such as CBI [25], BOOMERanG [26], and DASI [27], to observe the power on small scales in the universe. Together, they managed to measure the first peak and size distributions of the fluctuations to a precision sufficient to conclude that the universe is indeed flat ($k = 0$ in equation 1.2) [28].

Around the same time, large-scale surveys of the position of extra-galactic objects were carried out which concluded that the universe is not only expanding but that the rate of the expansion is accelerating [29]. The observation of this effect demanded a driving force behind the phenomenon if it were to adhere to Newton's second law. These arguments led to the introduction of the mysterious "dark energy" to the standard model. However, the most recent observations could not account for the amount of dark energy needed to explain the measured rate of the expansion. Moreover, balloon experiments could not obtain the sufficient angular resolution to study the details of the power spectrum which might reveal it. Thankfully, a new experiment called MAP promised the higher accuracy and resolution needed to pin down more critical cosmological parameters such as the age of the universe, the Hubble constant and the relative dominance of baryonic, dark matter and dark energy.

2.4.1 WMAP

In 1993 David Wilkinson set out to improve the results of COBE by first recruiting Chuck Bennett who had worked on the COBE DMR project. Together, they became the Principal Investigators of the next big satellite-borne CMB experiment which they named "Microwave Anisotropy Probe" (MAP). While COBE had probed the large scale anisotropies stretching across the whole sky, the balloon-borne experiments only measure small scales on patches in the sky. The goal of MAP was to close the gap of missing data between these sets of measurements. In addition to observing with improved angular resolution, MAP would also be observing in five different frequency bands, as opposed to COBE's three. Tragically, Dave Wilkinson passed away in September of 2002, and the team decided to rename the satellite "Wilkinson Microwave Anisotropy Probe" (WMAP) in his honor. Having mapped out the full sky, the image showed stunning detail. Among its many findings, WMAP pinpointed the age of the universe to 13.74 billion years, it found the Hubble constant to be 70km/s/Mpc, and that the universe was within 0.4 percent of being completely flat. In addition to this, it confirmed the idea of a universe dominated by dark matter, a full 72 % of its energy density, along with 23 % dark matter and only 4.6 % normal baryonic matter [30].

While WMAP slowly came to the end of its 9-year lifespan in 2010, it passed the torch of space-borne CMB research to Planck, a project funded by the European Space Agency (ESA), which aimed to once and for all pinpoint the cosmological parameters.

2.4.2 Planck

With the era of American space-based CMB-missions coming to a temporary end, two separate European teams of researchers submitted applications for funding for the European Space Agency (ESA) Horizon2000 program. The first one was called COBRAS and was an Italian/American collaboration. The second was SAMBA, a French proposal. ESA chose the best of both worlds, and ended up with the combined project; Planck, named after the Nobel Prize-winning physicist who was the first to explain the black-body spectrum. The main idea of the Planck mission was to make the definitive measurements of the CMB temperature anisotropies. The vast difference in resolution to WMAP and COBE are illustrated in figure 2.3. While the main focus would be on temperature observations, Planck would also be polarization sensitive. This means that it would be able to discern the polarization direction and intensity of CMB photons, which can be used to look for primordial gravitational waves, but more on that later. Additionally, Planck would study the Sunyaev-Zel'dovich (SZ) effect. The SZ effect is when low energy CMB photons scatter off high energy electrons in intergalactic gas, effectively distorting the CMB signal. This effect is especially prominent in galactic clusters or small angular scales. However, this effect, should not be mistaken for the actual small-scale fluctuations produced in the early universe. In order to tell them apart, Planck would observe the effect at different frequencies and angular scales. Furthermore, Planck would observe at nine frequency bands, an upgrade from WMAPs 5, but extending from 90 to 857 GHz as the highest frequency, allowing for more accurate observations of galactic dust, and hence better removal of foreground contamination. The nine frequency channels were split into two different instruments; the Low-Frequency Instrument (LFI), led by Nazzareno Mandolesi in Bologna, Italy, and the High-Frequency Instrument (HFI) led by Jean-Loup Puget in Orsay, France. Together they acted as principal investigators for the satellite.

The total experiment cost approximately one billion euros, compared to WMAPs 120 million. This is a result of Planck being an ESA mission, as opposed to a NASA mission because major ESA projects strive to include all member countries. In October 2012, the project had produced five full-sky maps (two more than planned). After gathering data for four years and six months, Planck was decommissioned and started drifting away from its observation position, between the Sun and the Earth. To illustrate the improvement in precision, we tabulate the most essential parameters found in the experiment along with the results presented by WMAP.

While the Planck experiment is officially coming to an end, there are still a plethora of improvements to be made to the data. This thesis aims to improve the legacy of Planck by correcting some of these issues by revealing systematic effects through component separation. Although there are no current space-based CMB missions, there are myriad of smaller, ground and balloon-based CMB experiments in the works attempting the uncover other secrets hidden within the signal, such as the elusive primordial gravitational waves.

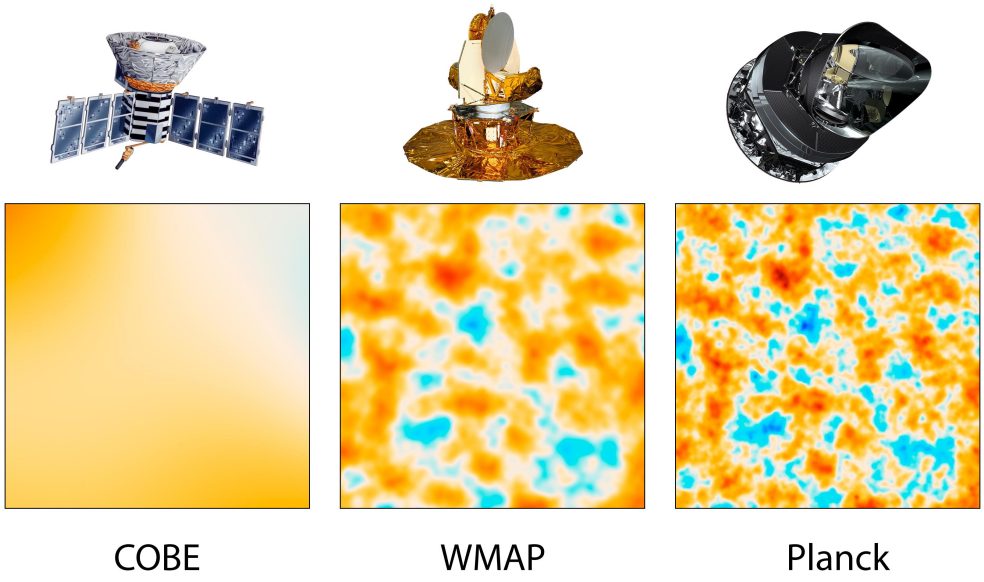


Figure 2.3: A patch of equal area taken from the full sky CMB image for each of the three space-borne CMB missions showing the immense difference in resolution. Image credit NASA/JPL-Caltech/ESA [31]

Parameter	WMAP value	Planck value	Description
	Fitted cosmological parameters		
$\Omega_b h^2$	0.02264 ± 0.00050	0.02225 ± 0.00016	Physical baryon density
$\Omega_c h^2$	0.1138 ± 0.0045	0.1198 ± 0.0015	Physical Dark matter density
Ω_Λ	0.721 ± 0.025	0.6844 ± 0.0091	Dark energy density
n_s	0.972 ± 0.013	0.9645 ± 0.0049	Scalar spectral index
τ	0.089 ± 0.014	0.079 ± 0.017	Reionisation optical depth
	Derived cosmological quantities		
t_0	13.74 ± 0.11	13.813 ± 0.026	Age of the universe [Gyr]
H_0	70.0 ± 2.2	67.27 ± 0.66	Hubble constant [km/s/Mpc]
Ω_b	0.0463 ± 0.024	0.04848 ± 0.00034	Baryon density
Ω_c	0.233 ± 0.023	0.2610 ± 0.0032	Dark matter density

Table 2.1: Best fit cosmological parameters from WMAP 9-year and Planck 2015 baseline analysis [30, 3]. Note that the Planck data incorporated the its polarization power spectra for better accuracy.

2.5 The future of CMB research

At this point, Planck has observed the sky at angular scales down to the limit of what is meaningful for primary CMB temperature fluctuations, and will, therefore, most likely remain the backbone of the field of observational cosmology for decades to come. It will serve as the reference for all smaller scale experiments, and its cosmological parameters will be of great value to many branches of physics. The fact that branches such as theoretical physics and particle physics both rely on these numbers being accurate underlines the importance of making the final product of Planck robust. However, while the CMB temperature data has undergone years of scrutiny, there is little space for fundamental discovery, other than achieving greater and greater precision. While such precision is vital in science, many cosmologists have shifted their attention to another property of the relic radiation.

When looking back in time towards the Big Bang, we observe electromagnetic waves. Because these electromagnetic waves only escaped the primordial soup 378,000 years after the big bang, we have no observations of the universe prior to that. However, events such as inflation are theorized to produce what is called *gravitational waves*, which rippled through space-time at the speed of light. In contrast to electromagnetic radiation, gravitational waves were not restricted by electron collisions during very early times. This allowed them to carry their information almost effortlessly through the universe. Gravitational waves were directly measured in 2016 by the LIGO experiment, a discovery that was rewarded with a Nobel prize [32]. Gravitational waves produced by inflation have yet to be observed, because the signal is theorized to be extremely weak, but if it does exist, they should be visible as swirly patterns in the polarized CMB data. Hence, the race is on to sufficiently clean the polarized CMB data in order to see the weak, swirly signal.

2.5.1 Polarized CMB

To understand the value of the polarized CMB signal, we once again travel back to the primordial universe. It is theorized that during inflation, strong gravitational waves caused by rapid expansion rippled through the fabric of space-time. These ripples stretched and compressed space, causing electrons to see radiation coming from the stretched direction as cooler, and those coming from the compressed direction as warmer. When this electron scattered and emitted a photon, its polarization direction was affected by the direction of the propagating gravitational wave which caused the stretch. This pattern shows up in the polarized CMB signal as swirly patterns or "B-modes" in the polarization direction of the photons, as seen in figure 2.4.

In 2014 the BICEP2 team announced that they had found these elusive swirls while pointing their South Pole radio telescope towards a patch on the Southern sky. This supposed discovery sparked widespread debate throughout the cosmology community about whether the signal was reliable or not. Unfortunately, a drawback of the BICEP2 instrument was that it only observed at one frequency, which made it harder to discern what is signal, and what is foreground. The Planck team argued that, based on its

seven polarization sensitive bands, there was a lot more contamination than what the BICEP2 team assumed. After suffering months of scrutiny, it was revealed that the BICEP2 B-mode signal was indeed an effect of polarized dust contamination [33]. This serves as a reminder of the importance of removing foreground contamination.

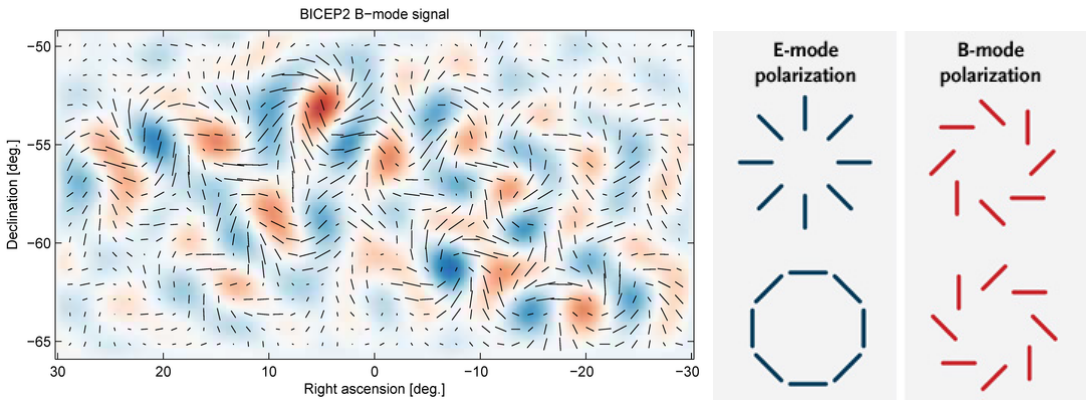


Figure 2.4: Left: Bicep signal showing signs of the swirly B-mode pattern in a patch of the polarized CMB sky. Black lines show the direction of polarization, it's length showing the intensity. Figure from Ade et al. [34]. Right: E- and B-Mode polarization patterns [35].

2.5.2 Future missions

Many new experiments are planned for the future, but as mentioned earlier, the focus has shifted towards observing polarized CMB. At the moment, the most advanced such experiment is the proposed Japanese (JAXA) satellite named LiteBIRD [36]. Its goal is to map out the full sky in polarization, using six bands with over 2000 sensors. As of today, it is the most promising of upcoming space-based projects. Smaller scale balloon experiments such as SPIDER [37], and LSPE[38], are also in the works. Both of which aims to observe the B-modes at different angular scales. In addition to space and balloon-based missions, efforts are also taken to observe the polarized CMB from the ground. Experiments such as the QUBIC experiment[39], observing from the Atacama desert along with the CMB stage 4 (CMB-S4) mission[40], with separate contributing arrays in both Chile and the Antarctic. Other upcoming experiments with indirect effects on CMB research, although beneficial for component separation are C-BASS[41], S-PASS[42] and QUIJOTE [43], which aim to map out the low-frequency domain and to study synchrotron, free-free and spinning dust effects. Another big experiment further into the future is ESAs Euclid satellite [44]. Euclid will study the effect of weak lensing to improve our understanding of dark matter and dark energy by a large-scale survey of galaxies. This experiment aims to reveal the later stages of cosmic evolution, as opposed to the early years observed by Planck and other CMB instruments.

Part II

Methods

Chapter 3

The microwave sky

"I always believe that the sky is the beginning of the limit."

— MC Hammer

As COBE, WMAP or Planck scans the sky with their microwave-sensitive eyes, they do not see the CMB signal with perfectly visible anisotropies. Instead, the microwave sky is contaminated by a mix of components such as thermal dust, synchrotron, and free-free radiation, illustrated in figure 3.1. As the name suggests, the Cosmic Microwave Background is the backdrop to the microwave sky, and by virtue of the cosmological relationship between distance and time, the first light ever emitted. Therefore, in order to study the CMB, we must extricate the foregrounds. For keeping track of the various sky components, we construct a simple mathematical model on the form,

$$\mathbf{d} = \mathbf{s} + \mathbf{n} + \sum_i \mathbf{f}_i, \tag{3.1}$$

where $\mathbf{d}$ is the observed data, $\mathbf{s}$ is the CMB signal, $\mathbf{n}$ is the Gaussian instrumental noise and $\mathbf{f}$ are the various foreground components. In this chapter, we will further investigate the last term of the model to gain an intuitive understanding of the composition of our data signal, as well as the relevant morphology of the foregrounds useful for the component separation analysis to come.

3.1 Foreground components

With a goal of distinguishing each type of signal in the added image, we must get to know the culprits of the contamination, because each of them possesses their unique morphology and physical properties. For example, the dipole is a result of our movement through space relative to the CMB, and should not be confused with the behavior of dust particles in the Milky Way galaxy. While our focus lies on the study of the CMB, the component separation process is an important tool also for studying other astronomical phenomena. The Planck satellite provides high-resolution full-sky measurements in a broad range of frequencies, as illustrated in figure 3.1. Across these frequencies, we see

many different types of radiation. This radiation carries important information on the composition of our galaxy and is therefore vital for many branches of astronomy.

One of the most important characteristics of these foreground components is that their intensity is strongly frequency dependent. By observing them at different frequency channels, one can map out their behavior and thereby differentiate them. As illustrated in figure 3.2, each component can be characterized by a unique frequency spectrum. In some cases, the lack of "uniqueness" of their model leads to problems, as is the fact with synchrotron and free-free emission. We characterize this lack of uniqueness as a *degeneracy*. When two frequency spectra are very similar and respond equally to change in parameter values, we call them degenerate, because they are almost impossible to distinguish from one another. Breaking degeneracies in component separation is paramount to achieve a physical model of the sky, and one of the most significant challenges we face. In addition to the colored lines in figure 3.2, the thin grey vertical lines correspond to the sharp line emission from carbon monoxide. This effect is visible at three different frequencies, corresponding to three different electron transitions.

Moreover, the foreground regime looks different for the temperature signal and the polarized signal. Note that in the figure to the right, the CMB signal is weaker than the foregrounds at all frequencies, which implies that doing component separation is much more demanding for polarization than temperature. This figure also underlines the importance of observing with many different frequency bands. The grey areas in the figure represent each of these bands for the Planck experiment, which spans the interesting parts of the CMB-frequency domain.

We will now go through the most relevant components for the analysis carried out in this thesis. They are ordered according to their relative importance and prominence in the total sky signal.

3.1.1 Monopole and Dipole

We already mentioned this effect, but it is so dominating that we have no choice but to include an in-depth discussion of it. The prominence of the dipole effect is evident in the raw sky map for the 217 GHz channel as observed by Planck in figure 3.3. As we have learned in chapter 2, the CMB dipole is the first foreground one encounters when wanting to observe the CMB, and perhaps also the simplest. The dipole is the Doppler effect produced by our movement through space relative to the CMB. It appears in our observations as a blueshift in the direction we are headed, and a redshift in the direction we are moving away from. The amplitude of this Doppler effect is approximately 3.4 mK [45]. This effect is sometimes removed during the process of converting time-ordered satellite data into maps, known as map-making. However, this it should preferably be left to component separation. Since the small uncertainties regarding the direction and amplitude of the dipole affect other components.

One often distinguishes between the orbital and the galactic dipole. In this thesis, we will always refer to the galactic dipole when specified otherwise. The orbital dipole is the effect of the satellite's orbit around the sun, which is accounted for before we get

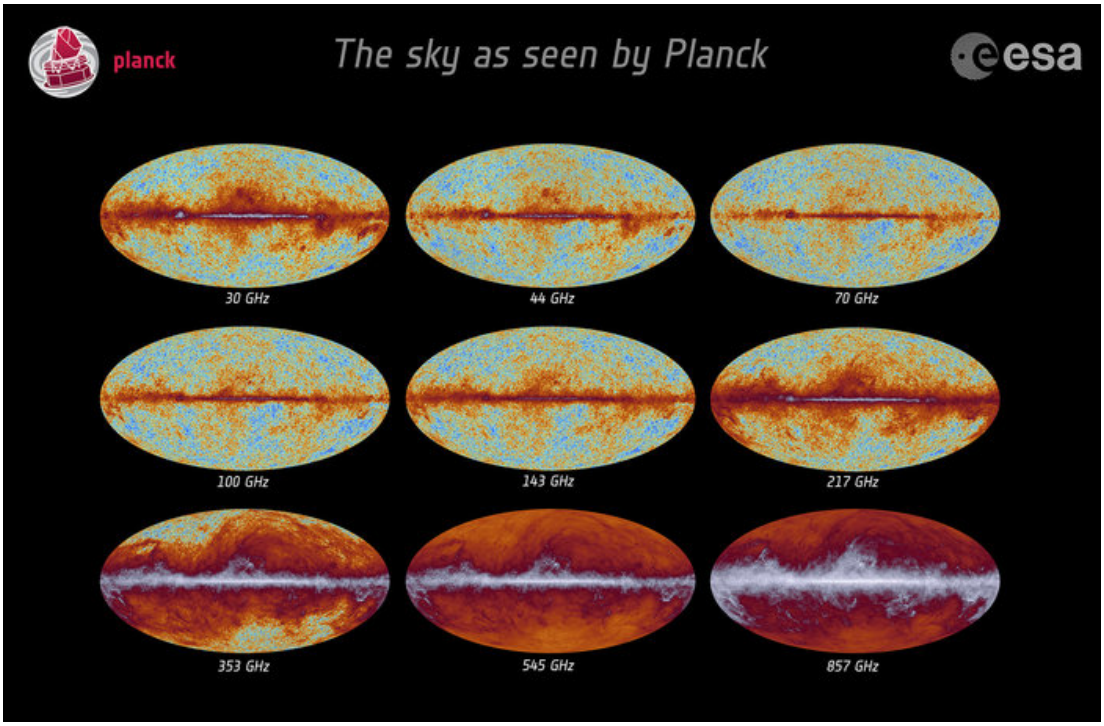


Figure 3.1: All 9 Planck frequency bands showing the various foreground contamination across frequency bands. Low frequencies are heavily dominated by free-free and synchrotron emission, while the highest are almost direct probes of thermal dust emission. Moderate frequencies contain most of the CMB signal. Image courtesy of ESA and the Planck Collaboration.

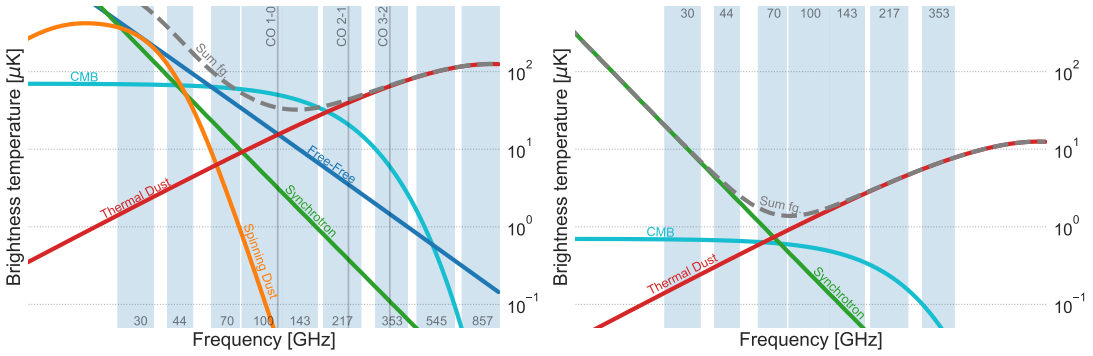


Figure 3.2: The brightness temperature frequency spectrum of each component of the microwave sky in temperature (left) and polarization (right). Grey bands are the frequency bands observed in by Planck. We note that the temperature CMB signal is brighter than combined foregrounds, while the polarized CMB signal is weaker, making component separation for polarization more challenging.

our hands on the data.

Because we wish to study the anisotropies of the CMB, we must account for the 2.7255 K offset across the sky. This is the CMB mean temperature monopole contribution. When accounting for this, we may not simply subtract a constant factor across the sky. The reason for this is that all current CMB experiments (except FIRAS) are differential, and only measure variations relative to an arbitrary reference (like Planck) or Saturns two positions on the sky (like WMAP). The true monopole is therefore unobservable and must be estimated through other means after map-making.

3.1.2 Thermal dust

As is illustrated in figure 3.1 and 3.2, the 100 GHz band, thermal dust becomes the dominating foreground and completely envelopes the sky. This type of radiation is produced by hot, vibrating silicate or carbonaceous dust grains scattered throughout the Milky Way galaxy. The current most popular model for its intensity behaviour as a function of frequency is a modified black-body (MBB), also known as a grey-body. This model has three variables; the amplitude A_{d}, the emissivity index β_{d} and the dust temperature T_{d}. The mathematical expression is

$$s_{\mathrm{d}}(\nu, A_{\mathrm{d}}, \beta_{\mathrm{d}}, T_{\mathrm{d}}) = A_{\mathrm{d}} \cdot \left(\frac{\nu_0}{\nu}\right)^{\beta_{\mathrm{d}}+1} \frac{e^{\frac{h\nu_0}{k_{\mathrm{b}}T_{\mathrm{d}}}}-1}{e^{\frac{h\nu}{k_{\mathrm{b}}T_{\mathrm{d}}}}-1}, \tag{3.2}$$

where ν_0 is the reference frequency, k_{d} is Boltzmann's constant, and h is Planck's constant. The β value was previously thought to be between 1.7 and 2.0, which leaves the model with a much steeper tail in figure 3.2, resulting in more CMB in the 100 GHz band. The value of T_{d} determines where the peak of of the model is. The latest analysis find the dust temperature to be approximately 21 K, and β_{d} to be around 1.6 [47]. The MBB sufficiently describes the thermal dust ensemble up to 857 GHz, beyond that point, the composition becomes too complicated for such a simplified model. In an ideal world, one would split this dust model into its different components, but because the various dust types are profoundly degenerate, it would require more data to model them correctly.

Along with the dust grain emission from our galaxy, the model also encapsulates CIB. CIB is the Cosmic Infrared Background radiation, produced by red-shifted dust from distant galaxies. Because of the likeness in frequency behavior to local thermal dust, we cannot easily distinguish the two. However, it is possible to exploit the different angular power spectra of the two signals, for instance as done with GNILC (Remazeilles et al. [48]), to perform a reasonable separation between the two.

The morphology of the thermal dust component can be described as a thick galactic emission band with a shark-fin like feature North of the galactic plane. Furthermore, when we encounter thermal dust in our analysis, we distinguish it by its faint morphology stretching far into high latitudes. Its prominence is quite evident in the 353 to 857 GHz sky maps in figure 3.1.

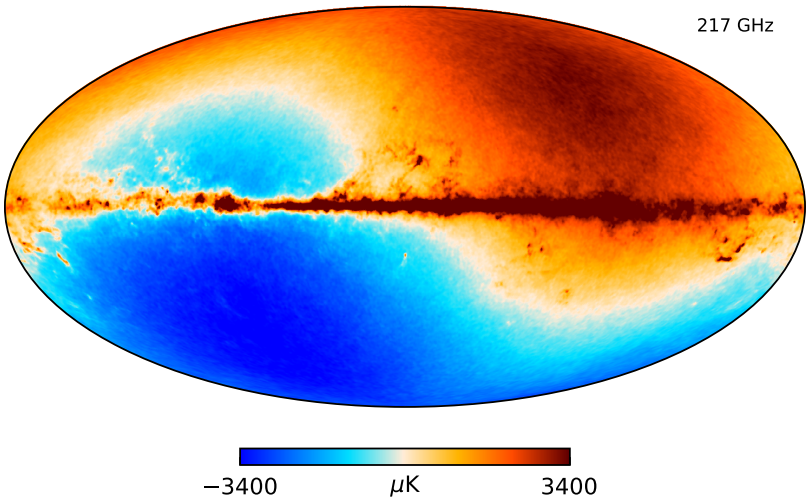


Figure 3.3: Actual sky at 217 GHz as seen by Planck, showing the prominence of the dipole effect.

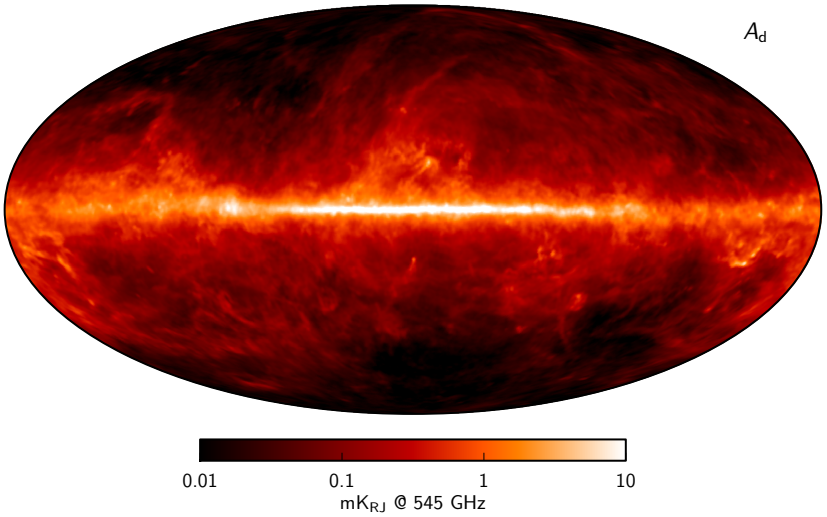


Figure 3.4: Maximum posterior thermal dust intensity maps from the joint 2015 analysis of Planck, WMAP and Haslam observations. Image courtesy of the Planck collaboration [46]. One distinctive feature is the shark-fin shape above and to the right of the galactic center. Additionally, thermal dust envelopes much of the sky, dust contamination is therefore often seen as a thick galactic plane.

3.1.3 Synchrotron radiation

One of the foreground components encountered at the lower frequencies observed by Planck is synchrotron radiation. This type of radiation is generated by relativistic cosmic-ray electrons spiraling in the galactic magnetic field and emitting photons. The identification of this foreground signal is especially important in the search for B-modes, as it is highly polarized, and therefore one of the two main contaminants in the polarized CMB signal.

We model the synchrotron component through a power law with $\beta_s \approx -3$, at frequencies higher than 20 GHz. Therefore, one of the challenges we face when trying to identify this signal in the weaker data is that Planck does not probe any of the frequencies dominated by synchrotron emission. As we see in figure 3.2, the low-frequency band observed by Planck is highly degenerate with free-free and spinning dust emission. Preferably, one would have an additional band at very low frequencies, probing the signal at its strongest. This is why the Haslam data was included in the 2015 joint Planck analysis. On top of the nine frequency bands observed by Planck, this analysis added 5 WMAP bands as well as the Haslam survey at 408 MHz from 1982 [46]. In the future, however, experiments such as S-PASS [42] at 2.4 GHz, C-BASS [41] at 5 GHz and QUIJOTE [43] at 10-40 GHz will bring new and vital data, which will help to distinguish the synchrotron signal from free-free and spinning dust adequately.

The morphology of synchrotron can be seen in figure 3.5. Characteristic traits include the strong signal in the very center of the galaxy, and a prominent column of radiation stretching up towards the galactic pole, known as the "North galactic spur".

3.1.4 Free-Free radiation

Diving deeper into the soup of degenerate foregrounds we encounter the free-free signal. Free-free emission, or Bremsstrahlung, is formed during electron-ion collisions, such as proton collisions accelerating the electron which in turn emits a photon. Thanks to its strong correlation with the $H\alpha$ line, the morphology and properties are reasonably constrained using a template derived by Dickinson et al. [49]. Hence, free-free emission can be modeled with only two parameters, resulting in an effective power law behavior in intensity with a $\beta_{ff} \approx -2.13$. The two parameters that make up the model are the emission measure, EM, which is the integrated squared electron density along a line of sight and the electron temperature, T_e. In the relevant frequencies for us, EM effectively adjusts the amplitude of the model while T_e slightly changes the power law index.

Notable morphology for the free-free component includes the Gum and Orion nebulae. The Gum Nebula is identified by its doughnut-like shape stretching South from the right side of the galactic plane. The Orion nebula is the half-moon shape further right of the Gum nebula. In addition to these, both the prominent blob North of the galactic center and the overall thin galactic disc are unique features of this component.

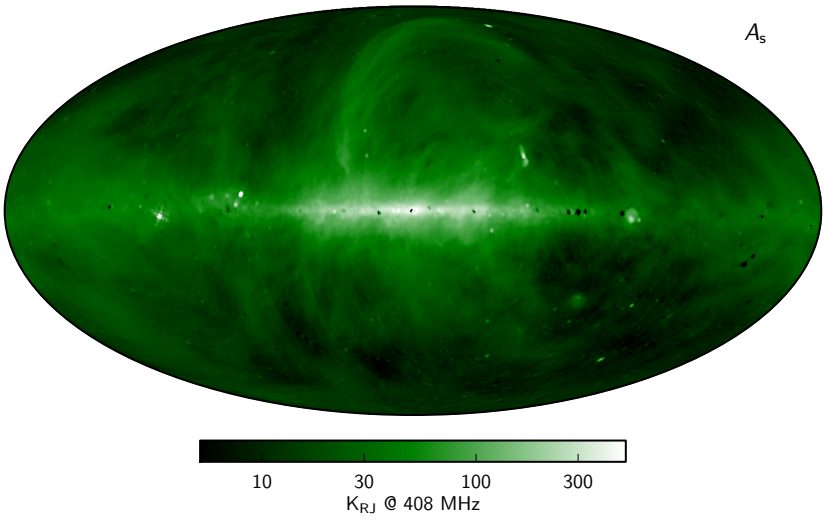


Figure 3.5: Maximum Posterior synchrotron intensity maps from the joint 2015 analysis of Planck WMAP and Haslam observations. Image courtesy of the Planck collaboration [46]. The characteristic morphology of synchrotron emission is its thicker shape around the galactic center, thinning out as we move towards the out edges of the Milky Way. Another feature is the "the North galactic spur"; a visible band of emission stretching up towards the North galactic pole.

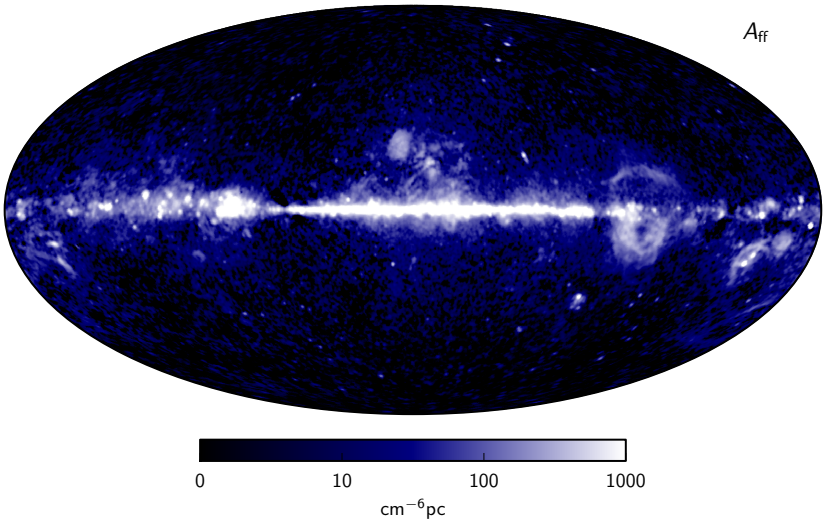


Figure 3.6: Maximum posterior free-free emission maps from the joint 2015 analysis of Planck WMAP and Haslam observation. Image courtesy of the Planck collaboration [46]. We note two distinct features of its morphology, the blob North of the galactic center, as well as the Gum nebula, which is the doughnut shape stretching Southward from the right part of the galactic plane. Next, to the Gum nebula, we also see the Orion nebula further towards the edge of the figure.

3.1.5 Carbon Monoxide (CO)

One of the main characters of this thesis is the carbon monoxide (CO) line emission. This emission is observed at sharp frequencies 115, 230 and 345 GHz, for transitions $1 \rightarrow 0$, $2 \rightarrow 1$ and $3 \rightarrow 2$, respectively. The combined emission for all lines is illustrated in 3.7. This elusive component was not taken into consideration when vital decisions about the detector bands of the Planck instrument were made, and therefore caused a great deal of headaches when the HFI data were revealed to be heavily contaminated by it. While they are illustrated in figure 3.2 as sharp lines, their spectra can be described as delta functions of decreasing amplitude. One of the primary goals of this thesis is to remove this contamination robustly, but contrary to the other foreground models, we cannot observe their behavior across frequency bands due to their sharp nature. More on this later.

Because both CO and interstellar dust were created during star formation, the morphology of their signal is similar. However, CO is most prominent closer to the galactic plane, exhibiting little emission on high latitudes. Furthermore, the CO has a distinct, sharp and thin morphology. Two prominent features are the shark-fin, just North and to the left of the galactic center, and the fan-region. The fan region is the noticeable signal stretching out from the galactic plane on the far left side of the map. Lastly, we see the Orion nebula underneath the galactic disk on the right side of the map.

3.1.6 Spinning dust

The spinning dust component was first introduced as the "anomalous microwave emission" or AME, by Leitch et al. [50]. In this thesis, we refer to it under the name of "spinning dust". Its origin is still not fully established in the scientific community. However, in recent years, more and more observational evidence point towards the spinning dust theory. Spinning dust emission is emission from small dust grains with non-zero electric dipole moment rotating at gigahertz frequencies.

Spinning dust is one of the three strongly degenerate foreground components encountered in the low-frequency domain, diminishing our ability to model its spectrum. However, for the majority of this work, the spinning dust component will be merged with free-free and synchrotron to form a joint low-frequency component. We will therefore not pay too much attention to it, as the elusive component is a tale of its own.

Its morphology, however, is worth noting. As can be seen in figure 3.1, the 30 GHz channel is contaminated by something that looks like dust. However, thermal dust is not dominant at such low frequencies. This signal is instead generated by the dust present at the same physical locations, but by spinning dust grains which emit light at a different frequency. The result is a similar morphology, which is easy to tell apart, as they appear in completely different frequency bands.

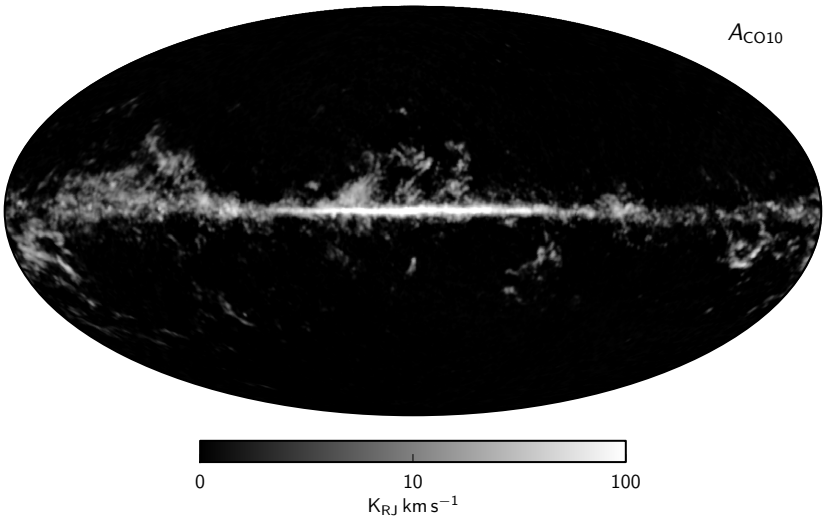


Figure 3.7: Maximum posterior for CO $J = 1 \rightarrow 0$, $J = 2 \rightarrow 1$ and $J = 3 \rightarrow 2$ line emission maps from the joint 2015 analysis of Planck WMAP and Haslam observations. Image courtesy of the Planck collaboration [46]. The CO emission exhibit very distinctive features, such as the shark-fin seen in thermal dust, but this time more prominent. In addition to this, a tell-tale sign of CO is the morphology of the "fan region" on the left side of the map, as well as the Orion nebula underneath on the opposite side.

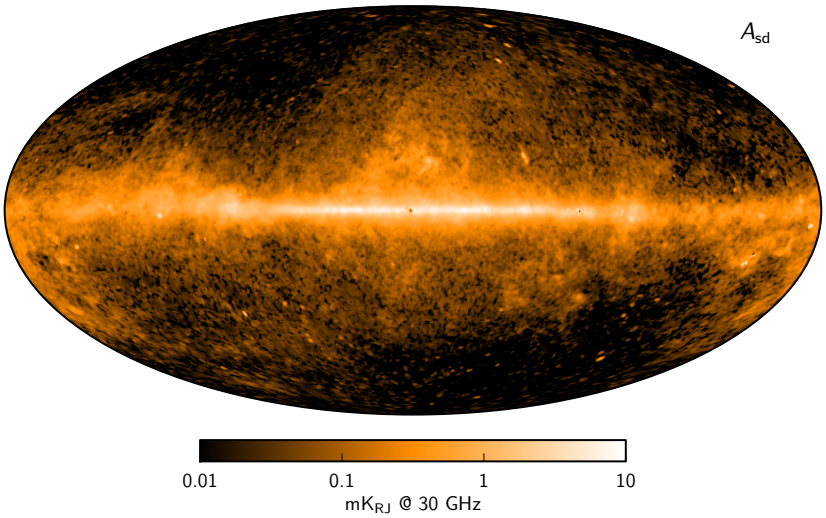


Figure 3.8: Maximum posterior spinning dust intensity maps from the joint 2015 analysis of Planck WMAP and Haslam observations[46]. Its morphology is reminiscent of thermal dust, which is not surprising considering it is generated at the same physical locations but through a different process, which emits radiation at different frequencies to its thermal counterpart.

3.1.7 Additional components

Zodiacal Light

In addition to all galactic and intergalactic foregrounds, we must account for one contaminant within our solar system. As Planck scans the sky, it attempts to avoid as much sunlight as possible. Unfortunately, sunbeams light up dust particles scattered around the solar system creating a distinct pattern visible in the full sky Planck data. To get rid of this imprint, we fit templates made explicitly for each full-sky data set. The reason for this is the Sun's variability with time and season, requiring different templates for each dataset.

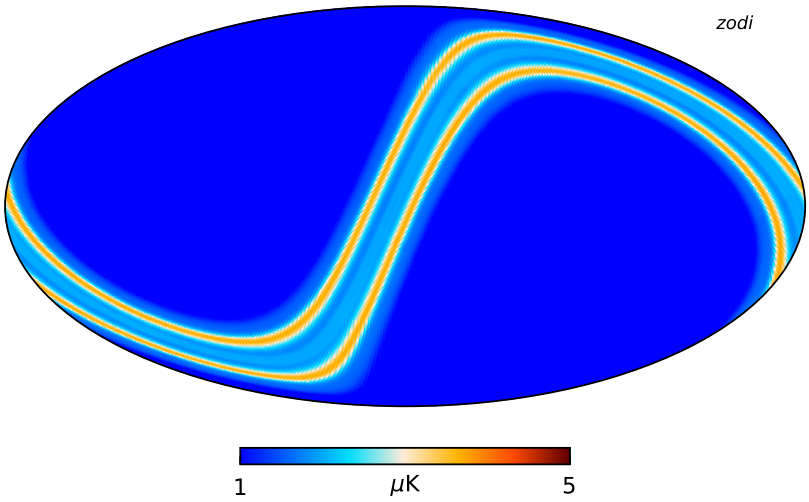


Figure 3.9: Zodiacal light emission template.

Thermal Sunyaev-Zeldovich

As mentioned in chapter 2, the thermal Sunyaev-Zel'dovich effect causes hot electrons in clusters to boost the CMB photons. This causes the signal to deviate from a perfect black-body. This effect is quite weak, and only significant from the very brightest clusters. In order to fully map out its entire effect during global component separation, we must accomplish better control over the systematic effects of the signal. In addition to this, at the angular resolution considered in this thesis, the effect is negligible.

Chapter 4

The Planck instrument

"Technology is cool, but you've got to use it as opposed to letting it use you."
— Prince

The Planck instrument and the process of data reduction bring further complexity to our data model. Some of these effects are just as vital as foreground contamination, and in some cases even more so. No instrument is perfect, and neither are the ones on board Planck. In this chapter, we will discuss the most critical systematic effects attributed to the instrument itself. While we focus on the Planck satellite in this thesis, the same operations are relevant for all similar instruments.

Before we receive our data on the form of full sky maps, it has undergone careful pre-processing, such as flagging and calibration. Flagging is the action of identifying and removing bad data from the data stream, while calibration is the process of translating the measured voltages of the satellite into correct temperature information. After these operations, the time-ordered data is assembled into full sky maps through the process of map-making. We will describe this process more in depth in the next chapter, but in a nutshell, the job of a map-maker is to interpret a "time-stream" of pointing vectors and temperatures into a full sky map. This signal jigsaw puzzle must account for many of the effects described in this chapter, and the rest is up to us to fix while doing component separation.

From a component separation point of view, we may write down our data model once again on the form

$$\mathbf{d}_\nu = \mathbf{s}_\nu + \mathbf{n} \tag{4.1}$$

$$\mathbf{s}_\nu = g_\nu \sum_{i=1}^{N} F_\nu^i(\beta_i, \Delta_\nu)\mathbf{a}_i, \tag{4.2}$$

where we have added a couple of instrumental corrections. Here, g_ν is a gain correction for the frequency channel ν, and Δ_i is a bandpass shift.

As a result of the digitization of the CMB photons, and the observation beam, our data is pixelized and smoothed. These operations may be described mathematically by a linear operator, $\mathbf{A} = \mathbf{PB}$, where $\mathbf{P}$ is the "pixel window", and $\mathbf{B}$ is the "beam".

The corrections described above, however, are only the corrections we make during component separation. There are a plethora of other instrumental effects which must be accounted for during map-making, some of which only become visible after component separation. In order to identify these, we must familiarize ourselves with the details of the instrument.

4.1 Instrument

All instruments have their pros and cons. The COBE and WMAP satellites chose to carry differential radiometers, which measures the sky in two different directions and observes the difference. Planck LFI also used radiometers, but instead used a reference load of 4 K on the satellite to measure signal difference. Planck HFI, on the other hand, used what is known as bolometers. Bolometers measure the total incident energy of incoming photons. Because of the low temperature of the CMB photons, the instrument needed to be cooled down to extreme temperatures of about 0.1 K, even colder than that of space. The process of keeping the instrument properly cooled at all times was a challenging feat and a process not without drawbacks.

While Planck looks at the sky and measures changes in wattage, we are more interested in a physical image of the temperature fluctuations on the sky in Kelvin. Therefore, we construct a model to interpret the raw satellite data, a_i, from each detector i,

$$a_i = K_i \epsilon_i \int d\nu (A\Omega)_\nu \tau_i(\nu) dI_\nu \ [\mathrm{K_{CMB}}], \tag{4.3}$$

where K_i is the photometric CMB dipole calibration factor for converting between watts and Kelvin, $(A\Omega)_\nu$ is the telescope beam for each frequency. The frequency dependence of the beam is illustrated in figure 4.1, where the angular size of the beam is dependent on the frequency observed. We must also model $\tau_i(\nu)$, which is the normalized spectral transmission or bandpass spectrum, and ϵ is the optical efficiency of the instrument describing how many photons survive to the end of the instrument system. Lastly, dI_ν denotes the integration over source intensity [51]. As we move through this chapter, we will refer back to this model, as we discuss its relevance in component separation in more detail.

4.2 Focal plane

The focal plane of the Planck instrument consists of the HFI instrument, seen in figure 4.1 as bolometer horns arranged on a disk surrounded by the more massive LFI horns. The HFI instrument consists of 36 horns with a total of 52 bolometers observing at six different frequencies; 100, 143, 217, 353, 545, and 857 GHz. The LFI instrument totals 11 horns covering the lower three frequencies 30, 44, and 70 GHz.

We also distinguish between two types of sensors, namely the polarization sensitive bolometers (PSBs) and spider-web bolometers (SWBs). The SWBs are designed to measure temperature, and therefore ideally polarization insensitive. However, the

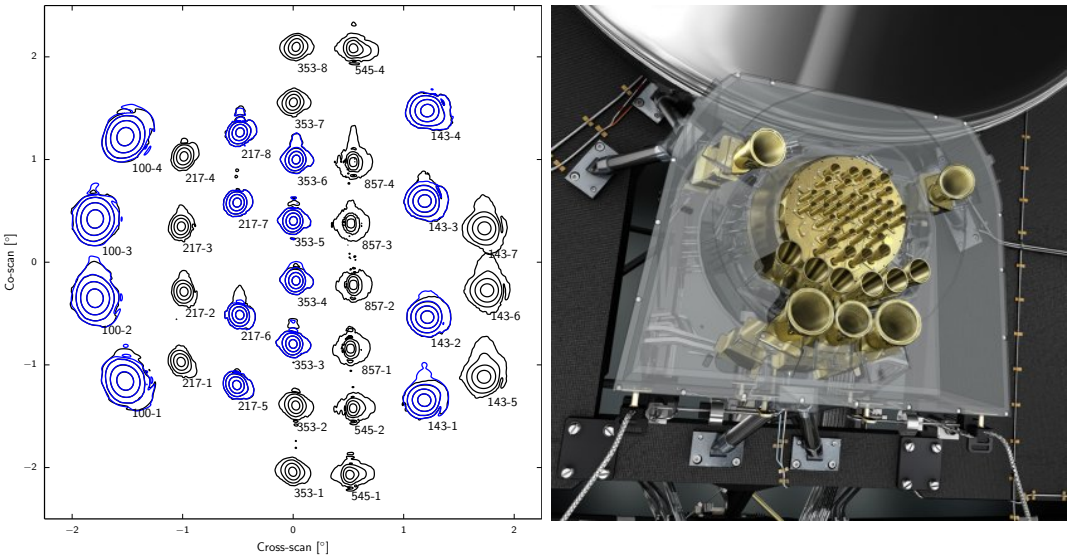


Figure 4.1: Left: HFI scanning beam contours, showing the asymmetry and position of the beams. Note that higher frequency beams are smaller, and blue signifies a spider web bolometer [47]. Right: the Planck focal plane [52].

polarization efficiency has been found to be 1.6% and 8.6% of the full signal, but with high uncertainty [53]. The PSBs consist of a rectangular grid aligned in a specific direction. Incoming photons with electrical fields parallel to this direction generate currents, and therefore deposit energy, allowing polarization reconstruction. All PSBs share their horn with another PSB aligned perpendicular to the other, so as to measure the Stokes parameters I and Q for polarization. Another PSB is rotated by 45 degrees as to measure the U stokes parameter along the same scanning line. The $Q - U$ stokes parameters measure linear polarization, while I measures intensity. In the analysis of the last chapters, we split the PSBs into single bolometers, and label them either a or b, such as $100 - 1a$ and $100 - 1b$.

4.2.1 Instrumental beams

Each detector carries subtle differences in a variety of ways. The first effect we shall talk about is the instrumental beam. Because a telescope does not register temperature at a single point, but rather from a finite solid angle, we must account for the size and shape of this registered area, mathematically denoted by $(A\Omega)_\nu$ in equation 4.3. Asymmetric beams are one of the most important sources of systematic error in CMB missions. The scanning beams of the HFI instrument can be seen in figure 4.1. The beam area can be described as a density, where more photons come in from the center than from the sides. As they measure the sky, we must also account for the orientation as the satellite rotates as well as its scanning strategy. Because we generally assume a symmetric beam during analysis, the observed signal $I(\hat{p})$ may be approximated by full-sky convolution

of the physical CMB signal $s(\hat{n})$,

$$I(\hat{p}) = \int s(\hat{n}) b(\hat{n} \cdot \hat{p}) d\Omega_{\hat{n}}, \tag{4.4}$$

where b is the beam density. This operation takes the form of the linear transform $\mathbf{B}$. This smoothing effect is however not a significant problem in this thesis, as all of our data is smoothed to 40 arc-minutes, sufficient for identifying large-scale systematics in our analysis. If we were interested in sampling the power spectrum, however, we would be forced to account for this in order to discern small scales and properly sample high multipoles.

4.3 Pixel window

Another similar effect to the beam is the smoothing that occurs when we digitize the image and convert photon information to pixel representation. The CMB has theoretically infinite resolution, but our computers work with pixelized maps. When we digitize our signal, the photon information is therefore binned into each pixel, effectively averaging over a small area of the sky. To correct for this, we apply the pixel window function. The pixel window is often combined with the beam effect in the data model as $\mathbf{A} = \mathbf{PB}$. As with the beam, our resolution suffices for this analysis. However, if we require an estimate of the power spectrum, this effect must be accounted for to obtain the physical CMB signal.

4.4 Gain calibration

Because we are dealing with many separate detectors, we must make sure that they are all calibrated to measure the same intensity. Relative calibration of an instruments response rate with time is something that must be accounted for in the TOD during map-making. However, we may attempt to remedy sub-optimal gain estimation by adjusting the mean value with a multiplicative factor g_ν in equation 4.2. This multiplicative factor changes both the CMB signal and the foregrounds equally. Fortunately, the CMB is a perfect black-body, and we, therefore, know its signal strength across frequencies, which we may calibrate against.

Unfortunately, calibration such as this becomes increasingly complicated with the lack of CMB in the sky maps. In the high signal-to-noise domain of the 545 and 857 GHz bands, we are in practice only observing one component, thermal dust, which leaves us with no reference signal to anchor our calibration. While we may increase the gain value for a map, we may also decrease the amplitude of the thermal dust component by an equal amount. This makes the gain and amplitude parameters perfectly degenerate in the case of the 545 and 857 GHz sky maps. Therefore, the only way of calibrating the component is with the intensity of itself along its predefined frequency spectrum. With a physical representation of the thermal dust frequency spectrum given by β_d and T_d, we may calibrate these channels at least to a certain degree of accuracy. Ideally, these channels should be calibrated by correcting with FIRAS.

4.5 Bandpass

The instrument bandpass profile appears as the τ factor in equation 4.3. A detector bandpass profile is the shape of its relative response across its full bandwidth. Once again, the instruments are not exact, and they do not measure at sharp frequencies, but rather in bands of $\pm 12.5\%$ of the center frequency for LFI and $\pm 15\%$ for HFI. This is illustrated by the width of the grey bands in figure 3.2 chapter 3.

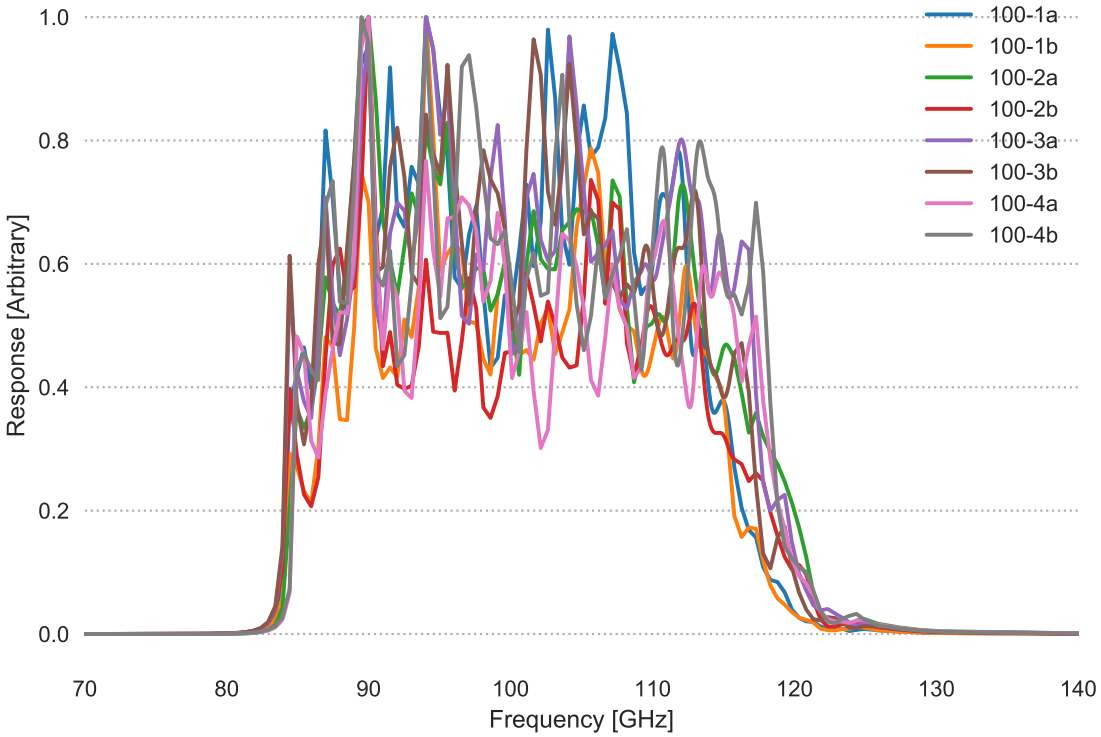


Figure 4.2: The bandpass profiles of all 100 GHz bolometers showing small differences in sensitivity and position.

The bandpass tells us the amplitude, and the bandwidth of each of the detectors, and the shapes of these are shown for all 100 GHz bolometers in figure 4.2. This profile enable us to translate the foreground spectrum we construct into the signal we observe through unit conversions and color corrections [46]. We also note the slight differences in each of the bandpass profiles in figure 4.2. These small differences must all be accounted for when constructing a complete map for each frequency, however, in some cases, by employing the maps separately, we may use these subtle differences to our advantage because they serve as separate data points. The bandpass profiles are inherent properties of the detectors, measured to the best of our ability before the instrument left earth. This process is, as many facets, not completely accurate and can lead to many different issues. In this thesis, we distinguish between two common problems regarding the bandpass profiles, namely bandpass error and bandpass mismatch.

4.5.1 Bandpass Error

One of the effects attributed to a wrongful measurement of the bandpass profile is the "bandpass error" [46]. This effect causes a multiplicative spectral-index dependent scale error, which means it is highly degenerate with gain errors. Fortunately, there is a way of breaking this degeneracy. Because frequency maps are calibrated against the CMB dipole, the anisotropies are independent of frequency; hence the only component sensitive to shifts in the bandpass is the foregrounds. Therefore, when shifting the bandpass back and forth across frequencies, it affects only foregrounds, as opposed to gain corrections, which also scales the CMB. This effect breaks the degeneracy and allows for proper bandpass calibration. We parameterize this translation as

$$b(\nu) = b_0(\nu + \Delta_\nu), \tag{4.5}$$

where $b_0(\nu)$ is the nominal bandpass, $b(\nu)$ is the fitted bandpass, and Δ_{nu} is our applied frequency shift. We stress that this is not an accurate description of the bandpass error phenomenon, but rather a sufficient description for our component separation purposes. A physical model of bandpass errors requires information about tails and tilt of the bandpass profiles. Similarly to the gain calibration, bandpass shifts struggle to reach a physical result in the high signal-to-noise domain, where there is no CMB to anchor to. A bandpass error is very much a non-linear effect, as opposed to gain calibration, which makes the correction even more challenging.

4.5.2 Bandpass Mismatch

The bandpass mismatch problem is a similar issue, but with different effects on our component maps. Bandpass mismatch is attributed to the difference in the overall shape of the bandpass profile. While two detectors measure the same pixel, their bandpass profiles differ, which in practice means that they observe different temperatures for the same patch of the sky. The measured difference between detector is interpreted by our map-making algorithm as correlated noise, which the map-making algorithm then tries to suppress across the sky. The type of correlated noise removal applied in recent CMB analysis, therefore, causes the signal to be smudged out, following the scanning strategy of the satellite. The bandpass mismatch is easily recognized as a dolphin-line morphology in frequency residual maps. Unfortunately, we can not correct for this in component separation, since it is defined in time rather than space.

4.6 Instrumental noise and the RMS Map

Finally, we have the instrumental noise, $\mathbf{n}$, which has been present in our model since chapter 3. This noise is assumed to be Gaussian distributed with a known covariance matrix N_ν. Although the noise does not vary much in time, the effect of Planck's scanning strategy causes some areas to be scanned more times than others, reducing its variance more in some pixels than others. This phenomenon is illustrated in 4.3, as a root-mean-square (RMS) map of the variance of the noise. Because of the relationship

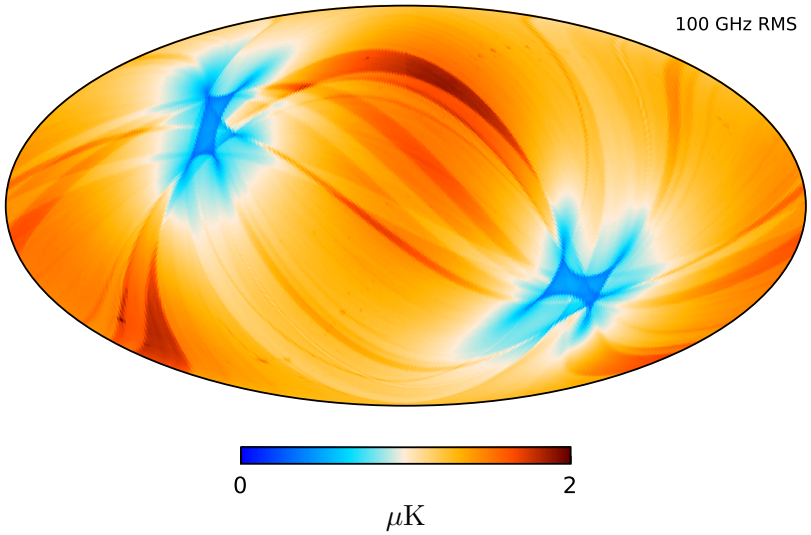


Figure 4.3: Standard deviation of instrumental noise for the 100 GHz band. The variance of the noise is also correlated with the scanning strategy of the satellite, because blue regions have been scanned many times, and therefore have low instrumental noise.

between observations and noise variance, this RMS map also serves as a useful illustration of the Planck scanning strategy. From the estimated noise level of the detector σ, we can express the noise in each pixel as

$$\sigma_i = \frac{\sigma}{\sqrt{n_i}}, \tag{4.6}$$

where n_i is the number of times that pixel has been scanned. The number of times a certain pixel has been scanned is a time-dependent variable because Planck made several full-sky scans during its lifetime, the 1-year data has more noise than the 4-year data.

4.7 Additional systematic effects

There are a plethora of systematic effects that may arise at any point in data processing. New errors might even occur while attempting to remedy the old ones, but many are of a lesser degree of interest for this thesis. However, a couple of systematic effects that are of some concern in this thesis and deserve a short explanation. They are the far side lobe (FSL) signal, and SWB polarization sensitivity estimates. The far side lobe component is a spurious signal contaminant from outside the main beams. This effect is removed during map-making. Additionally, while SWBs are in principle only polarization sensitive to less than 10%, these may vary across detectors. Using wrong sensitivity values in accounting for the polarization sensitivity during map-making will lead to polarization contamination. This is because the leaked polarization signal is

removed using templates scaled with the sensitivity of the SWB. Subtracting the wrong amount of polarization signal will, therefore, contaminate the maps.

Chapter 5

CMB analysis - from raw data to cosmological parameters

"Don't try this at home."

— Steve-O

There exists a vast number of numerical and mathematical tools for efficiently analyzing the CMB. In this chapter, we will give an overview of the data processing from time-ordered sky signal to parameter estimation. More specifically, we will discuss how our data is handled on a sphere and ordered in map-making, followed by a discussion of the backbone of this thesis; the concept of Bayesian analysis and the Gibbs sampling algorithm. Finally, we will elaborate on the Blackwell-Rao estimator used for determining the angular power spectrum and constraining cosmological parameters from the CMB.

5.1 Spherical Harmonics and the Power Spectrum

When we look at maps produced from COBE, WMAP or Planck data, we see obvious differences in resolution and smoothness, but the physical properties of the universe are not intuitive, and impossible to deduce by eye. To disentangle these, we introduce a few useful mathematical tools, namely spherical harmonics, and the angular power spectrum. By decomposing the full sky maps of COBE, WMAP or Planck into many separate maps of waves of different scale, we can represent the initial map as a linear combination of basis functions, Y_{lm}. The l parameter is the number of waves along a meridian, and m is the number of modes along the equator. This is similar to a Fourier transform in flat space, where any function can be reproduced as a sum of different sine and cosine waves, but this time, it is expanded onto a sphere.

For example, we may recreate the CMB signal as produced by COBE by adding together smooth maps at big scales. In order to do the same for WMAP, we need increasingly smaller scales in the mix, as is visualized for the first eight multipoles of WMAP in figure 5.1. In the final linear combination, all multipoles are assigned different

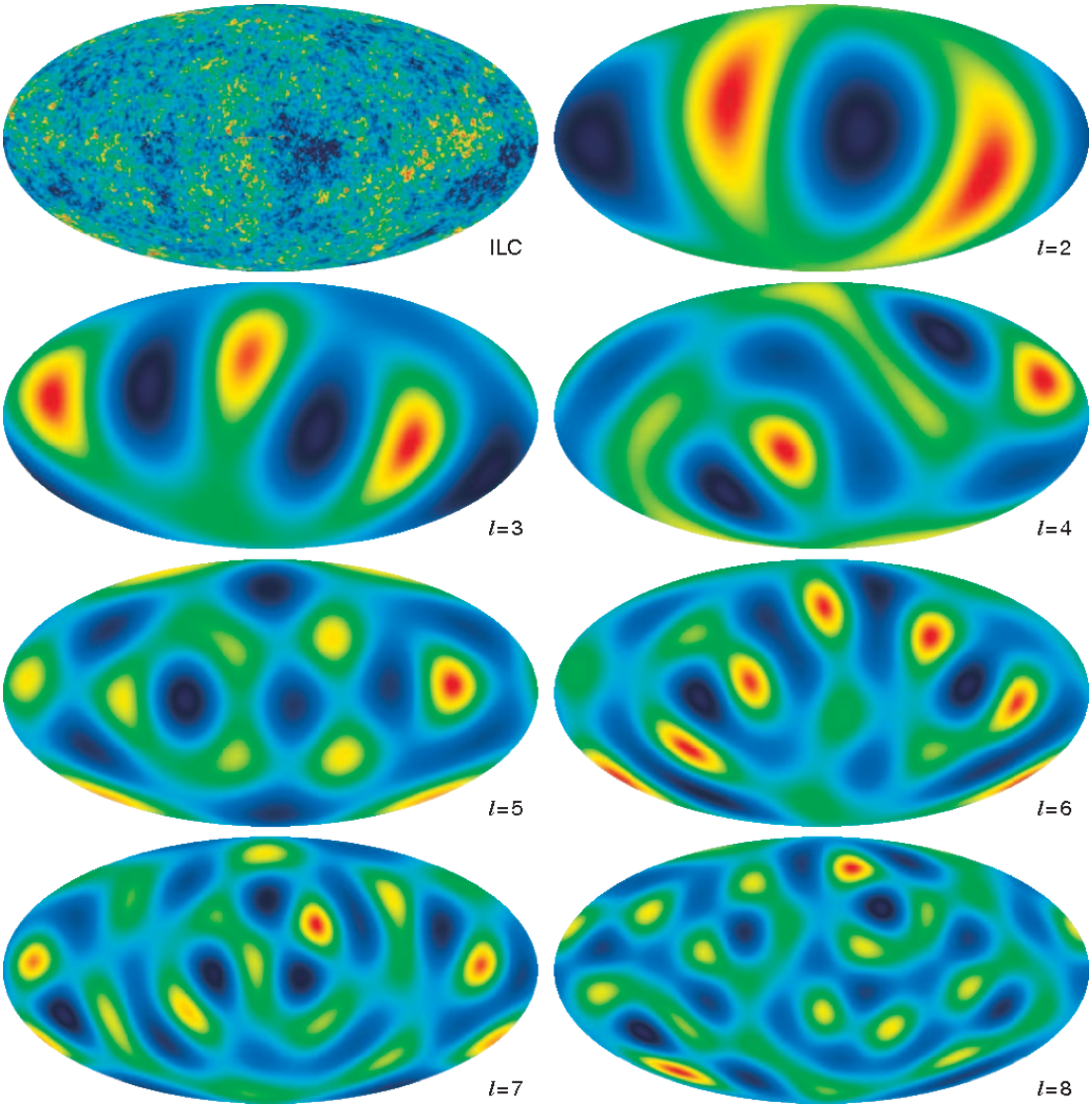


Figure 5.1: We can represent any spherical image as a sum of spherical harmonic basis functions Y_{lm}. These images represent the modes from the decomposed CMB map on the top left. This CMB map is from the ILC 3-year analysis of WMAP. The rest of the images illustrate the various spherical harmonic multipoles from $l = 2$, the "quadrupole", to multipoles as high as $l = 8$. To produce a map such as the Planck CMB map, we would need thousands of modes. Image taken from the WMAP temperature analysis of Hinshaw et al. [54]

weights, or amplitude represented mathematically as a_{lm}. The dipole, for example, requires a stronger weight than the quadrupole, and each of them is coupled to different physical phenomena. The square of these amplitudes assigned to each multipole spans out the *power spectrum*, and this power spectrum reveals crucial information on the composition of the universe and the processes that created all structures. The associated power spectrum for WMAP and the 2013 Planck results are shown in figure 5.2. We define the spherical harmonic temperature anisotropies in the CMB mathematically as

$$\frac{\Delta T}{T}(\theta, \phi) = \sum_{l} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \phi), \tag{5.1}$$

where each of the a_{lm} in the case of the CMB represent the amplitude attributed to each multipole. Each multipole is associated with physical effects, a_{0m} is, for example, the monopole and a_{1m} is the amplitude of the dipole. When computing the angular power spectrum, C_l, we sum over m, because the universe is assumed to be isotropic, and direction is therefore irrelevant. Besides, all a_{lm}'s are assumed to be Gaussian and statistically independent. We commonly represent the CMB temperature power spectrum as average over amplitudes for each multipole l;

$$C_l = \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}|^2, \tag{5.2}$$

where the factor of $2l + 1$ is the number of m's per l.

The CMB anisotropies are stochastic, which means that to minimize the uncertainty of the power spectrum we should ideally observe an ensemble of universes. Unfortunately, as far as we know, there is only one. We can think of the universe as one number drawn from a probability distribution and our models of the universe attempt to recreate this probability distribution. However, the only available data point is that single number. This uncertainty in our cosmological models is called "Cosmic variance", $\mathrm{Var}(C_l)$, and limits the accuracy of our measurements and model fit. This effect is particularly prominent on large scales (low l's) because we have fewer samples. For small scales, however, we may probe many different parts of the sky independently and deduce their statistical properties.

5.2 Pixels on a sphere

For a digital representation of the signal on a sphere, we need to choose a pixelation scheme. How we chose to represent our data have significant implications on our ability to analyze it efficiently. In recent years, the standard representation of CMB data has been the HEALPix scheme by Górski et al. [55]. The acronym stands for Hierarchical Equal Area isoLatitude Pixelation of a sphere, implying that it describes data on a sphere in terms of pixels of equal area. One refers to the resolution of such a map by N_{side}, where the number of pixels per map is given by $N_{\mathrm{pix}} = 12 \cdot N_{\mathrm{side}}^2$.

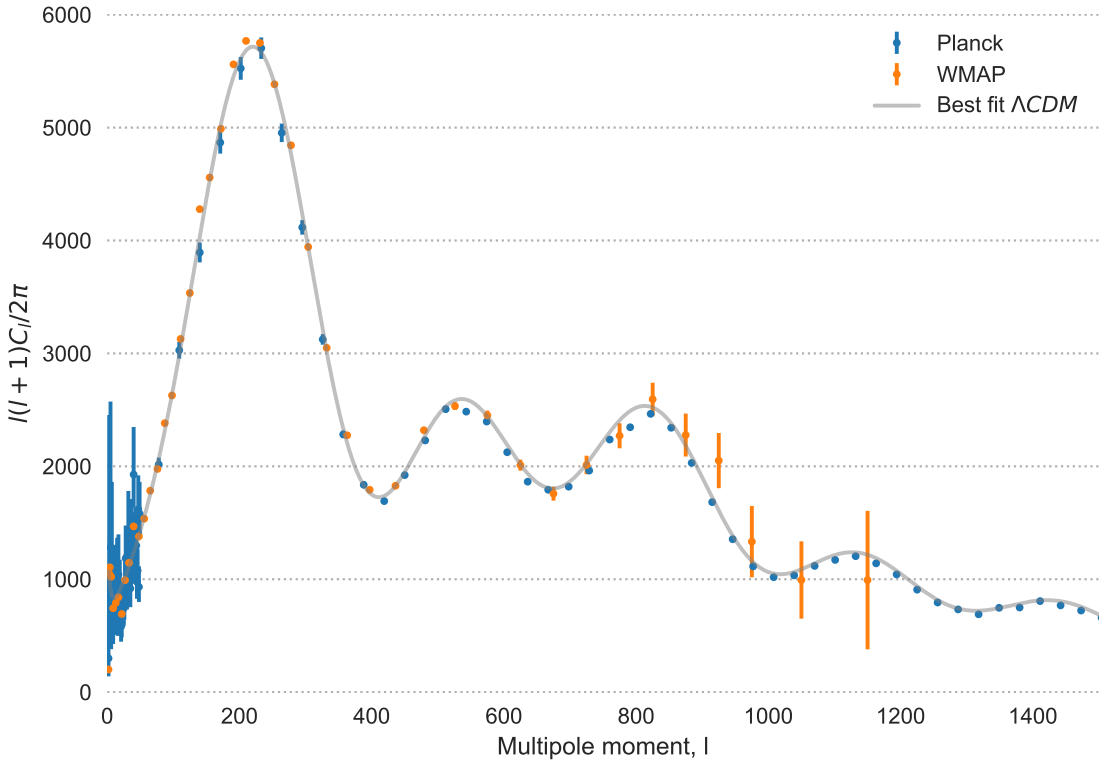


Figure 5.2: Best fit ΛCDM, WMAP and Planck power spectra. Note how WMAP was unable to probe small scales (high l's) due to lower resolution measurements. The high error at high l's is due to the low amount of possible measurement of big scales, known as cosmic variance. Data courtesy of NASA's LAMBDA service [22]

This type of pixelation scheme has a set of interesting mathematical properties which allows us to speed up the process of spherical harmonic transforms, which is one of the most significant challenges in constraining data usage in CMB analysis. For example, the fact that all pixels are placed on rings of constant latitude around the sphere allows for computationally efficient estimation of Legendre polynomials which is vital in evaluations of spherical harmonics [56]. In addition to this, the hierarchical nature of the pixelation allows for easier wavelet transformations, and the equal area of each pixel facilitates sampling of data points with regional independence.

5.3 Map-making

The data we download from the satellite is not perfectly arranged into sky-maps, but rather a stream of *Time-Ordered Data* (TOD). The data assembly process is called map-making, which is the process of arranging temperature and pointing information into a full map while - to the best of one's ability - accounting for the potential instrumental effects. For large data sets, this is computationally demanding, which requires smart

map-making algorithms. One such algorithm is called "destriping", a method utilizing the instrument scanning strategy to minimize correlated noise. When applying this method, we model our data as

$$\mathbf{d} = \mathrm{P}\mathbf{m} + \mathrm{F}\mathbf{a} + \mathbf{n}, \tag{5.3}$$

where $\mathbf{d}$ is the TOD, $\mathbf{m}$ is the sky map which we want to solve for, and P is the pointing matrix that converts our ideal map into TOD. The second term in the equation signifies the correlated noise in our model. Had we not included correlated noise, map-making would be simple, and just a matter of binning the TOD onto a full sky pixelization. Considering that the noise is correlated in time, following the scanning strategy, destriping attempts to model it by dividing it into a series of constant offset values called *baselines*. In our model, $\mathbf{a}$ is the amplitude of each baseline, and F projects them onto the TOD. A significant amount of the effort of map-making goes into modeling this correlated noise in the best way possible. The $\mathbf{n}$ factor in our model, is once again uncorrelated white noise. Solving for $\mathbf{m}$ is not trivial, and even if the correlated noise is appropriately modeled, other errors may not be evident until after component separation, when the signal has been decomposed into separate components. While what remains after ideal component separation is theoretically uncorrelated noise, in practice, it leaves behind the breadcrumbs of instrumental errors not correctly modeled during map-making and more often than not, traces of mismodeled foregrounds.

When constructing full-sky maps, it is common to combine the different detector data from each instrument to minimize systematic noise. The maps are in practice being averaged over, lowering the Gaussian noise. When these sets of maps show signs of errors, such as a bandpass mismatch morphology or polarization leakage, we can separate them again and identify the offending detector-map. Ideally, one would include all detector maps separately, in order to maximize the data volume. In this thesis, we analyze sky maps from three different map-making pipelines. The first is DX11, applied to the 2015 Planck release [57, 47]. Since then, the official Planck Data Processing Center (DPC), has developed DX12, using the SROLL code [58], which we will also analyze. However, the primary focus of this thesis will be the analysis of the NPIPE maps, constructed using the NERSC pipeline at Lawrence Berkeley National Laboratory by Reijo Keskitalo [59]. The specific differences between these will be discussed in the results part.

5.4 Posterior distribution sampling

As new CMB experiments gather increasingly precise data, the analysis tools are forced to keep pace. During the years of COBE, removal of foregrounds was carried out using brute force methods of posterior distribution mapping, which were viable due to the low resolution of the data. While the COBE-DMR mission observed with an angular resolution of 7°, corresponding to 6144 pixels per map, each Planck HFI map consists of $50 \cdot 10^6$ pixels. This immense increase in resolution demands new, efficient, and comprehensive foreground removal techniques. Therefore, a plethora of algorithms have

emerged throughout the years to solve these problems. One prominent recent example is the Gibbs sampling algorithm, a Markov chain Monte Carlo (MCMC) algorithm for Bayesian analysis proposed to the CMB community by Jewell et al. [60] and Wandelt et al. [61].

5.4.1 The problem

Before delving into the discussion of the statistical tools, we define the problem. First, we reformulate the data model from the last chapters without foregrounds and instrumental systematics, simply as

$$\mathbf{d} = \mathbf{s} + \mathbf{n}, \tag{5.4}$$

where $\mathbf{d}$ is the observed data, $\mathbf{s}$ is the CMB signal and $\mathbf{n}$ is the Gaussian noise. Furthermore, we assume both $\mathbf{s}$ and $\mathbf{n}$ to be Gaussian random fields with covariance matrices $\mathbf{S}$ and $\mathbf{N}$, respectively. In harmonic space, as we have seen before, the signal can be expressed as $\mathbf{s} = \sum_{l,m} a_{lm} Y_{lm}$, and the power spectrum is a property of this signal related by $C_{lm,l'm'} = \langle a^*_{lm}, a_{l'm'} \rangle = C_l \delta_{ll'} \delta_{mm''}$. Following a measurement of our signal, we are interested in knowing the probability of a set of parameters, θ, given those data, $\mathbf{d}$. This probability is known as the *posterior distribution*, and can be expressed through Bayes' theorem, the foundation of Bayesian analysis,

$$P(\theta|\mathbf{d}) = \frac{P(\mathbf{d}|\theta)\, P(\theta)}{P(\mathbf{d})} \propto \mathcal{L}(\theta) P(\theta). \tag{5.5}$$

This equation relates the *likelihood* $\mathcal{L}(\theta) = P(\mathbf{d}|\theta)$, which is the probability of the data given the parameters, and prior $P(\theta)$, to our posterior. The prior is any knowledge we already have about the distribution of θ. The expression also includes a constant normalization factor $P(\mathbf{d})$, which we disregard for simplicity. If we apply this theorem to the problem of CMB data and the power spectrum, we obtain

$$P(\mathbf{s}, C_l|\mathbf{d}) \propto P(\mathbf{d}|\mathbf{s}, C_l) P(\mathbf{s}, C_l) = P(\mathbf{d}|\mathbf{s}, C_l) P(\mathbf{s}|C_l) P(C_l). \tag{5.6}$$

This relation, assuming both the signal and noise to be Gaussian, results in the joint posterior distribution of the two parameters,

$$P(\mathbf{s}, C_l|\mathbf{d}) \propto e^{-\frac{1}{2}[\mathbf{d}-\mathbf{s}]^t \mathbf{N}^{-1} [\mathbf{d}-\mathbf{s}]} \prod_l \frac{e^{-\frac{2l+1}{2}\frac{\sigma_l}{C_l}}}{C_l^{\frac{2l+1}{2}}} P(C_l), \tag{5.7}$$

where the first exponential part of the expression corresponds to $P(\mathbf{d}|\mathbf{s}, C_l)$, and simply measures the goodness-of-fit between model and data; $\mathbf{n} = \mathbf{d} - \mathbf{s}$. The last part is simplified through $P(C_l|\mathbf{s}, \mathbf{d}) = P(C_l|\mathbf{s})$, because we want to extract the power spectrum, and therefore only need to look at the signal $\mathbf{s}$, and not the full data including the noise $\mathbf{d}$. In order to map out this full posterior, we require an efficient sampling algorithm.

5.4.2 The Metropolis-Hastings algorithm

The Metropolis-Hastings (MH) algorithm is a form of MCMC algorithm and works by randomly sampling from distribution proportional to the posterior when direct sampling is not an option. In a nutshell, it works by proposing a sample from a distribution, that is either accepted or rejected to form the posterior, based on an acceptance ratio. This process is then repeated until convergence. Mathematically, we describe this acceptance ratio as

$$a = \frac{P(\theta^{i+1})}{P(\theta^i)} \frac{T(\theta^i|\theta^{i+1})}{T(\theta^{i+1}|\theta^i)}, \tag{5.8}$$

where $P(\theta^{i+1})$ and $P(\theta^i)$ are the probability densities of the new and old parameters. Next, $T(\theta^i|\theta^{i+1})$ is the proposal distribution, which give the probability of proposing θ^{i+1} given θ^i. We initialize the algorithm by setting an initial condition x_0, starting at $i = 0$, before iterating using the accept-reject method:

1. Generate a candidate sample from $T(\theta^i|\theta^{i+1})$
2. Calculate the acceptence ratio a.
3. Generate a random number between 0 and 1, and test it against the acceptence probability $\mathrm{A} = \min\{1, a\}$. If it is accepted then the new parameter becomes θ_{i+1}, if it is rejected, then $\theta_{i+1} = \theta_i$.

After a sufficient number of samples, this algorithm will converge towards $P(\theta)$.

When we deal with the CMB, we want to sample many different parameters. This renders the sampling problem multidimensional. However, in applying the MH algorithm, we must be able to evaluate the complicated, full posterior distribution of these parameters. To make this process more tractable, we propose an alternative, where we break the multidimensional problem into separate well-known distributions. Specifically, we want to propose for conditionals of the full posterior. In doing this, we may rewrite the proposal distribution for a problem with two parameters as

$$T(\mathrm{A}^{i+1}, \mathrm{B}^{i+1}|\mathrm{A}^i, \mathrm{B}^i) = \delta(\mathrm{B}^{i+1} - \mathrm{B}^i) \cdot P(\mathrm{A}^{i+1}|\mathrm{B}^i), \tag{5.9}$$

where A and B are parameters from different distributions. This method now expresses the proposal distribution by keeping one component constant, while sampling the other from the exact conditional distribution. This again allows us to rewrite the acceptance ratio as

$$a = \frac{P(\mathrm{A}^{i+1}|\mathrm{B}^{i+1})}{P(\mathrm{A}^i|\mathrm{B}^i)} \cdot \frac{\delta(\mathrm{B}^i - \mathrm{B}^{i+1})P(\mathrm{A}^i|\mathrm{B}^{i+1})}{\delta(\mathrm{B}^{i+1} - \mathrm{B}^i)P(\mathrm{A}^{i+1}|\mathrm{B}^i)} \tag{5.10}$$

$$= \frac{P(\mathrm{A}^{i+1}|\mathrm{B}^i)P(\mathrm{B}^i)}{P(\mathrm{A}^i|\mathrm{B}^i)P(\mathrm{B}^i)} \cdot \frac{P(\mathrm{A}^i|\mathrm{B}^i)}{P(\mathrm{A}^{i+1}|\mathrm{B}^i)} \tag{5.11}$$

$$= 1. \tag{5.12}$$

Here we have used the delta function to set $\mathrm{B}^{i+1} = \mathrm{B}^i$, before applying Bayes theorem to the first fraction. In practice, this algorithm accepts all proposals and requires only prior knowledge about its conditional distributions, which immensely simplifies our problem. This type of sampling is known as Gibbs sampling.

5.4.3 Gibbs sampling

The application of the Gibbs sampling scheme allows us to rewrite our CMB analysis problem into separate conditional distributions $P(\mathbf{s}|C_l, \mathbf{d})$ and $P(C_l|\mathbf{s}, \mathbf{d})$. These conditional distributions are often much simpler, and often have analytic distributions. Once the conditional distributions are known, the process works by keeping either of the parameters fixed, and iteratively sampling the conditionals. A mathematical representation of the iterative sampling scheme is thus represented as

$$\mathbf{s}^{i+1} \leftarrow P(\mathbf{s}|C_l^i, \mathbf{d}) \tag{5.13}$$

$$C_l^{i+1} \leftarrow P(C_l|s^{i+1}, \mathbf{d}), \tag{5.14}$$

where the arrow indicates sampling from the distribution on the right hand side. An additional reason to favor the Gibbs sampler is the fact that no samples are rejected, as opposed to general MCMC methods, where each sample must pass a certain threshold to be accepted. In principle, all Gibbs samples are accepted, apart from a small portion removed during the burn-in phase.

5.5 The `Commander` Gibbs sampling code

In this thesis, we will use the Gibbs sampling algorithm through the `Commander` code; a joint foreground-CMB Gibbs sampler code by Eriksen et. al. [1]. In order to explain the inner workings of this code, let us consider the signal model

$$\mathbf{s}_\nu(\theta) = \mathbf{s}_\nu(\mathbf{a}_i, \beta_i\, g_\nu, \Delta_\nu, \mathbf{m}_\nu), \tag{5.15}$$

$$\mathbf{s}_\nu = g_\nu \sum_{i=1}^{N} F_\nu^i(\beta_i, \Delta_\nu)\mathbf{a}_i, \tag{5.16}$$

where, once again, $\mathbf{a}_i$ is the amplitude vector for component i, where β_i is the associated set of spectral parameters for each component. The g_ν is the multiplicative gain factor per frequency band ν, Δ_ν is the bandpass shift, and $\mathbf{m}_\nu$ is the monopole offset per frequency band. Lastly, F_ν^i is the projection operator which translates the amplitude of each component onto the basis of the observed data, accounting for their spectral effects and frequency dependence.

With our model in place, we explore the relevant `Commander1` Gibbs sampling scheme, as applied in this thesis. Looking at our data model, we want to sample the joint CMB-foreground posterior $P(\mathbf{s}, C_l, \mathbf{a}_i, g_\nu, \Delta_\nu, \mathbf{m}_\nu|\mathbf{d})$. By sorting the parameters into conditional sampling distributions, we may solve the system iteratively as

$$\{\mathbf{s}, \mathbf{a}_i, g_\nu, \Delta_\nu, \mathbf{m}_\nu\}^{i+1} \leftarrow P(\mathbf{s}, \mathbf{a}_i, g_\nu, \Delta_\nu, \mathbf{m}_\nu|C_l^i, \beta_i^i, \mathbf{d}), \tag{5.17}$$

$$\beta_i^{i+1} \leftarrow P(\beta_i|\mathbf{s}^{i+1}, \mathbf{a}_i^{i+1}, g_\nu^{i+1}, \Delta_\nu^{i+1}, \mathbf{m}_\nu^{i+1}\mathbf{d}), \tag{5.18}$$

$$C_l^{i+1} \leftarrow P(C_l|\mathbf{s}^{i+1}). \tag{5.19}$$

Sampling from each conditional requires different approaches, by far the most computationally demanding of which, is equation 5.17. In order to fit the linear amplitudes, and sample from the conditional distribution $P(\mathbf{a}_i|\mathbf{d}...)$ we employ the conjugate gradient (CG) solver for the following equation,

$$(\mathbf{S}^{-1} + \mathbf{P}^T\mathbf{N}^{-1}\mathbf{P})\mathbf{a} = \mathbf{P}^T\mathbf{N}^{-1}\mathbf{d} + \mathbf{P}^T\mathbf{N}^{-1/2}\omega_1 + \mathbf{S}^{-1}\omega_2. \tag{5.20}$$

We solve for the amplitudes $\mathbf{a}$, where $\mathbf{P}$ projects amplitude to data, $\mathbf{S}$ and $\mathbf{N}$ are the signal prior and noise covariance matrices respectively. The ω_i parameters are random Gaussian vectors with zero mean and unit variance. While we will not explore the details of the CG method, we recommend the thorough and entertaining review given by Shewchuk (1994) [62] for any interested reader. All non-linear parameters, such as spectral-parameters in equation 5.17, are sampled via an inverse sampler. The sampling step for spectral parameters, employs the distribution

$$P(\beta_i|\mathbf{s}^{i+1}, \mathbf{a}_i^{i+1}, g_\nu^{i+1}, \Delta_\nu^{i+1}, \mathbf{m}_\nu^{i+1}\mathbf{d}) = e^{-\frac{1}{2}(d-A\nu^\beta)\mathbf{N}^{-1}(d-A\nu^\beta)}, \tag{5.21}$$

where $A\nu^\beta$ is a simplification of the signal model, where A is the previously solved amplitudes and β are the non-linear spectral parameters. Solving this system is done by mapping out the corresponding cumulative distribution, $F(\beta)$, and drawing a random value u such that $F(\beta) = u$, and hence $\beta = F^{-1}(u)$.

Lastly, we sample the C_l, by assuming the sky to be Gaussian and isotropic, which reduces the conditional distribution to a well-defined inverse gamma distribution.

5.5.1 Limitations of `Commander1`

In this thesis, we will be dealing with maps of $N_{\text{side}} = 256$, and 64, which are downgraded from the full resolution maps of $N_{\text{side}} = 2048$ and 1024. Additionally, the version of `Commander` applied to this thesis is not compatible with multiple resolution input maps across frequencies, a major drawback which scales down the output maps to the lowest resolution input map. A newer version, `Commander2`, corrects for this, by sampling in spherical harmonic space, as opposed to pixel space. However, `Commander2` lacks some of the functionality of `Commander1`, we therefore employ the latter in this thesis.

Of course, as with all things, the Gibbs sampling is not perfect. It struggles with nearly degenerate parameters, which slows down the convergence time significantly. As previously discussed, we often encounter such degeneracies, an example being the similar effect of gain and amplitude sampling of high-frequency bands in Planck. This is not a significant problem in our analysis, as we apply low-resolution data. For a high-resolution analysis, such degeneracies would impair our computation speed catastrophically.

5.6 Parameter estimation with the Blackwell-Rao approximation

After producing a full realization of the CMB sky through the Gibbs sampling scheme, we are tasked with estimating the cosmological parameters encoded within it. This

is partly done through the sampling of the power spectrum in `Commander`. However, due to high risk of foreground contamination along the galactic plane, we are forced to disregard a large portion of our data. When we sample the power spectrum from the CMB using `Commander`, we apply a mask over the high-risk foreground contaminated area and fill in the missing signal by sampling from a prior distribution, as illustrated in figure 5.3. One such realization of the sky has an associated power spectrum

$$\sigma_l^i = \frac{1}{2l+1} \sum_{m=-l}^{l} |s_{lm}|^2. \tag{5.22}$$

However, in order to account for uncertainties from the mask and noise, we wish to average over an ensemble of sky realizations. This minimizes the variance of our final power spectrum and hence the associated cosmological parameters. Therefore, we generate many such sky realizations σ_l^i using `Commander`, by filling another set of values in the masked area. Finally, in order to obtain the full density, we average over the distributions for all sky realizations

$$P(C_l|\mathbf{d}) \approx \langle P(C_l|\sigma_l^i)\rangle. \tag{5.23}$$

This method of computing the likelihood of the power spectrum is called the Blackwell-Rao approximation. The process is illustrated in figure 5.4, which shows the probability distribution of each power spectrum, given that of one sky realization.

By constraining the power spectrum $C_l(\theta)$, we also constrain the cosmological parameters θ encoded within it. These parameters can thus be estimated by applying a simple MCMC method, with sufficient knowledge of $P(C_l|\mathbf{d})$. The main advantages of employing the BR estimator as opposed to applying a brute-force evaluation of the likelihood is the better scaling with $l_{\max}$.

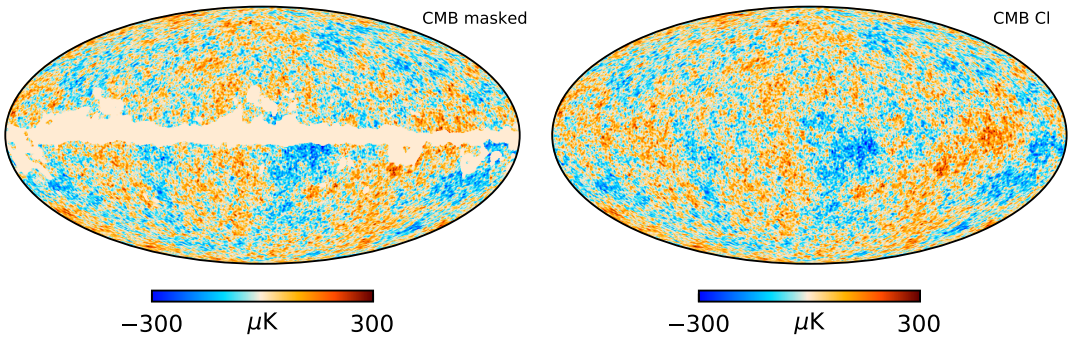


Figure 5.3: Left: masked CMB maximum posterior. Right: one CMB sky realization, filled in signal in masked region from random sample of a prior.

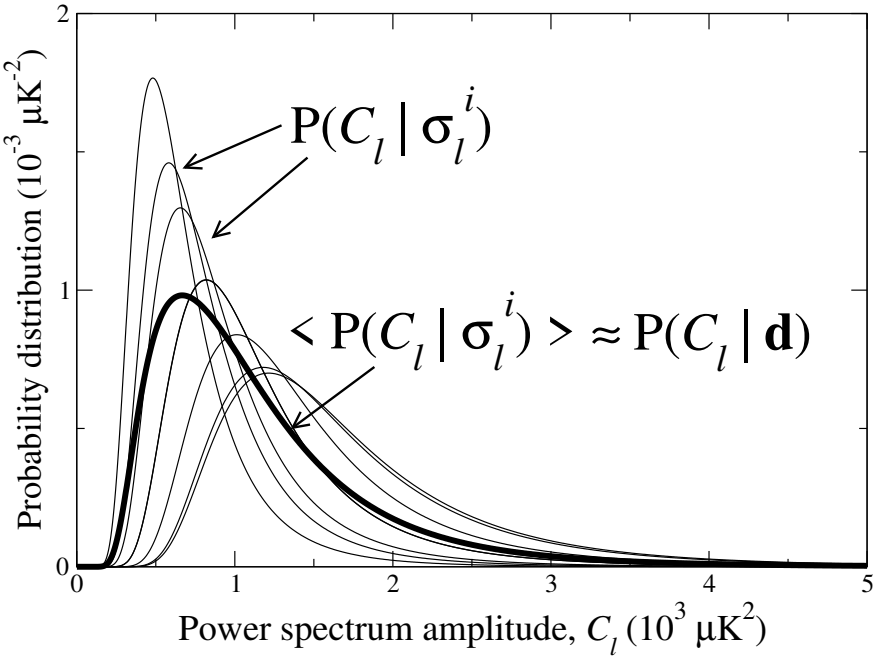


Figure 5.4: Illustration of the the BR approximation. Thin lines show the probability of a power spectrum amplitude C_l, given the sky realization specific power spectrum σ_l^i. The thick line shows the average of the ensamble, and the true density $P(C_l|\mathbf{d})$ Figure take from Chu et al. [63]

Part III

Results

Chapter 6

Old state-of-the-art sky maps

"We've got a situation."

— Mike "the Situation" Sorrentino

We split the results part of this thesis into four chapters; "Old state of the art sky maps", "Evolution of NPIPE sky maps", "New state of the art sky maps", and "Conclusion and future prospects". The first chapter describes the pedagogical exercise of applying component separation on a well-known set of sky maps. The first map set employed is the latest set produced by the Planck data processing center (DPC). This sky map set consists of nine full frequency maps, on which we fit a simplified sky model. Subsequently, we apply an identical component separation model to the previous generation of sky maps, presented in the official 2015 Planck release, and discuss the changes that have been made since. Our overarching goal is to achieve a good model while using as many data points as possible. Moreover, we wish that the component maps describe a physical sky and that residual maps exhibit as little contamination as possible. However, our ability to include parameters is reliant on the number and frequency coverage of the available maps, which greatly differs from set to set.

The most significant challenge in component separation at the moment lies in the interplay of instrumental systematics and foreground contamination. The 2015 Planck release presented detector maps exhibiting substantial systematic effects, which motivates the central part of this thesis, chapter 7. In chapter 7, we present a new set of sky maps, produced by Reijo Keskitalo at Berkeley Lab using my results as inputs, and the results of applying component separation on these. The chapter aims to resolve some of the issues of the 2015 release by identifying problems and subsequently adjusting our component separation model in conjunction with the creation of new and improved sky maps from Keskitalo. Finally, in the two last chapters, we present our best-fit sky model and associated component maps, and a discussion of the future of component separation.

6.1 DPC map-making - DX11 and DX12

The Planck DPC sky maps represent the current public state-of-the-art. In this thesis, we discuss two iterations of these maps. First, we will look at DX12, which is the latest iteration, and the maps to be presented in the 2018 Planck release. Second, we look at DX11, which are the maps applied to the analysis of the 2015 Planck release [46]. Both of these sets rely on the same general idea of map-making, utilizing the destriping method explained in section 5.3. However, in 2016, the Planck collaboration was introduced to the `SRoll` algorithm, which is similar to the one used on DX11, but simultaneously solves for common errors such as calibration and leakage of temperature to polarization [58]. This method of fitting everything at once, through a "steam-roll" approach (hence the name), has its advantages and disadvantages.

While the advantage is that one uses all of the data, a possible disadvantage is that the templates used for each component are degenerate. This introduces the possibility that the additional templates remove true signal and not only systematics.

6.2 Component separation for DX12

The first set of sky maps we analyze are the nine full frequency maps for DX12. As we peel off layer by layer of foreground components, modeled by our algorithm using our chosen priors, we see less and less signal in the residual maps. An ideal result from each run of `Commander` would be a complete white noise residual with each component perfectly separated into full-sky maps. However, the imperfections of our model reveal residual signal in the range of $\pm 10\mu$K for each frequency map illustrated in figure 6.3. The squared residual summed across all bands is presented as the χ^2 map, expressed mathematically as

$$\chi^2(p) = \sum_\nu \left(\frac{\mathbf{d}_\nu(p) - \mathbf{s}_\nu(p)}{\sigma_\nu(p)} \right)^2 . \tag{6.1}$$

Here, $\mathbf{d}_\nu(p)$ is the data signal for each frequency band at a given pixel, $\mathbf{s}_\nu(p)$ is the signal model per frequency at a given pixel including all foregrounds, and $\sigma_\nu(p)$ is the standard deviation of the noise per pixel.

The χ^2 map is an internally consistent check of the goodness-of-fit of our model. While we may fit a model that is utterly ridiculous, we may still achieve low values of χ^2, but the component maps would not be physical. We can constrain our model by adding more frequency bands, which will always increase the χ^2, or we may introduce more parameters, which will also make the model more flexible. As long as we compare models with the same number of free parameters and a similar number of maps, the χ^2 serves as a suitable method of comparison. Furthermore, the χ^2-map serves as an excellent reference for identifying errors as we adjust our model. In the following chapters, we refer to the χ^2-map when evaluating the progress of our sky model, while at the same time, making sure our component maps describe a physically meaningful system. Furthermore, the high signal-to-noise region along the galactic plane is hard to

model. Hence we will present a separate summary χ^2 values for high latitudes and the full sky.

6.2.1 Component maps

All maps presented in this thesis are smoothed to a common resolution of $N_{\text{side}} = 64$ and a beam of $40'$ FWHM. The DX12 maximum posterior component maps will be presented in the following section, generated using the following sky model, with priors listed in table 6.2.1

- We fit a joint low-frequency component consisting of spinning dust, synchrotron, and free-free emission, approximated by a power law with parameter β_s.
- We fit the thermal dust blackbody parameter β_d while freezing the temperature T_d
- We fit a single CO-multiline component consisting of all three lines visible at 115, 230 and 345 GHz with fixed line ratios at 0.52672 and 0.16997 respectively.
- No additional mono- or dipole corrections are done after map-making.
- We do not fit any gain values or bandpass shifts.
- All-sky maps are full frequency per band.

Component	Prior
CMB	$\mathrm{T}_{\mathrm{CMB}} = 2.755\mathrm{K}$
Low-frequency	$\beta_s = -3.1 \pm 0.5$
Thermal dust	$\beta_d = 1.55 \pm 0.1$
	$T_d = 23\mathrm{K}$
CO	Spatially uniform line ratios

Table 6.1: The DX12 sky model priors.

With these factors in mind, we present the component maps produced by `Commander` for this set of sky maps. The maximum posterior CMB map in figure 6.1 shows anisotropies on the scale of $\pm 300\mu$K with a prominent, red line along the galactic plane displaying the difficulty of foreground removal in the high signal-to-noise regime. The associated χ^2 goodness-of-fit map shows signs of mismodeled free-free foregrounds, which is not surprising considering we fit a joint low-frequency component. In addition to the free-free morphology, the galactic band is thickened, suggesting that other foregrounds such as thermal dust are also contributing to the high χ^2 values.

Figure 6.2 shows the complete set of fitted foreground components of our model. Because each pixel is fitted with a separate model per foreground component, we also

present the model parameter values per pixel. Firstly, we see the amplitude of the low-frequency component displaying the unmistakable morphology of free-free with a thickening of the galactic center as an effect of synchrotron emission, with spinning dust stretching out from the galactic band. Next, we see the corresponding low-frequency model β parameter per frequency, followed by the amplitude of the CO emission. On the bottom, we see the three maps related to the thermal dust component, with the amplitude A_{d}, followed by the modified black-body parameters, β_d and T_d.

For each `Commander` run we also compute a set of residual maps $\mathbf{d}_\nu - \mathbf{s}_\nu$, corresponding to each of the frequency maps used in the analysis. These residual maps are shown in figure 6.3. The residual maps are paramount in identifying mismodeled foreground components. With the ability to identify contamination at each frequency band, we can decompose the morphology seen in the χ^2 map and adjust our model for the next run.

The associated residual maps of the DX12 analysis indicate that the thickening of the galactic band in the χ^2 is likely attributed to dust contamination from the 545 GHz channel. The LFI channels exhibit free-free morphology, which stretches all the way to the 217 GHz band. This is an effect of inadequate low-frequency β_s fit, which model extends up to 217 GHz. Moreover, this set of residual maps exhibit a range of monopole effects, visible as the reddening of the 143 GHz band, and the blue residual in the 217 GHz map. We also note the red morphology at high latitudes on the 857 GHz map, which is most likely due to a bad fit of the thermal dust β_d parameter.

From looking at the residual maps in figure 6.3, we deduce that this model does not clean the residual maps sufficiently to reveal any present systematic effects. Because we are applying full-frequency maps, systematic effects are less prominent, which in principle is a good thing, but comes at the cost of our ability to model the foregrounds, as it leaves fewer data points for us to fit our model against. We, therefore, stress the fact that our overarching goal is to include as many data points as possible while maintaining a low χ^2 with components describing a physical sky. We conclude that this first component separation analysis has yielded sub-optimal results. Ideally, we would continue the model fit until we were able to identify instrumental errors, or at the very least achieved a better foreground fit. Moreover, the goal of the comparison was to construct a set of maps to serve as a comparison for the upcoming analysis of NPIPE. The current state of these residual maps diminishes our ability to compare these sky maps set and draw definite conclusions on their relative improvement. In this thesis, we prioritize the analysis of the NPIPE maps, therefore, we were unable to apply the same precision analysis to the DX12 sky maps. However, the attempt served as an enormously useful training-ground for component separation, and the application of the `Commander` code.

6.3 Component separation for DX11

As a comparison to the DX12 sky maps, we have included an analysis of the previous iteration of Planck sky maps as presented in the Planck 2015 data release, henceforth referred to as DX11. In order to conveniently compare the two sets, we have implemen-

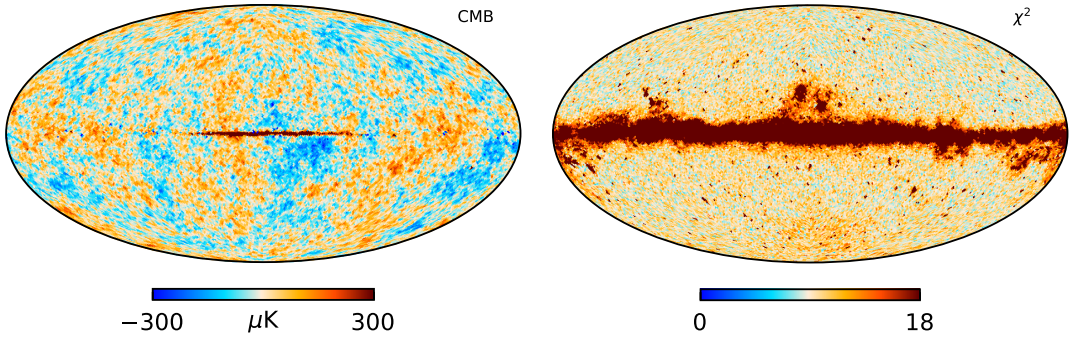


Figure 6.1: DX12 CMB and χ^2 map produced through component separation with `Commander`. First (top left) we see the CMB maximum posterior with anisotropies on the scale of $\pm 300\mu$K with a thin red line of unmodeled foreground signal along the galactic plane. Next, (top right) we see our goodness-of-fit χ^2 value for each pixel, showing clear free-free morphology and a thick galactic band.

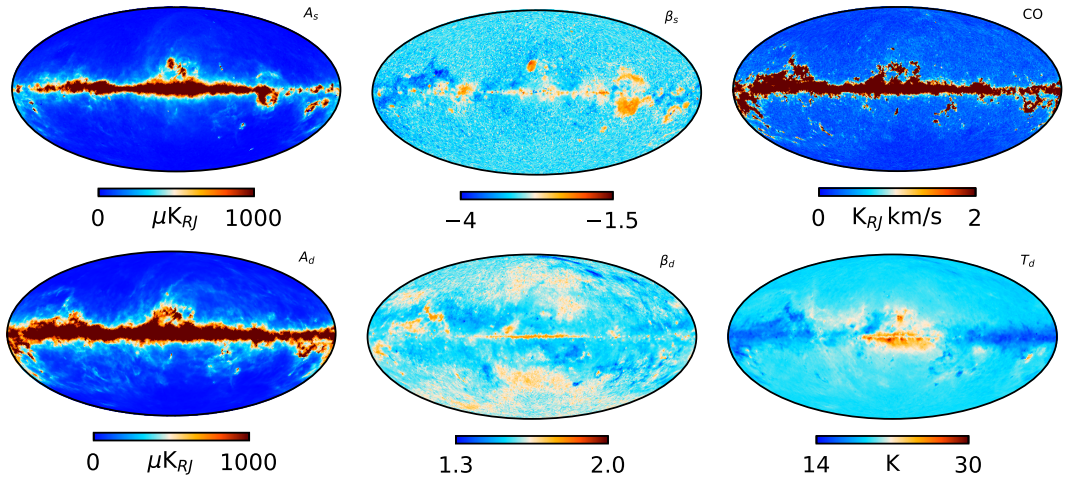


Figure 6.2: Component maps for DX12 produced through component separation with `Commander`. On the top, from left to right we first see the amplitude of the low-frequency component, displaying clear free-free morphology as well as a thickened galactic center as a result of synchrotron emission. Next, we have the β parameter of our low-frequency model for each pixel, followed by the amplitude of the CO signal. On the bottom, we see the parameters attributed to the thermal dust model. First, we have the amplitude, which clearly displays the morphology of dust. Followed by the β and Temperature component of the thermal dust modified black-body model respectively.

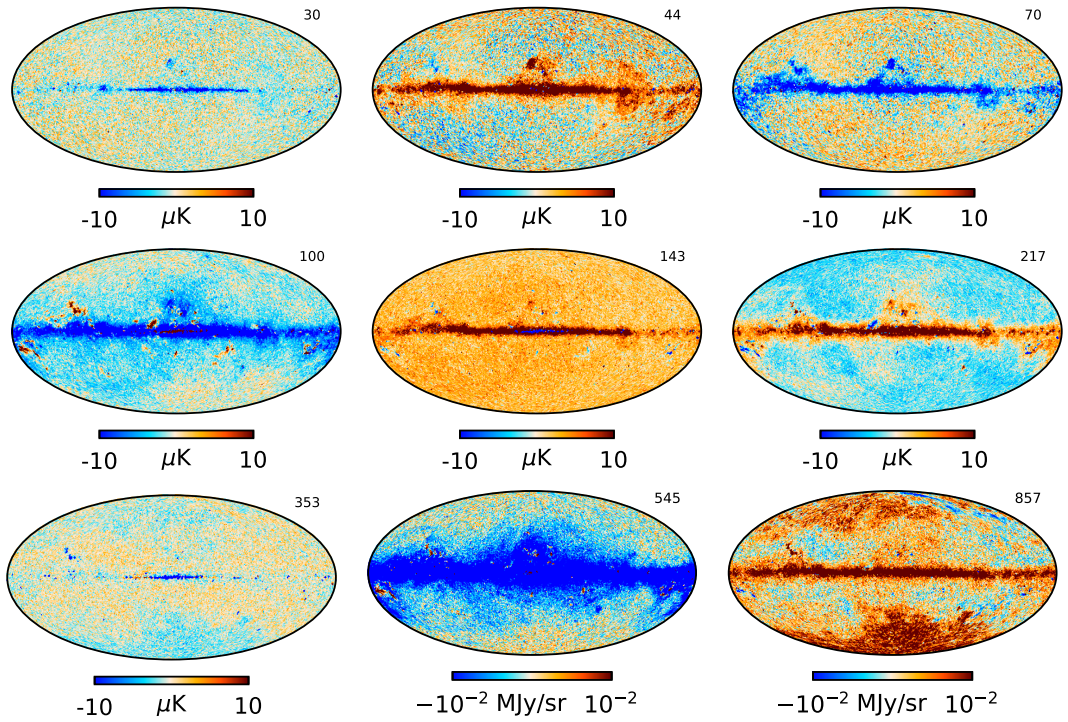


Figure 6.3: DX12 residual maps for each frequency channel used in the `Commander` analysis. Notable features are the free-free morphology visible from 44 GHz all the way to 217 GHz, as well as a monopole causing a reddening of the 143 GHz map and blue-shift in the 217 GHz. Additionally, there is clear dust contamination in the 545 GHz residual and prominent red CIB features on high latitudes of the 857 GHz map, most likely a result a bad thermal dust β_d parameter fit.

ted the DX11 sky maps with the same bandpass profiles, RMS-maps, and calibration masks as those used in the analysis of DX12. Additionally, our sky model is identical to that of DX12, using the same sky-components with the same priors, and the same number of frequency bands. However, this set of sky maps include the galactic dipole. In order to retain a consistent dipole across map sets, we removed this by calibrating against a masked difference map with DX12. This method ensures robust comparison as both monopole and dipole terms are accounted for.

6.3.1 Component maps

The DX11 analysis is strictly intended for comparison purpose, we therefore present the component maps as the difference to their DX12 equivalent, illustrated in figure 6.6. However, the χ^2 and CMB maps are included figure 6.4 as they were produced by `Commander`. The most intuitive in terms of difference to DX12 is the comparison of the CMB maps in the top left in figure 6.6. While it shows no significant changes on

high latitudes, it exhibits a thick red galactic band, signaling a discrepancy in terms of thermal dust fit. We note that the CMB map still exhibits a thin red line spanning the galactic center. Its presence, however, is due to the difficulty of model fitting in the high signal-to-noise domain. While there appears to be little change compared to the CMB of DX12, the difference map reveals major changes along the galactic plane. Our χ^2 map now also shows a bad fit of CIB, evident from the characteristic red areas at high latitudes. This is either an effect of monopole-errors or a bad thermal dust β parameter. The apparent asymmetry of the signal between the North and the South hemisphere will be discussed in a later section. Table 6.3.1 shows the average χ^2 values for the full map, as well the average for high latitudes. This shows that the model applied in these analyses are a better match for the DX11 map set across the full sky, but slightly better at high latitudes for DX12. We bear in mind that this sky model does not sample the thermal dust temperature component. Therefore, the difference is not illustrated, as it is simply zero.

Map set	χ^2_{full}	χ^2_{hilat}
DX12	18.508	1.128
DX11	16.145	1.160

Table 6.2: χ^2 values for the full sky and for high latitudes for the DX12 and the DX11 analysis.

As a result of applying the same sky model, gain values, RMS map, and bandpass, the visible differences can only be attributed to map-making. Looking at the component difference maps, the most significant difference lies along the galactic plane. Most notable in the residual maps this time, is the strong red high latitudes in the 545 and 857 GHz channel, along with the monopole features discussed previously. From the χ^2 values presented above, and the component maps below, it is hard to say with confidence if the DX12 are, in fact, worse than those of DX11.

From the Planck 2015 release, we know that these sky maps exhibit significant systematic effects [46]. Our inability to properly fit the foregrounds, and the lack of single detector maps, makes these systematic effects difficult to deduce from our results. Therefore, we have included a set of three residual maps presented in the Planck 2015 joint component separation foreground article, which illustrate the problems faced in the Planck DPC maps. These maps are shown in figure 6.8, and underlines the motivation for this thesis; supply a set of maps without systematic effects by applying component separation and map-making iteratively.

As a final comparison between the sets, we investigate the large-scale differences between DX11 and DX12 by computing the angular power spectrum using the Blackwell-Rao estimator. Figure 6.5 illustrates the small differences on high multipoles between the DX12 and DX11 analysis. Furthermore, because the processing is identical between the two, this is a direct comparison between the effects of map-making strategies on large scales. We conclude that for large scales all three maps are statistically equivalent.

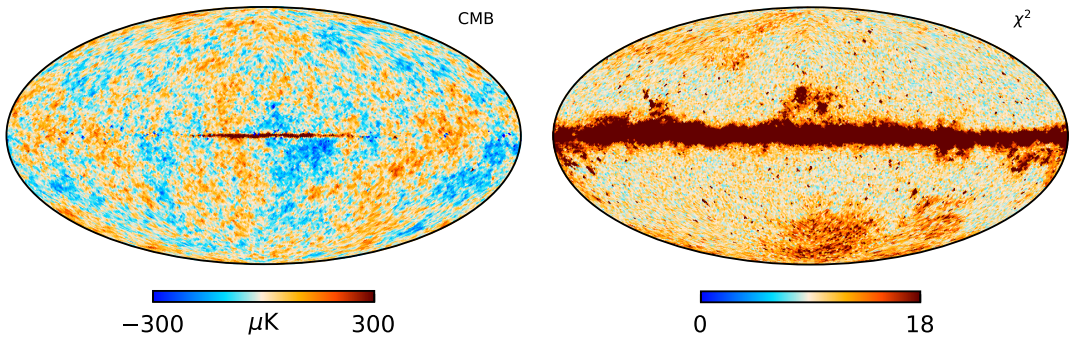


Figure 6.4: DX11 CMB and χ^2 maps, with a visible red line still present in the galactic center of the CMB. The χ^2 now exhibits higher values at high latitudes

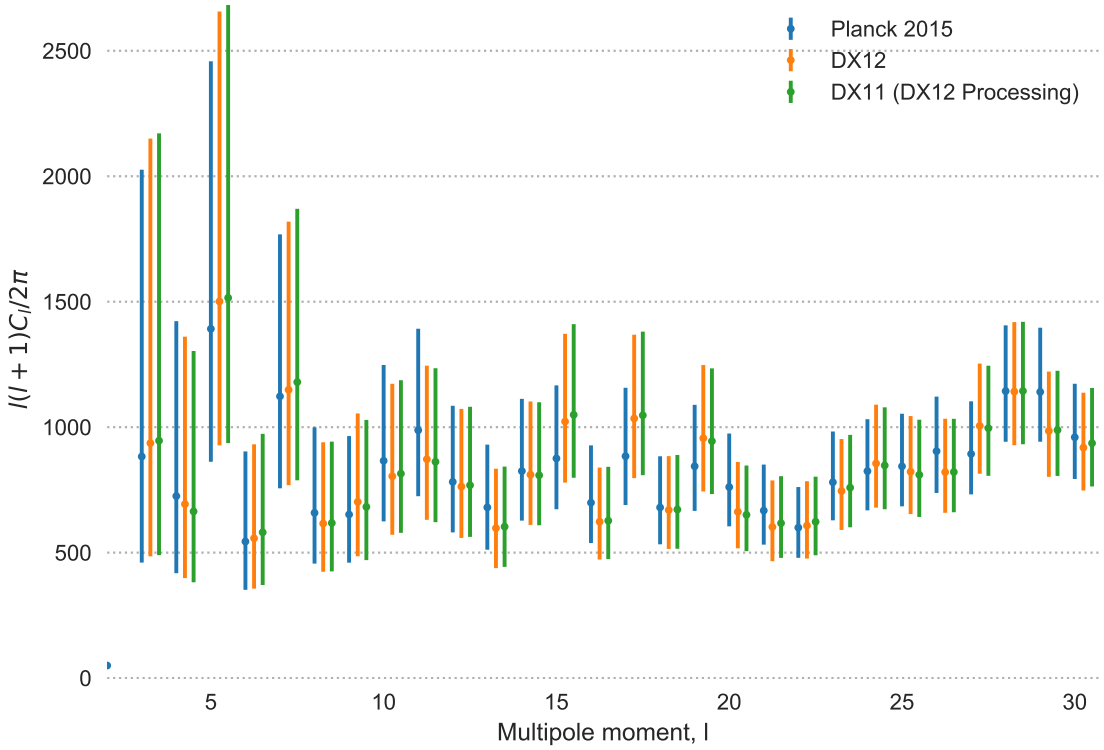


Figure 6.5: Comparison of Blackwell-Rao estimated CMB temperature power spectrum for DX12, DX11 and the official Planck 2015 release at low multipoles.

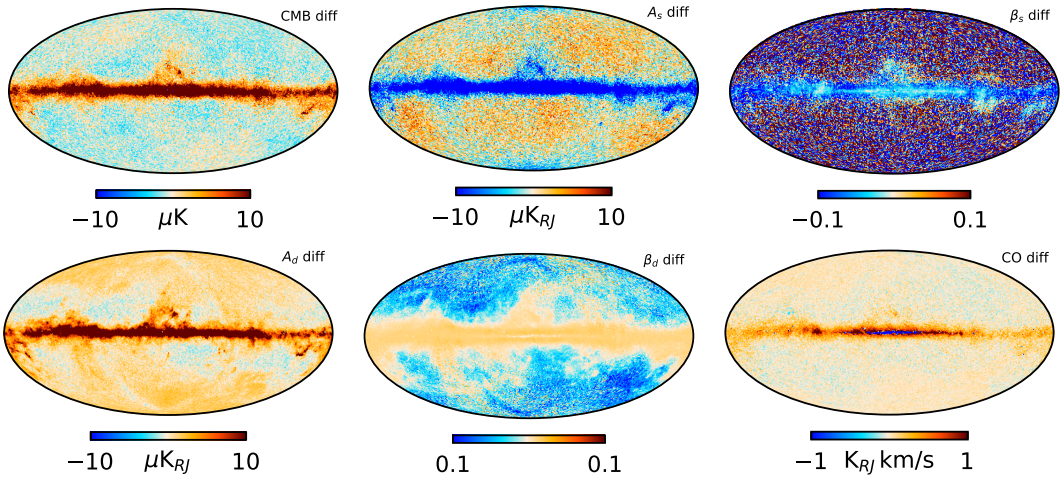


Figure 6.6: Component difference maps (DX12-DX11) for each model component. Most notable differences lie along the galactic band. Thermal dust β parameter sees the most change at high latitudes.

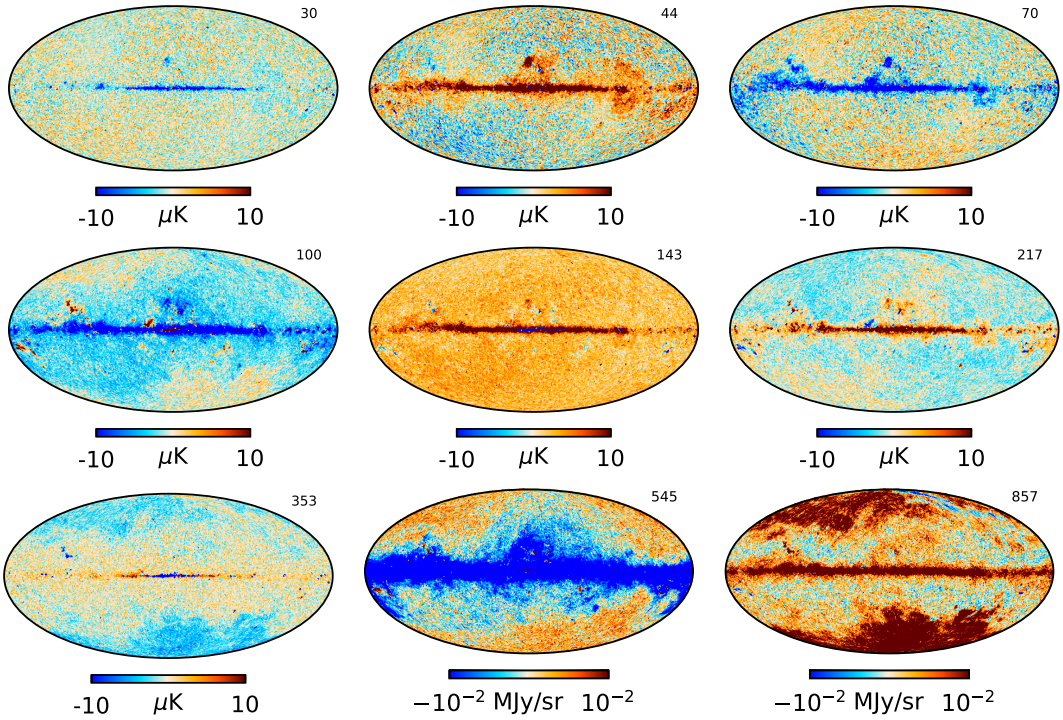


Figure 6.7: The DX11 residual maps are quite similar to those of DX12, including the free-free morphology stretching from 30 to 217 GHz with decreasing amplitude, as well as monopole offsets on 143 and 217. However, there are visible changes in high latitudes on the 545 and 857 GHz channels.

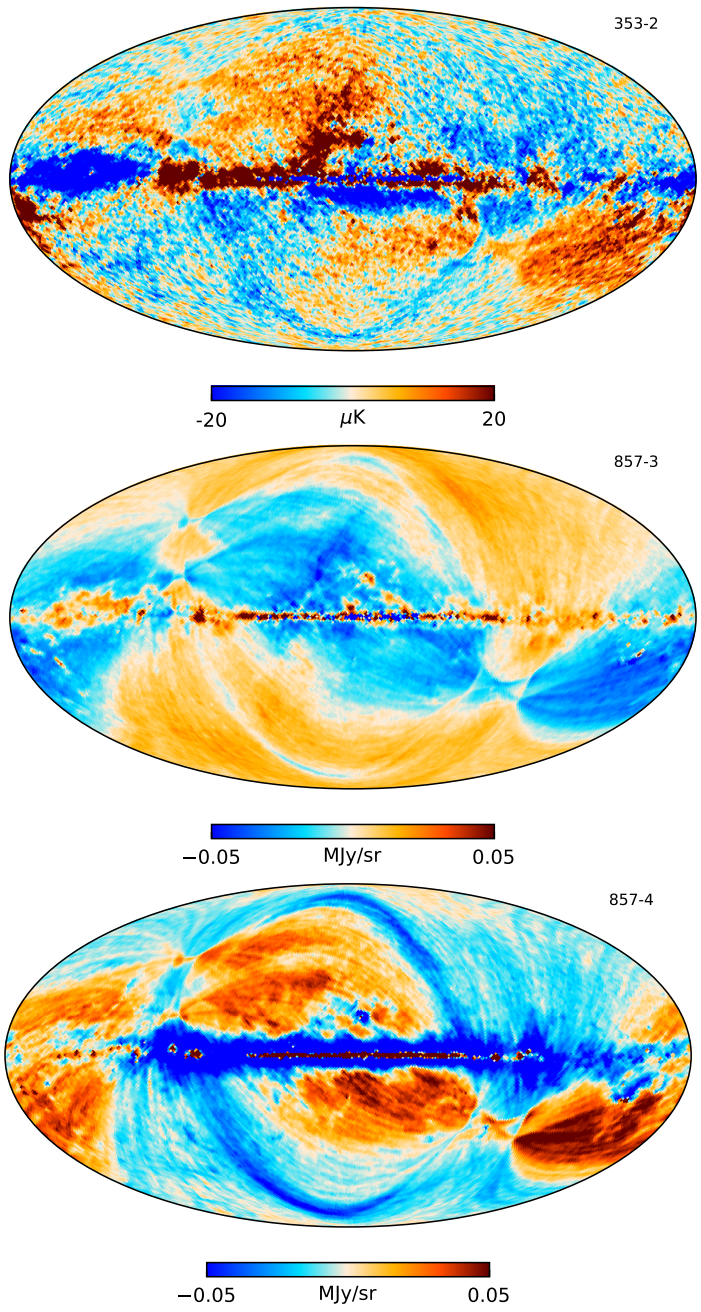


Figure 6.8: Residual maps from the Planck 2015 release exhibiting significant systematic errors. From top to bottom: 1) The 353-2 detector map residual showing polarization contamination in temperature. 2) The 857-3 detector map residual showing bandpass mismatch; dolphin morphology due to differences in bandpass profile. 3) The 857-4 detector map residual showing transfer function error. Plots courtesy of the Planck collaboration [46].

Chapter 7

The evolution of NPIPE sky maps

"People don't understand the pressure on me to look perfect"

— Kim Kardashian

Everything that has been done in this thesis so far has led up to this point. In this chapter, we will describe our efforts of component separation on the NPIPE sky maps. We will present problems encountered on the way, and the solutions we applied. We begin our analysis at a point in the progression of the NPIPE maps where map-making had reached a stand-still. At this point, instrumental effects could not be cleaned without encountering strong degeneracies with foregrounds, which could only be resolved with a new sky model. In this chapter, we present the details of our journey towards the best fit sky model for the NPIPE maps.

7.1 NPIPE map-making

The NPIPE sky maps are made by Reijo Keskitalo at Lawrence Berkeley National Laboratory, using the NERSC Pipeline (hence the name). These maps have been produced through a different approach to map-making compared to those used at the Planck DPC. While the core algorithm is the same, many small differences in correlated noise calibration and glitch removal add up to a notably different set of maps. Another difference between NPIPE and DX12 is the inclusion of more of the time-ordered data. NPIPE, for example, includes data gathered during repointing maneuvers of the satellite. The main difference between the two, however, lie in calibration. It is hard to attribute the visible changes to one specific process, as it is a result of a sum of many small ones. NPIPE is still very much a work in progress, and this thesis is a significant part of improving them.

While we take no part in the actual map-making, the process of component separation reveals systematic effects and provides a sky model used in calibration. This symbiotic relationship with map-making in the case of NPIPE has serious advantages, which will become evident as we progress through the generations of NPIPE component map results.

7.1.1 Sky-map comparison

In Figure 7.1 we present the difference between the NPIPE sky maps and those of DX11 from the Planck 2015 release, and the newest Planck DPC-maps, DX12 produced by SRoll. The most striking features in figure 7.1 are residual dipoles. Two separate effects are causing the dipole in these maps. The first one is the overall calibration differences, the second is due to fitting a far side lobe template in the map-making process of NPIPE, which induces an x-dipole effect. The second most prominent feature is the residual zodiacal light morphology, which is a result of the seasonal component of the Zodiacal light not being removed from the NPIPE maps. Additionally, due to changes in calibration and overall noise reduction we see a general static across the sky, but most prevalent in the LFI frequencies. On the HFI channels, we also see stripes following the scanning direction of the satellite, an effect attributed to the calibration of the baselines for correlated noise. Finally, we see what looks like transfer function error on the 353 GHz difference for DX12, which may be an effect of the steam-roll procedure.

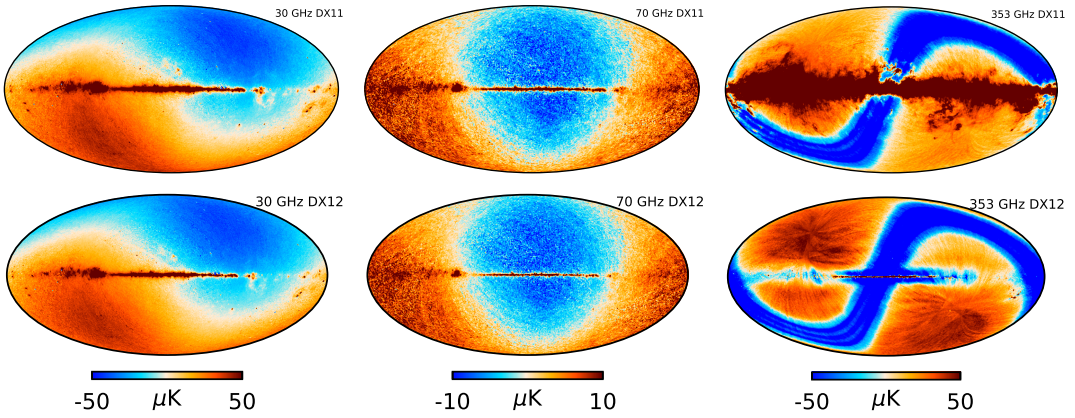


Figure 7.1: Top: Difference between DX11 and initial NPIPE sky maps for selected frequency bands. Bottom: Difference between DX12 and initial NPIPE sky maps for the same bands. The most prominent features in these maps are the residual dipoles. They arise from small differences in map-making techniques. For example, the dipole in 30 GHz is an effect of overall calibration differences. In addition to that, the center dipole seen clearly in 70 GHz, is a result of the use of a far sidelobe template in NPIPE. Differences along the galactic plane on higher frequencies are due to calibration differences in the DPC maps and NPIPE. The 353 map comparison displays some transfer function error, this is likely due to the application of the `SROLL` procedure.

7.2 NPIPE component separation

Naturally, we do not determine the best sky model on the first try. The component separation process is slow and steady. With a whole new set of sky maps, we must undertake the challenge of fine tuning every aspect of our model, from gain adjustments,

to bandpass shifts, monopole corrections and of course tweaks of the model parameter initial conditions. In this section, we outline the steps taken to obtain the final set of results. In order to summarize our analysis, we divide it into six milestones. An overview of these milestones are presented in table 7.2, which describes the main features of each analysis as well as the corresponding goodness-of-fit value obtained.

Milestone	N_{maps}	Notable features	χ^2_{full}	χ^2_{hilat}
1	9	Gain calibration on 545 and 857	1514.936	0.634
2	9	Dipole correction and gain calibration	1909.057	0.584
3	38	Single bolometer maps	1.989	0.516
4	30	NPIPE version 5 and single horn maps	1.575	0.439
5	33	More high-frequency sky maps	8.013	0.579
6	33	NPIPE5v21 sky maps	6.701	0.506
(7)	24	NPIPE5v21 fewer sky maps	0.650	0.270

Table 7.1: Overview of NPIPE analysis progression through six main sets of component separation results. Included are notable features applied to the milestones, number of sky maps N_{maps}, as well as their corresponding goodness-of-fit χ^2 value. An extra run is included to illustrate the effect of applying fewer high-frequency maps in the milestone six analysis.

The significant improvement underlines the advantage of this iterative process between map-making and component separation with the introduction of each new sky map version, and associated model fit. However, the progress illustrated in table 7.2 is not directly intuitive. High-frequency maps carry a higher weight in terms of χ^2, as they are harder to fit, and the goodness-of-fit is largely tied to the number of input maps. We once again remind ourselves that we wish to include as many sky maps as possible while maintaining a low χ^2, while at the same time obtaining clean frequency residual maps and physical component maps. During the processing of these, many errors were discovered and fixed, and are discussed in the better part of this chapter.

7.3 Preparing the first generation of sky maps

We start off at a point where the map-making for the NPIPE maps are quite close to that of DX12. We only have nine maps, and seven of them are full-frequency maps. Because the rest of the 545 and 857 NPIPE maps showed heavy instrumental contamination, they are left out while we construct the baseline model. The first set of NPIPE maps

used are NPIPE4v207 for LFI, and NPIPE4v205 for HFI, both smoothed to $N_{\text{side}} = 64$ and 40' FWHM.

The first step in preparing the NPIPE sky maps is the removal of the mono- and dipole. While `Commander` is equipped to do this, we choose to calibrate the dipole against DX12. We do this by taking the difference between DX12 and NPIPE for each frequency, mask over the galactic plane, and fit the dipole. This dipole is then removed from its corresponding NPIPE frequency map. Furthermore, because the NPIPE data do not have full frequency maps for 545 and 857 GHz, and hence no DX12 map to compare it to, the mono- and dipole correction was done by the less accurate method of approximating the dipole directly from the relevant sky map. Thankfully, the dipoles in the 545 and 857 GHz maps are extremely weak, as a result of the high signal-to-noise ratio of thermal dust emission from within our galaxy.

7.4 Remarks on presentation layout

As we finally present the NPIPE component separation data, we first remark upon the layout of its presentation. In order to get an impression of the progression made between each milestone, every set of maps produced are presented with its corresponding χ^2 and CMB map. We also illustrate our progress through difference maps between the current CMB map to the best-fit NPIPE map and that of the previously presented DX12 analysis. The full evolution consists of 6 sets of component maps. Furthermore, each subsection in between the presentation of each milestone outlines the measures taken to improve the results at each step of the analysis.

7.5 Milestone 1 - Gain calibration on 545 and 857 GHz

We start off by modeling synchrotron, dust and CO amplitude and line-ratios, as well as a joint low-frequency component. We are also fitting β_s, β_d and T_d per pixel. The main difference in the model compared to DX11 and DX12, is the latter.

Figure 7.2 shows the first meaningful results of component separation for the NPIPE sky maps. Starting from the top, we see that the χ^2 has a much smaller galactic band compared to that of DX12, and is overall much bluer than the either of the DPC counterparts. The galactic band now exhibits distinct free-free morphology, which is not surprising considering the fact that we only fit a joint low-frequency component. The reason for doing this is the strong degeneracies between foregrounds on low frequencies and too few frequency bands with Planck alone to break them.

The goodness-of-fit values are now

$$\chi^2_{\text{full}} = 1909.057$$
$$\chi^2_{\text{hilat}} = 0.584,$$

where the high latitude value has halved, which is promising for the end result. The full χ^2, in this case, is a result of a few pixel outliers. However, with a threshold removal of these outliers, the value is still significantly greater those of DX12 and DX11.

Next, the CMB map shows a clear blue line along the galactic plane. This is very typical, as the high-intensity signal at the center of the galactic disc is extremely hard to model. This shape is also very much degenerate with β_d.

Lastly, we have a look at the difference maps, the most interesting of which is the difference with DX12. The most striking feature of this map is the and quadrupole seen as blue areas in northern and southern hemispheres with a red band along the ecliptic. This points

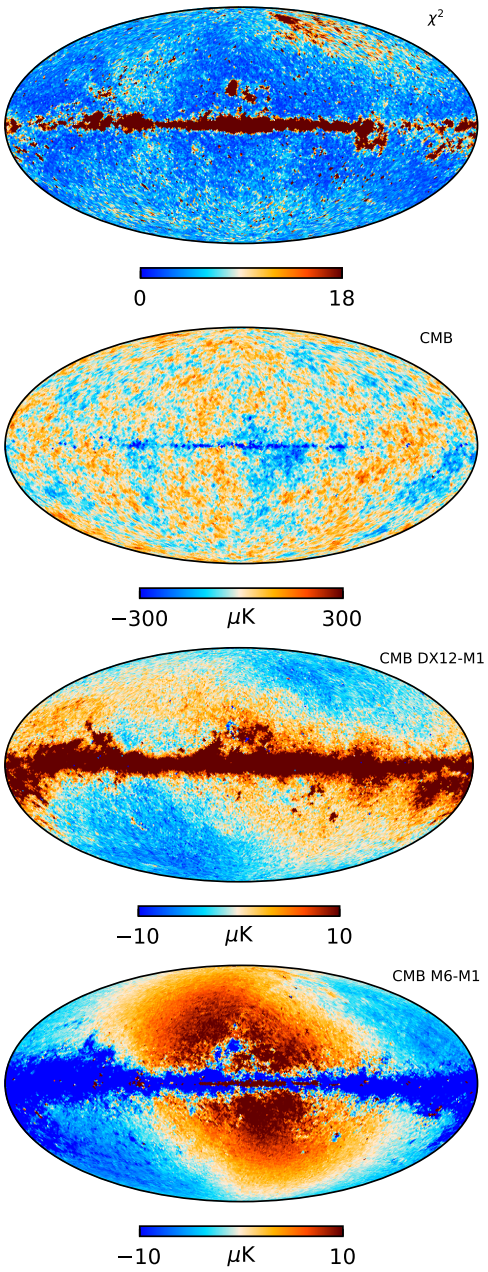


Figure 7.2: From top to bottom: 1) χ^2-full sky map, 2) CMB map, 3) DX12 CMB difference map, 4) CMB difference to best NPIPE map

towards the kinetic Doppler quadrupole [64] being present in our maps, which is most likely the reason for the red area seen in the χ^2 and should be corrected for in the dipole removal process. The intrinsic dipole difference between NPIPE and DX12 has been removed.

Additionally, there is a sharp, thick line across the galactic plane with dust/CO morphology. When encountering a signal like this in the difference maps, one must be careful not to treat it as a definite positive effect. This could also be attributed to over-estimation of the CO signal in NPIPE. However, judging from the clean χ^2 in this set, it is reasonable to assume that this is a result of successfully NPIPE model fit. Looking at the difference between the best fit NPIPE model, we see the full extent of the changes applied in the rest of the analysis. The residual x-dipole visible is an intrinsic difference in the NPIPE and DX12 maps discussed in section 7.1.1 and for our purposes not relevant. It will be removed from the reference map of DX12 in the rest of the analysis in order to clean up the difference maps, but the dipole will still be present in the actual data.

7.5.1 Gain calibration

In preparing this run, extensive gain calibration was carried out for the 545 and 857 GHz bands. As opposed to the lower frequencies, a gain calibration for the 545 and 857 GHz bands is difficult to fit within the Gibbs sampling framework due to degeneracies between gain and $\{A_{\text{dust}}, \beta_d, T_d\}$ as discussed in chapter 4. While lower frequencies may be calibrated against the CMB, the 545 and 857 channels have too strong dust contribution. Therefore, no more than 25 separate Commander runs were carried out in order to pin down the best combination of gain corrections. The result of this is illustrated in figure 7.4. This contour plot illustrates the combinations of gain values yielding the best goodness-of-fit value at high latitudes, interpolated from all 25 data points. Figure 7.3 also shows the full sky effect of changing the gain from 1.01 for both 545 and 857 to 1.0 and 1.1 respectively. The latter is the best fit, and the numbers we will use for the rest of this analysis. These numbers are most likely substantial, as a 10% gain calibration of 857 seems high but gives a reasonable fit. As evident by the vast differences between these maps, gain calibration is a vital part of the component separation process and will play an important part in the rest of the analysis. For this part of the analysis, all other gain values are set to 1.0.

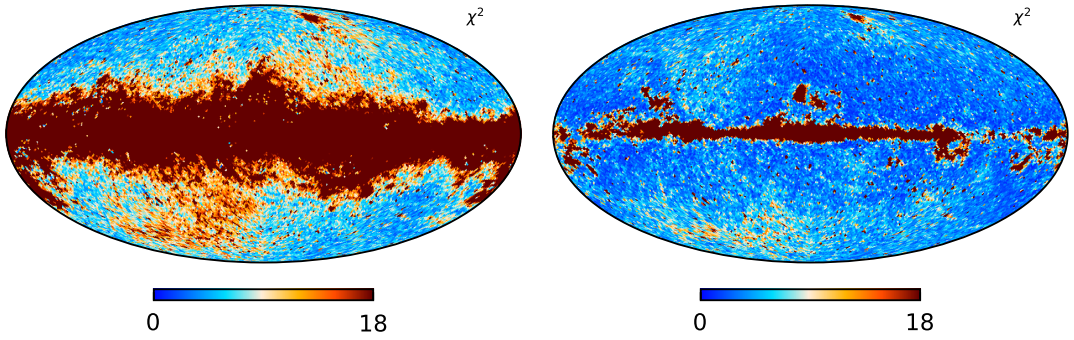


Figure 7.3: χ^2 of gain calibration on 545 GHz and 857 GHz. Left: Applying gain corrections of $g_{545} = 1.01$, $g_{857} = 1.01$, Right: Applying gain corrections of $g_{545} = 1.0$, $g_{857} = 1.1$. The figure shows how better gain calibration leads to our dust model picking up more of the signal, hence leaving a cleaner residual.

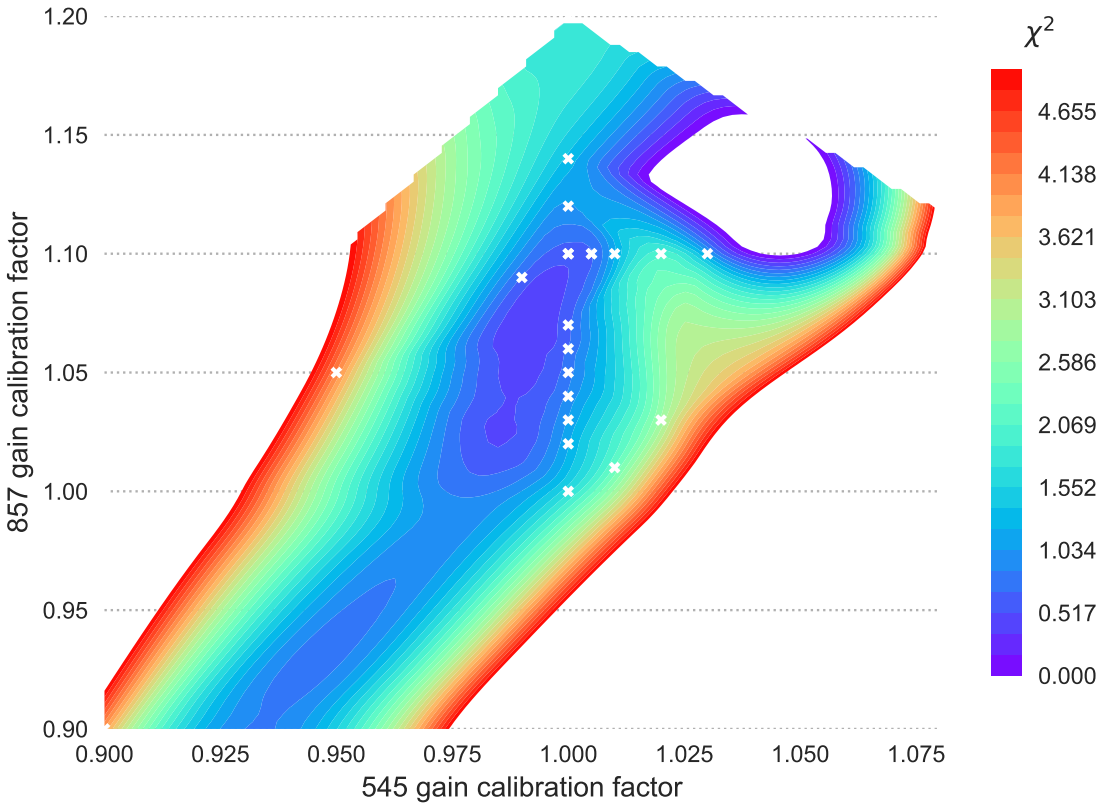


Figure 7.4: Contour plot of gain calibration numbers for the 545 and 857 sky maps with associated χ^2 value. The figure is interpolated from 25 separate data points. White markers denote a data point. The hole in the top right is an error in the interpolation due to few data point.

7.5.2 Quadrupole removal

Having identified a quadrupole in the previous milestone, we attempt to remove it. The kinetic Doppler quadrupole is an additional non-negligible effect which is induced by our proper motion through space [64]. Although the signal is weak, it should have been taken into account during the dipole-removal process. The monopole, dipole, and quadrupole can be expressed mathematically through the expansion of linearized temperature,

$$\frac{\delta I'(\nu')}{K} = \epsilon \frac{\delta T(\hat{n})}{T_0} + \beta z + \beta^2 z^2 Q(\nu') - \frac{1}{2}\beta^2. \tag{7.1}$$

Where the first term on the right-hand side is the monopole, the second is the dipole, the third and fourth is the frequency dependent kinetic quadrupole, the sum of the two latter terms is not equal to zero, but this is not an issue as all monopole factors are removed in `Commander`. For a more thorough discussion of the kinetic quadrupole, see Notari and Quartin [64].

With this knowledge, we reconsidered our dipole-removal strategy and instead construct plausible templates, using amplitude and direction values derived through a low-resolution dipole calibration run with `Commander`, by Harald Thommesen. This analysis fits the physical CMB dipole with all frequencies between 30 and 353 GHz. Figure 7.5.2 shows the process of dipole and quadrupole removal from the raw 100 GHz sky map.

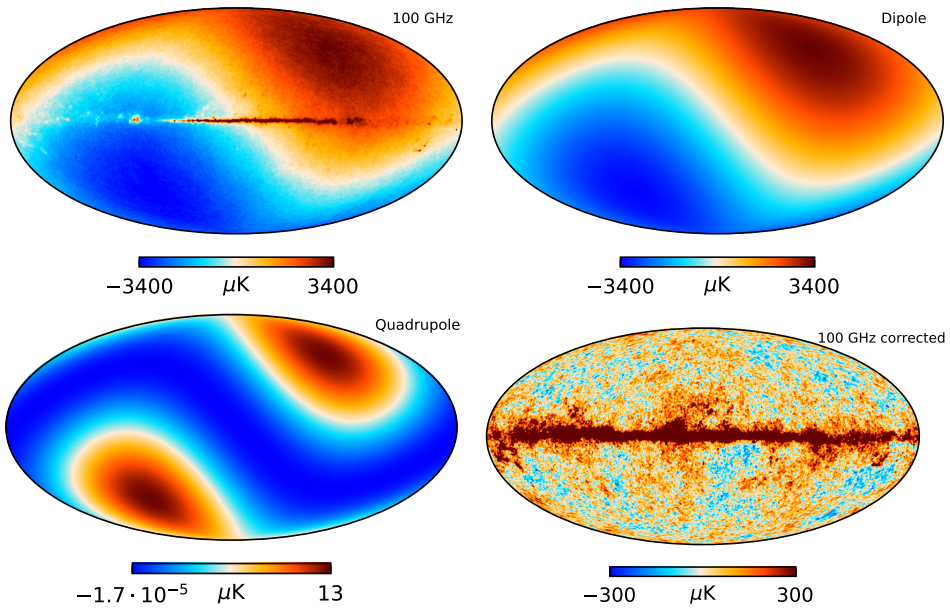


Figure 7.5: Dipole and quadrupole removal process for the 100 GHz NPIPE full sky map. Top left: uncorrected 100 GHz map. Top right: the computed dipole template. Bottom left: the computed quadrupole map. Bottom right: Corrected 100 GHz map.

7.6 Milestone 2 - New dipole correction method

Although we revamped the dipole removal process and also accounted for the quadrupole, we are still seeing dipole residuals in LFI, illustrated in figure 7.6, due to gain misestimation. Because of this, we now fit the residual LFI dipoles within `Commander`. In addition to the changes in dipole fitting, we make minor adjustments to our model. We set the monopole values to those of DX12 in order to have a starting point for sampling these, in addition to this, we calibrate gain across all channels, which up until this point have been set to 1.0. We always want to fit gain and monopoles where it is possible, and the reason for not doing it yet is simply because we want to modify our system one correction at a time. Additionally, because the CO is difficult to fit without probing the lines at different intensities using single-bolometer maps, we freeze their line ratios to 0.6 and 0.3 for 217 GHz and 353 GHz respectively.

Looking at the results, the χ^2 displays an overall improvement in high latitudes, the red patch from the quadrupole is now gone. The overall χ^2 has increased, however, with the numbers now being

$$\chi^2_{\text{full}} = 1909.057$$
$$\chi^2_{\text{hilat}} = 0.584.$$

Furthermore, the galactic plane looks thicker, and with more CO morphology, judging from the characteristic fan region on the far left, where we do not expect structure from free-free. This is expected with frozen CO line ratios and full frequency maps. In order to accurately model CO, we must apply single bolometer maps. By splitting up the full-frequency maps, we are left with one sky map per detector. This allows us to use the subtle differences in their bandpass profiles to probe the CO lines at different intensities. This is extraordinarily powerful for the 115 GHz line,

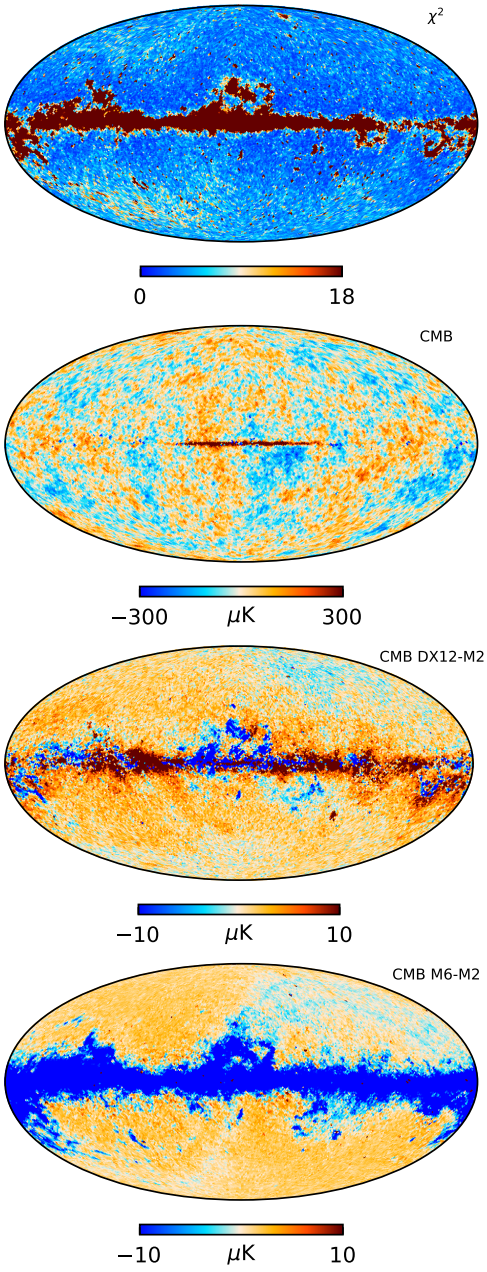


Figure 7.6: From top to bottom: 1) χ^2-full sky map, 2) CMB map, 3) DX12 CMB difference map, 4) CMB difference to best NPIPE map.

as it lies right on the edge of the 100 GHz Planck frequency band. Such an analysis is a natural next step in the analysis, made possible by the availability of single bolometer maps for NPIPE.

The CMB map shows a sharper red line in the galactic plane. A possible solution to this might be a better model of the thermal dust component, which should ideally be split into separate dust particle models. The DX12 difference map shows the galactic plane is a lot narrower this time around and suggests that the gain calibration at least brought us closer to DX12.

7.7 The caveats of single detector maps

Before presenting the results of component separation with single bolometer NPIPE maps, we discuss the caveats of such analysis. When splitting a polarization sensitive bolometer (PSB) into two separate detectors maps, we unleash a world of systematic effects. Because our main goal is to apply as many sky maps as possible, with as little residual signal as possible, we are poised to resolve these instrumental errors. These are all phenomena presented in the previous chapter which elaborates on instrumental errors.

In figure 7.8 we see some of the common errors encountered with the introduction of single bolometer maps. First, we look at the top left figure which shows clear signs of transfer function error. The tell-tale sign of transfer function error is that the galactic plane has been dragged out because the bolometer does not have time to sufficiently cool before registering a new photon. This effect should be accounted for by a transfer function correction during map-making but is showing a bad fit in this case. Secondly, the top right panel shows significant polarization-to-temperature leakage evident by the color change as we move along the galactic plane. This is the 353-2 Spiderweb bolometer, which should in principle only be weakly polarization sensitive, but in this case, its polarization sensitivity has been wrongfully estimated. The bottom two maps show a bandpass error in the process of being removed with the iterative `Commander` bandpass shift. The map on the left shows the first `Commander` sample, while the second is sample 50, suggesting that this effect can be significantly reduced during component separation. While bandpass error is an effect that may be accounted for in `Commander`, polarization leakage and transfer function problems are impossible to correct for on our end. If the errors are significant, we leave these maps out and report our findings to the map-maker. Removing detector maps does, however, come at the expense of our ability to finely measure the CO and thermal dust.

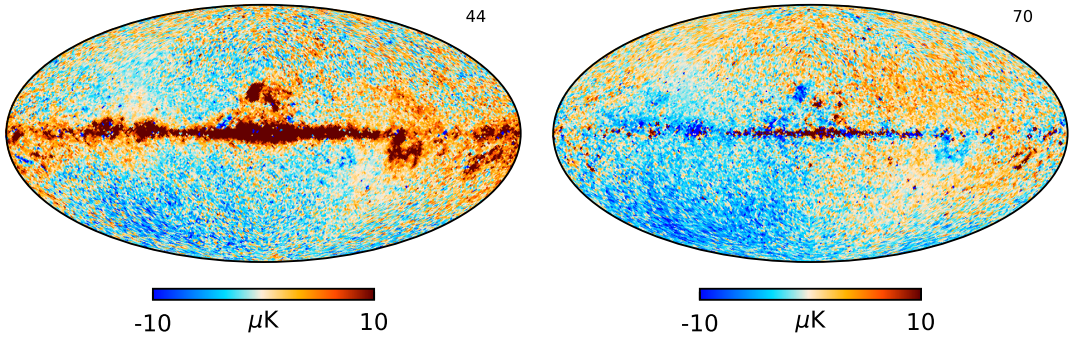


Figure 7.7: Weak dipole signature in the 44 GHz residual (left) and the 70 GHz residual map (right). This effect is a result of wrongful gain estimation for the applied dipole templates. These dipoles are now sufficiently fitted within `Commander`.

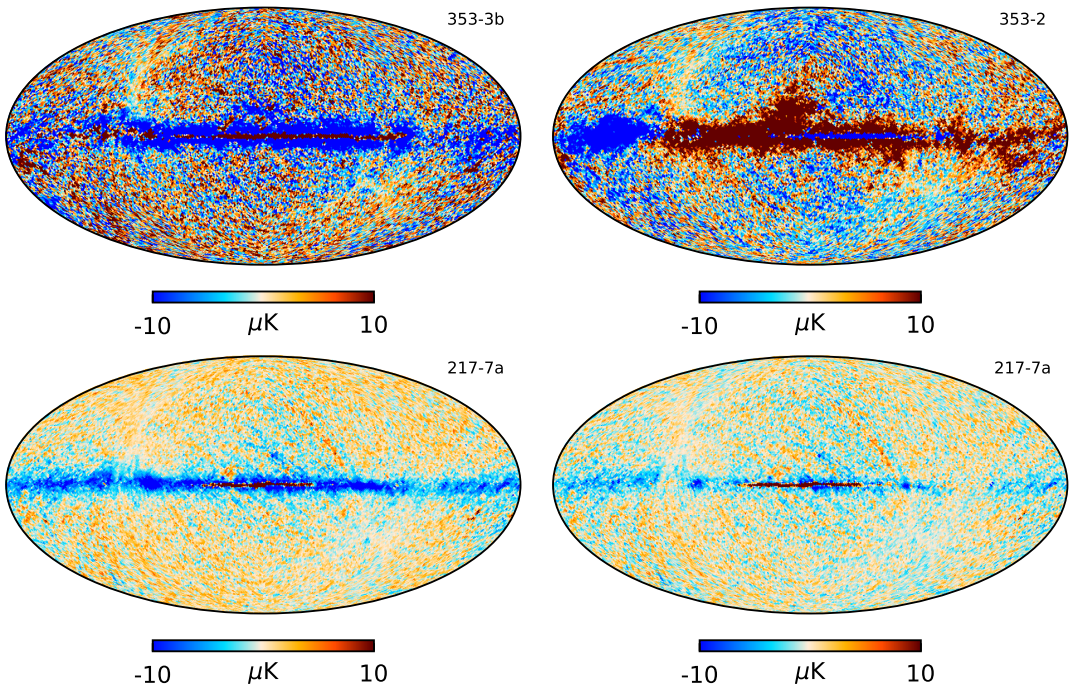


Figure 7.8: Residuals for selected detector maps, showing the most prevalent systematic errors in the map set. Top left we see Transfer function error on 353-3b, with the characteristic widening of the galactic band. Top right we see polarization-to-temperature leakage on the 353-2 Spider Web bolometer, which should in principle is only weakly polarization sensitive. The bottom two maps show Bandpass mismatch on 217-7a sample 1 and sample 50 respectively, visualizing the progression of `Commander`s iterative bandpass-shift fit.

7.8 Milestone 3 - Single bolometer maps

This milestone marks a significant step in our component separation analysis. We have split the 100, 217 and 353 GHz bands into single bolometer maps. This corresponds to a total of 8 maps for 100 GHz, where all detectors are polarization sensitive bolometers (PSBs), 12 maps for 217 with four spider-web bolometers (SWBs) and 8 PSBs, and 353 with another 12 maps from 4 SWBs and 8 PSBs. The inclusion of more maps allows us to fit all three CO lines separately. With one map serving as the reference for the line, we fit line ratios between detector maps, as opposed to fitting line ratios across frequency bands as we have done previously. In the calibration runs with single bolometer maps leading up to this version we have fine-tuned the gain and monopole values as well as applied bandpass profile shifts and fitted CO line ratios. The effect of the bandpass profile shift can be seen in figure 7.8.

In the χ^2 map we have less CO morphology, but it still exhibits a thick galactic plane. This thickening is an effect of adding more high signal-to-noise maps to our system. We note, however, that the χ^2_{full} value has decreased by a factor of 1000 with the removal of CO, effectively removing our outliers, and dramatically improving it over the DX11 and DX12 values. With two extra CO lines being modeled, we are also effectively adding two degrees of freedom to our system. The values are now

$$\chi^2_{\text{full}} = 1.989$$
$$\chi^2_{\text{hilat}} = 0.516.$$

Furthermore, if we take a look at the difference map between DX12 and NPIPE, we see strong evidence of CO morphology. This is a result of the extra maps being able to sample the CO line effectively at different amplitudes, which DX12 can not.

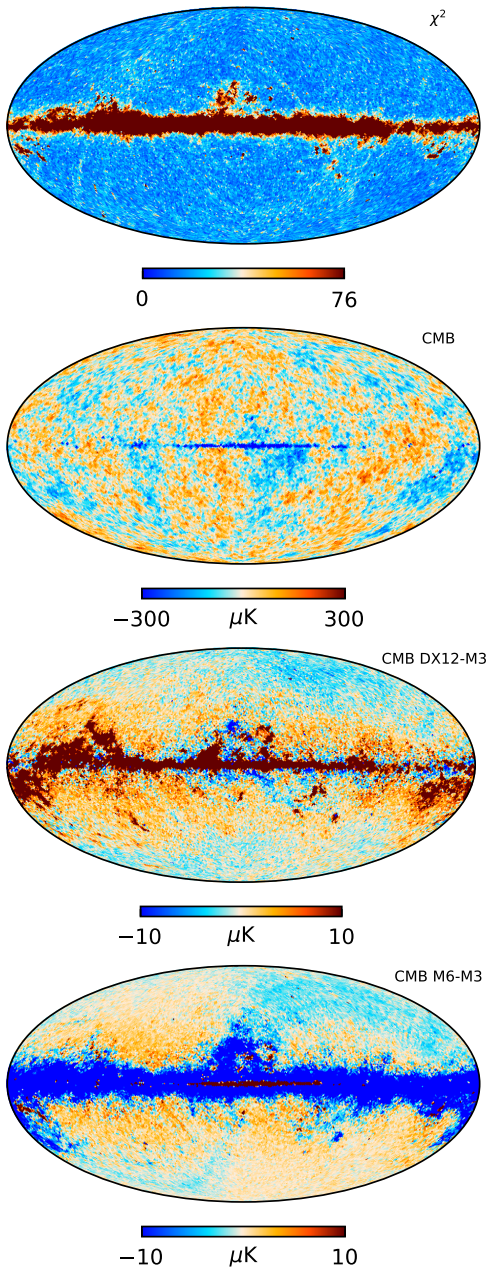


Figure 7.9: From top to bottom: 1) χ^2-full sky map, 2) CMB map, 3) DX12 CMB difference map, 4) CMB difference to best NPIPE map.

7.8.1 Fitting the three CO lines

Because each CO emission line is sharp defined by its respective rest frequency, we can not probe it across many frequency bands, which severely limits our ability to model them. The main motivation for splitting into single bolometer maps is that we gain the ability to probe the CO lines at many different intensities because their bandpasses differ slightly. This is a robust way of sampling CO that was not done in the Planck 2013 analysis, nor are they planned for the 2018 data release. In 2015, only two sets of combined detector sets at 100 GHz were used, which only probes the intensity at two points [46]. Our method is exceptionally robust for the $1 \rightarrow 0$ line, as it lies just on the edge of the 100 GHz Planck band, making it more readily discernible from its background. Figure 7.10 shows the three lines as sampled using the first generation of single bolometer maps. We especially note the sharpness of the $1 \rightarrow 0$ line as a result of its position on the 100 GHz Planck Band, and the diffuse morphology in the $3 \rightarrow 2$ line as a result of strong dust-foregrounds on the 353 GHz band. These lines are expected to be similar but with decreasing amplitude for higher transitions. Within each frequency band, we also assume a spatially constant fractional difference between the CO lines at each contributing detector, known as line-ratios.

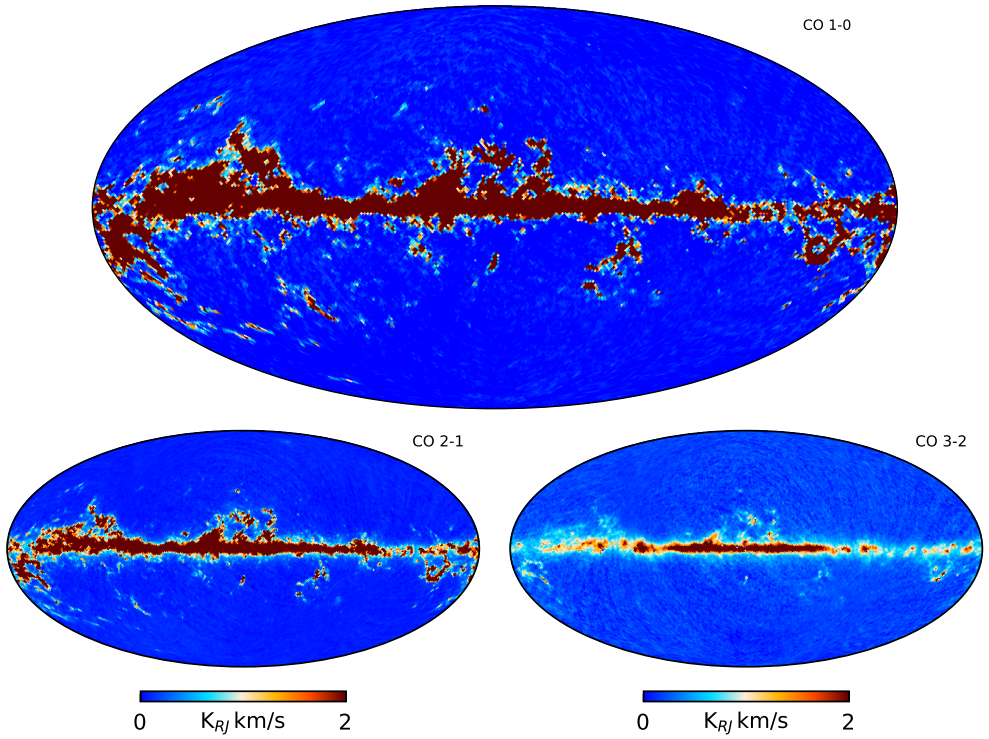


Figure 7.10: Three CO emission lines modeled using single bolometer maps. $1 \rightarrow 0$ being very sharp due to its position on the edge of our frequency band. $3 \rightarrow 2$ shows more dust contamination due to strong dust foregrounds in the 353 band.

7.9 Milestone 4 - NPIPE v5 and single horn maps

With the application of single bolometer maps, we were able to reveal significant instrumental errors present in NPIPE, illustrated in figure 7.8. For this set of results, we applied a new version of the NPIPE sky maps. These are the NPIPE5v4 and NPIPE5v5 maps corresponding to LFI and HFI respectively. Some of the improvements to this set includes the removal of less zodiacal light, in addition to not applying a variable gain correction, which was attempted on the previous maps without success. But most notably, these are single horn maps, ie. in the cases where the single PSB residual maps are dirty, we combine them with their partner which significantly reduces polarization-dependent signals. In this run, we have removed 217-1, 353-2, 353-7 and 353-8 and combined 217-6a/b, 217-8a/b, 353-5a/b and 353-6a/b.

Looking at our progression maps, the χ^2 shows slight improvement, with a thinner galactic band progressing towards free-free morphology. This is most likely an effect of using fewer maps at high-frequencies and not necessarily attributed to the new map set. The inclusion of this set of maps is mainly to underline the fact that removing maps in many cases improves the overall χ^2. This run shows a significant decrease compared to the previous with χ^2 values of

$$\chi^2_{\text{full}} = 1.557$$
$$\chi^2_{\text{hilat}} = 0.579.$$

At this point, the χ^2 values are looking very promising. With such a good model fit, we now shift our attention towards cleaning all the sky maps, in order to reveal systematic errors in every map. To do this, we must include maps that show strong signs of instrumental contamination, which will affect our χ^2 in the next set of results.

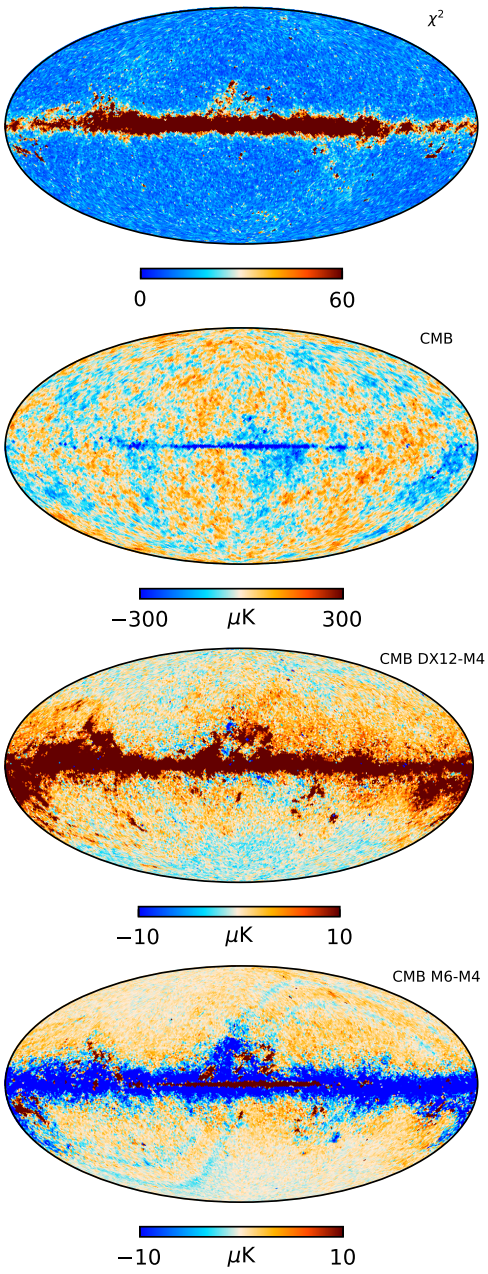


Figure 7.11: From top to bottom: 1) χ^2-full sky map, 2) CMB map, 3) DX12 CMB difference map, 4) CMB difference to best NPIPE map.

In the CMB map not much has changed, the blue line has gotten a bit bigger since the last time. The CMB DX12 difference map is showing more CO, which shares its morphology with dust, making it hard to discern if this is a better CO fit, or merely an effect of thermal dust. With the removal of maps, we do, however, expect our ability to model CO to worsen. The comparison with the best fit NPIPE now shows a slightly thinner galactic band, which we attribute to the reduction in detector maps.

7.9.1 Changing the dust β_d prior

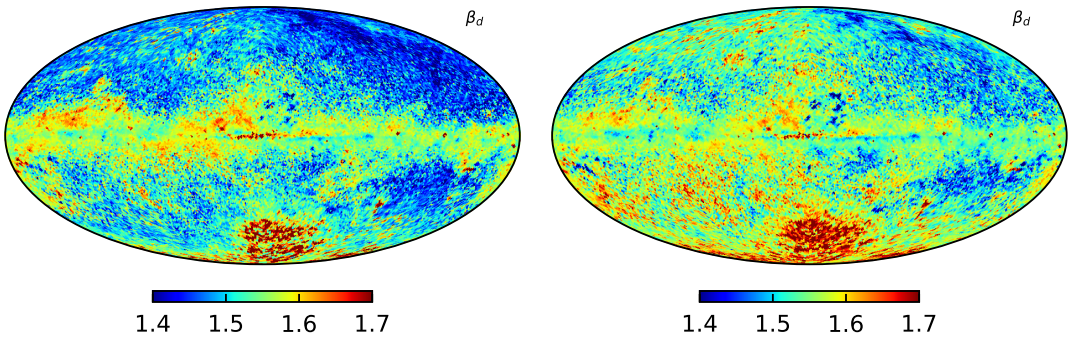


Figure 7.12: Adjusting the initial dust beta parameter gives a better β_d fit on high latitudes. Here, we adjust from 1.55 to 1.6, showing a more isotropic distribution, and therefore probably a better result for our model.

The slow and steady progress of improving our model, and subsequently removing instrumental errors, allows the employment of new maps at high frequencies. If we include more maps, we are better able to fit model parameters. With the introduction of single detector maps, we gained the ability to sample the dust β_d parameter with better accuracy. The parameter determines the slope of the thermal dust frequency spectrum, and therefore, greatly affects our models' estimate of thermal dust on lower frequencies. In addition to this, an accurate measurement of the β_d parameter is paramount in gain corrections, as it gives the component an intensity reference across frequencies when we have no CMB to calibrate against.

For the next analysis, we changed the prior for β_d to 1.6, up from 1.55. This change yielded positive results, as figure 7.12 illustrates a better overall agreement of the converged values across the sky. We note that the β_d-parameter was in the official Planck release observed to be 1.55 [47]. Although our measured value differs from that of the official Planck analysis, we do not expect this to be final, as we judge it quite volatile from our experience. However, we do expect it to converge towards a final value with less systematic errors in high-frequency maps allowing for more data points and a more rigorous model.

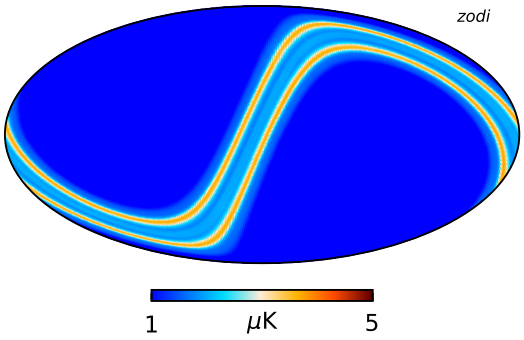


Figure 7.13: Template of zodiacal light applied in `Commander`. An effect of interstellar dust reflecting sunlight.

7.9.2 Template correction of zodiacal light

With the recent changes in the newest version of NPIPE maps, we must also account for an additional zodiacal light contribution. In the most recent set of sky maps, the decision was made to not include the seasonally varying part of the zodiacal light emission. The zodiacal emission is light reflected from dust particles within our solar system and into the detectors of Planck. The locality of this emission, therefore, make it seasonally dependent. We now remove the residual zodiacal light through pre-generated templates in `Commander` for each frequency channel, illustrated in figure 7.13 [65].

7.9.3 Applying a new calibration mask

With our focus shifting towards cleaning residual maps for use in the next iteration of map-making, we want to include more maps at higher frequencies. Earlier we have commented on how all but one of each of the 545 and 857 GHz maps have been left out of the analysis due to excessive systematic contamination. These are now re-introduced. In order to sufficiently clean these of foreground contamination, we anchor the 857-2 map, and use it as a reference for thermal dust, while we calibrate monopole and gain shift of the remaining channels. This only partially cleaned the residual maps, because the mask applied in the calibration was too big. Therefore, we applied a threshold to one of the 857 sky maps and created a smaller mask which can be seen along with the results of the application of it in figure 7.14.

Additionally, we now let the 353 band calibrate against thermal dust. This significantly improved all the high-frequency residual maps. However, noticeable systematic effects remain such as strong polarization leakage in the 353 GHz sky maps illustrated in figure 7.15. We are unable to correct for polarization systematics on our end, and this signature can therefore only be removed during map-making. The morphology visible in figure 7.15 is a sign of sub-optimal calibration of the polarization sensitivity of the SWB, which leads to excessive leakage [59].

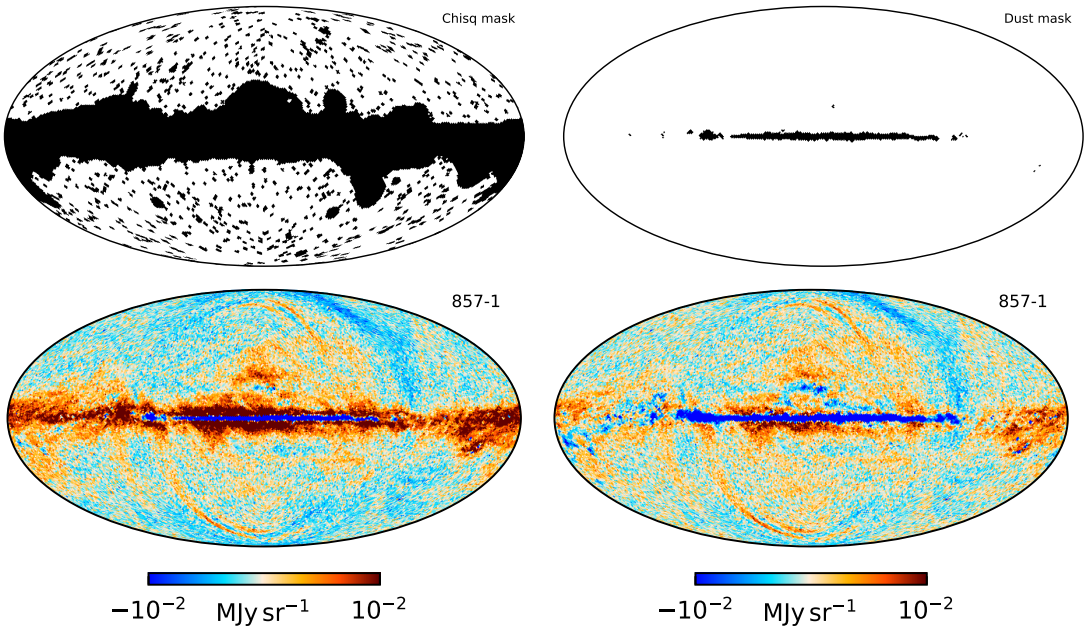


Figure 7.14: We apply a new calibration mask to high-frequency bands in order to clean their residual maps. Top left: the "Chisq" mask, which was previously used for this task. Top right: The new mask created from by applying a threshold to an 857 sky map. Bottom left: the 857-1 residual map using the old mask. Bottom right: the 857-1 residual map when gain and monopoles are calibrated using the new mask.

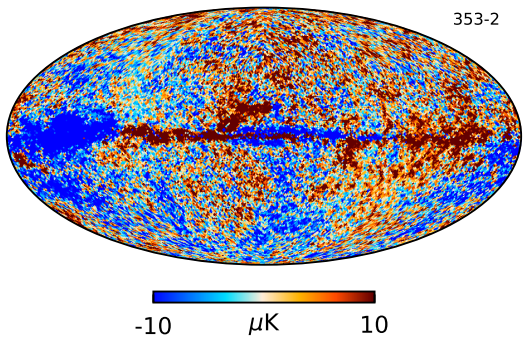


Figure 7.15: Residual map of 353-2 showing the prominence of polarization as changes in color along the galactic plane. This is attributed to the miscalculation of polarization sensitivity in the SWBs. Although we identify the polarization leakage by its features along the galactic plane, the effect of the contamination spans the full sky.

7.10 Milestone 5 - More high-frequency sky maps

With our model converging towards a god fit, we start including more data to our model. In this set we reintroduce all 353, 545 and 857 single horn maps, and attempt to remove foreground contamination by sampling gain and bandpass effects with the new calibration mask. The introduction of more maps leads to an increase in χ^2 as expected, with values

$$\chi^2_{\text{full}} = 8.013$$
$$\chi^2_{\text{hilat}} = 0.579.$$

While these values tell the story of a worsened χ^2, our map is now showing morphology matching that of free-free, with a narrower galactic band due to an improved thermal dust fit. We hypothesize that the blue and red lines visible in the center of the CMB require a more complicated dust model to be removed.

Moving on to the CMB difference to DX12, we see a very thick galactic band with obvious dust morphology, also indicative of an improved thermal dust fit. Because the DX12 analysis employed full frequency maps, it only probed the thermal dust component at three different data points (353, 545 and 857 GHz), this set measures at 13 data points. Unfortunately, there are drawbacks to this. While some maps show significant improvement with the implementation of our new calibration mask, the 353 GHz maps still exhibit significant polarization contamination.

The improvements made to the thermal dust fit underscores the value of error-free detector maps on our system. We must of course not attribute all these changes to the new maps since we also adjust the dust β_d prior. We expect both changes to contribute to the result. We also remark upon the lack of any zodiacal light in the difference maps, illustrating a successful removal with our template fit. The difference to the best fit NPIPE sky maps

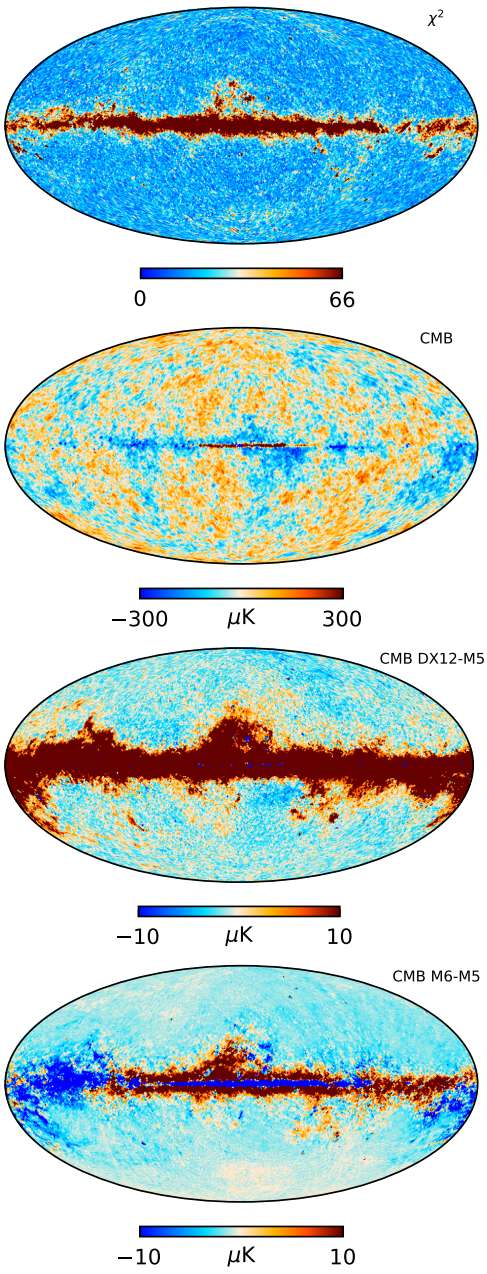


Figure 7.16: From top to bottom: 1) χ^2-full sky map, 2) CMB map, 3) DX12 CMB difference map, 4) CMB difference to best NPIPE map.

shows significant changes since the last time. The two milestones seem to agree in morphology on high latitudes, with a slight offset, while the main differences lie along the galactic band. Additionally, we note the vague power of the CIB signature in this map. The last milestone we introduce is our final result, and many changes can cause these differences, the most important one; the final sky maps from map-making.

7.10.1 North and South pole signal asymmetry

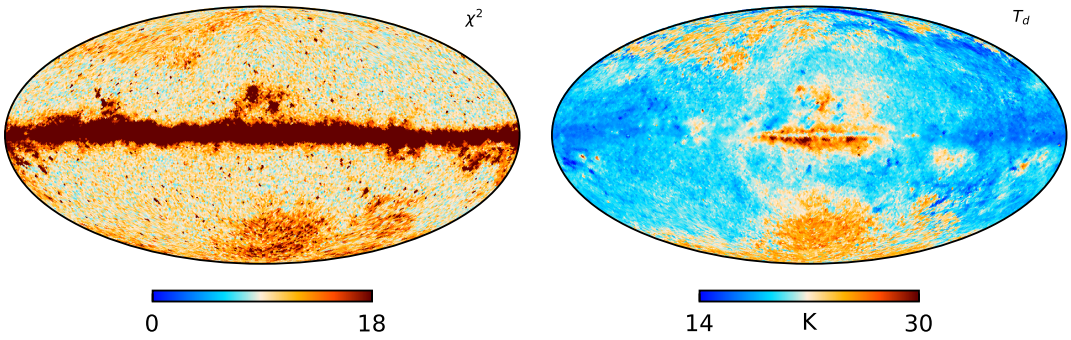


Figure 7.17: This figure shows the apparent north-south asymmetry in the CIB structure. Left: DX11 χ^2 map illustrating the difference in strength of the CIB morphology on the north and south hemisphere. Right: Dust temperature map from milestone 5 exhibiting the same CIB morphology asymmetry.

Throughout the analysis, we have noticed a peculiar north-south power asymmetry with apparent CIB structure at high-frequency residuals and also the dust component maps. The reason why we deem this peculiar is that we expect such a signal to be uniform. On further inspection, this effect is present in both DX11 and DX12, but more prominent in the former, as can be seen in figure 7.17.

Chapter 8

New state-of-the-art sky maps

"Once you replace negative thoughts with positive ones, you'll start having positive results."

— Willie Nelson

We now present the final set of NPIPE sky maps for this thesis, which we boldly refer to as the new state-of-the-art sky maps. In this chapter, we will discuss the changes made since the last milestone, before presenting the results of the final component separation analysis employing these sky maps to justify our claim.

8.1 The NPIPE5v21 sky maps

In conjunction with our efforts of component separation, Reijo Keskitalo has created a new set of NPIPE sky maps. By using the previous results from component separation as templates in the map-making process, a new, and "final" version of NPIPE have been born. This set of maps is called NPIPE5v21, henceforth referred to as npipe21. This set of maps appear to be the converged set of NPIPE temperature sky maps [59]. One of the most prominent changes in this set is the re-calibration of the polarization sensitivity of the SWBs, which adequately removes polarization leakage effects from our 353 GHz SWB residual maps. The latest estimation of the polarization sensitivity of the SWBs, shows a change in over 3% for the 353-1 SWB, reducing it to 0.18%, rendering it practically insensitive to polarization, resulting in a colossal difference in the residual map. Prior to this, the template fitting of polarization has in effect attempted to remove polarization signal that was not present in the data.

Moreover, corrections have been made to both bandpass mismatch and noise calibration effects, as well as improvements on transfer function residuals. In order to visualize the immense improvement, we look to figure 8.1, which compares the omitted detector residual maps of the 2015 official Planck analysis to the same detector residual maps of npipe21 (milestone 6), where the middle column is the corresponding milestone 5 residuals. From top to bottom we see a massive reduction of polarization leakage in the 353-2 detector, an effect of roughly equal magnitude across all 353 SWB detector

maps in npipe21. The second row shows the correction of bandpass mismatch, which was already significantly improved upon in earlier iterations of NPIPE, but now shows even better removal. The bottom row shows the major advancements made in the removal of transfer function residuals. Even the difference from milestone 5 to 6 is very encouraging.

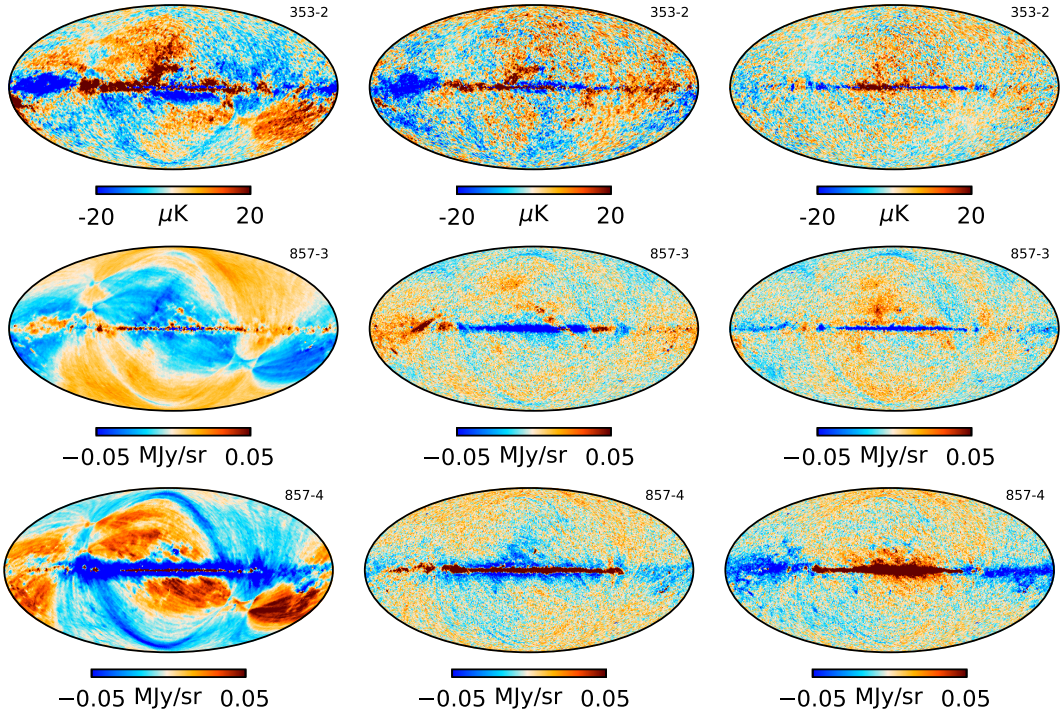


Figure 8.1: Comparison of the same three detector map residuals as presented in the official Planck DX11 analysis [46], an older NPIPE version (Milestone 5), and the best fit NPIPE (Milestone 6). From top to bottom we have: 1) 353-2, 2) 857-3, and 3) 847-4. This figure best illustrates the improvements made to the detector sky maps in polarization contamination, bandpass mismatch and transfer function errors respectively.

8.2 Milestone 6 - Final NPIPE sky maps

Milestone 6 marks the end of our journey of improving the NPIPE sky maps. In this milestone, we have employed the npipe21 sky map set in component separation, and the results are presented in the remaining part of this chapter. On top of the introduction of new sky maps, we have changed some important aspects of our model. Priors for thermal dust have been adjusted to $T_d = 20.1$, $\beta_d = 1.53$, and the low-frequency parameter to $\beta_s = -3.0$, down from $T_d = 23$, $\beta_d = 1.6$, and $\beta_s = -3.1$. These changes are justified in section 8.2.2.

First and foremost, we note that milestone 6 (M6) will rely on the component maps derived from the application of 33 sky maps. However, to show an ideal model fit using the same set of data, we have carried out an analysis employing fewer input maps. The χ^2 value results are

N_{maps}	χ^2_{full}	χ^2_{hilat}
33	6.701	0.506
24	0.650	0.270

The main difference between these two is the introduction of more sky maps on high frequencies. The low-frequency domain is the same, and hence the model fit for the low-frequency component. Therefore, the difference we see in these two maps can be attributed to the constraints applied to our thermal dust component with the inclusion of more data. We include the set with fewer sky maps, to illustrate the possible improvement in model fit for a simplified system, with the application the new set of sky maps. The values presented show immense improvement, not only does the simplified system display χ^2 values dramatically lower than our initial set of sky maps, but the full run also improves

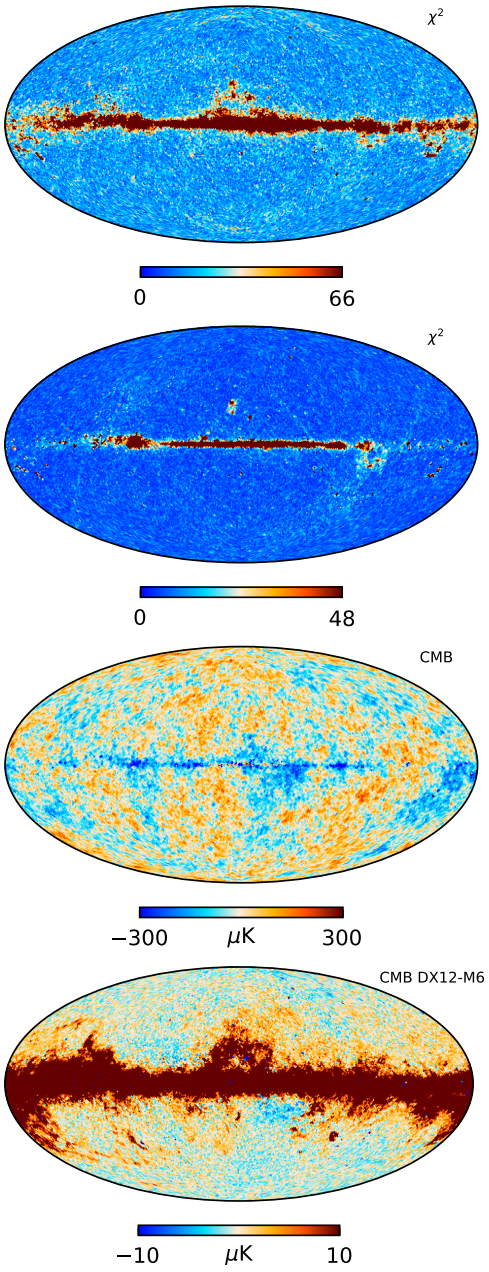


Figure 8.2: From top to bottom: 1) χ^2-full sky map, 2) χ^2 with fewer sky maps, 3) CMB map, 4) DX12 CMB difference map.

upon the dust fit of the previous model. The χ^2 map with fewer sky maps, second in figure 8.2, illustrates how the remaining model fit on low-frequencies can be resolved by accurately probing free-free emission, discussed in section 9.2.1. Furthermore, a direct comparison of the full M6 χ^2 map to that of M5, shows a thinner galactic band owing to our improved ability to fit thermal dust following the removal of systematic errors in high-frequency maps.

The M6 CMB map shows a slight improvement over the previous milestone in regards to the red line in the galactic center. Finally, we note the dramatic differences between DX12 and the best fit NPIPE model. This map shows yellow thermal dust morphology on high latitudes along with a strong red galactic band. We conclude that this is the result of more thermal dust removal in M6, likely a result of the changes made to the sky model priors, deduced from our improved ability to model thermal dust.

8.2.1 Milestone 6 residual maps

For this specific milestone, we also present the full set of residual maps illustrated in figure 8.3, showing what is only white noise in the majority of the maps. There is, however, some free-free residual morphology present in the LFI channels, as a result of the inaccuracy of the low-frequency component. Unfortunately, due to a misstep in the definite `Commander` run from which these results are taken, the 353 band exhibits improper gain adjustment. This is a result of sampling gain and bandpass at the same time while calibrating the sky maps against thermal dust. This causes the gain calibration to compensate for the band-pass shift. This is in principle possible when calibrating against the CMB because the nature of the phenomena breaks the degeneracies as previously discussed. Fortunately, this effect does not have dramatic implications for the overall result. There are, however, not only bad news. The residual high-frequency maps show improvement across the board, with significant improvement in regards to polarization leakage in the 353 GHz band.

Along with all residual maps, table 8.2.3 shows an overview of the final calibration numbers, consisting of estimated gain, monopole corrections, and bandpass shift values for all included sky maps. With new and clean sky maps, the estimated gain values also change. We note how the table illustrates the 353 GHz degeneracy between bandpass and gain calibration. We expect gain calibration to be within one percent, and the bandpass to not exceed 1 GHz shift.

8.2.2 Milestone 6 component maps

On top of the new version of NPIPE sky maps, there are quite a number of changes to our model, the effects of which are most prominent in the component maps. First of all, we change the priors for all parameters on both the low-frequency and the thermal dust component. The thermal dust parameter priors are now changed to $\beta_d = 1.53$ and $\mathrm{T}_d = 20.1$ as opposed to 1.6 and 23.0 in the previous runs. These values show a visually acceptable fit, with the β_d being almost perfectly uniform, and the high latitude CIB morphology removed in both maps as seen in figure 8.4. The low-frequency prior is

changed to $\beta_s = -3.0$ as opposed to -3.1, which also improves the uniformity of the map, the yellow areas are showing strong free-free morphology. Because of the importance of the estimated CO lines, in regards to this study, with the application of single detector maps, these are presented in their separate section.

Sky map set	β_s	T_d	β_d
DX11	1.60	19.7	-3.06
DX12	1.58	19.7	-3.07
npipe21	1.55	21.1	-2.99

Table 8.1: Estimated pectral parameters from DX11, DX12, and npipe21 sky maps.

Furthermore, table 8.2.2 shows the converged estimates of the component parameter values for the low-frequency and thermal dust components compared to those of DX11 and DX12, along with the value for a `Commander` run applying more maps on high frequencies. The changes in parameters are quite significant, especially that of the β_s low-frequency parameter. These values are, should not be regarded as definite for the DX11 and DX12 map sets, because the analysis of which fell short of ideal, as previously discussed.

8.2.3 Best fit CO component

Figure 8.5 shows a T-T scatter plot of two sets of $1 \rightarrow 0$ CO emission maps compared to the Dame et al. full sky CO survey [66]. From this, we deduce that our DX12 map is unable to adequately sample CO, as expected, considering we are only using full-frequency maps, and the foreground removal process has not yielded satisfying results. Next, we show a comparison with the best fit NPIPE model of milestone 6. Both of these are presented with their associated Pearson correlation coefficient, with an impressive $r = 0.997$ for the case of milestone 6. The central plot in figure 8.5, of the $2 \rightarrow 1$ line, shows almost equally impressive precision. The difference between Planck and Dame et al. was in the latest Planck release attributed to a combination of bandpass uncertainties in Planck, and an overall calibration uncertainty in Dame et al., as well as thermal dust leakage [46]. Finally, the bottom plot of figure 8.5 shows the case of the $3 \rightarrow 2$ line, where the correlation is no longer as significant showing high scatter, a combined result of the present thermal dust, as well as the difference in amplitude of the line. The actual full sky CO maps are illustrated in 8.6 with impressive accuracy.

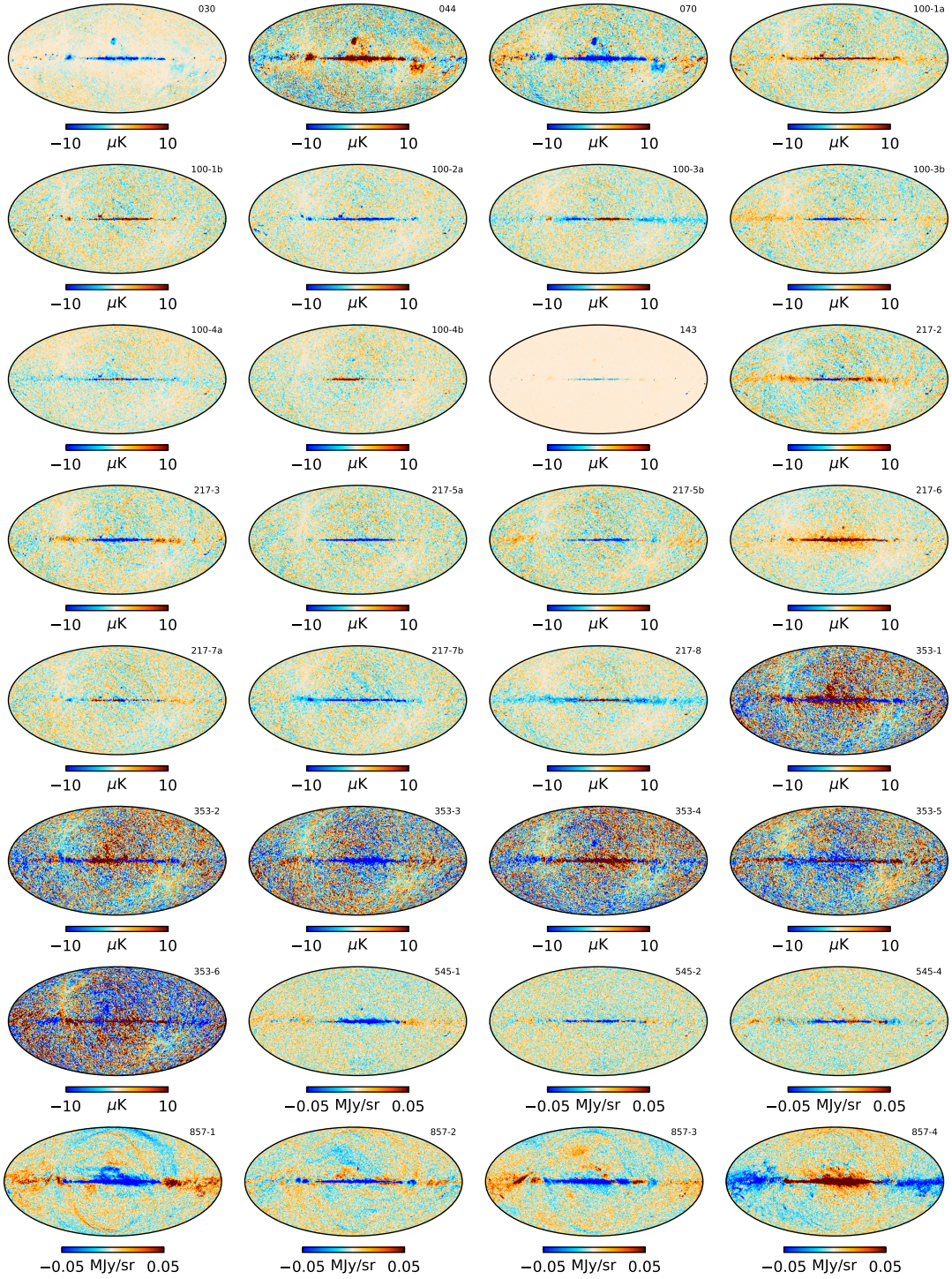


Figure 8.3: Milestone 6 residual maps. Visible free-free in LFI, visible calibration issues in the 353 GHz band, but overall improvement on HFI. 100-2b removed for formatting.

Instrument	Frequency [GHz]	Detector label	Calibration factor (gain)	Monopole correction [μK]	Bandpass shift [GHz]
Planck LFI	030		0.982432	0.000000	
	044		0.984135	14.11390	
	070		0.995503	19.54450	
Planck HFI	100	1a	1.002459	8.297117	
		1b	1.001939	7.204038	
		2a	1.003558	9.445082	
		2b	1.003431	9.540048	
		3a	1.002403	8.666812	
		3b	1.002277	7.164984	
		4a	1.004222	10.38996	
		4b	1.003666	9.791668	
	143		1.000816	22.00000	
	217	2	0.999717	68.65903	-0.779400
		3	0.999586	74.37238	
		5a	0.999047	68.79398	
		5b	0.999735	67.74893	
		6	1.000076	63.44133	
		7a	0.999051	67.94434	-0.541300
		7b	0.998537	68.93946	
		8	0.999484	65.88998	-0.289800
	353	1	1.047304	416.8616	-3.018900
		2	1.030700	419.9327	-2.401600
		3	1.008896	414.3506	-0.558700
		4	1.016336	432.7418	-0.670800
		5	0.997274	410.0113	0.150100
		6	1.034466	402.8273	-2.411900
	545	1	1.001981	0.3596256	
		2	1.026543	0.3598056	
		4	1.022230	0.3592131	
	857	1	1.077533	0.6226399	
		2	1.100000	0.6200000	
		3	1.081117	0.6248062	
		4	1.118268	0.6309479	

Table 8.2: Milestone 6 gain, monopole and bandpass corrections. We note correction errors on 353, due to simultaneous sampling of bandpass and gain while calibrating against thermal dust. Monopole corrections for 545 and 857 detectors are given in MJy/sr.

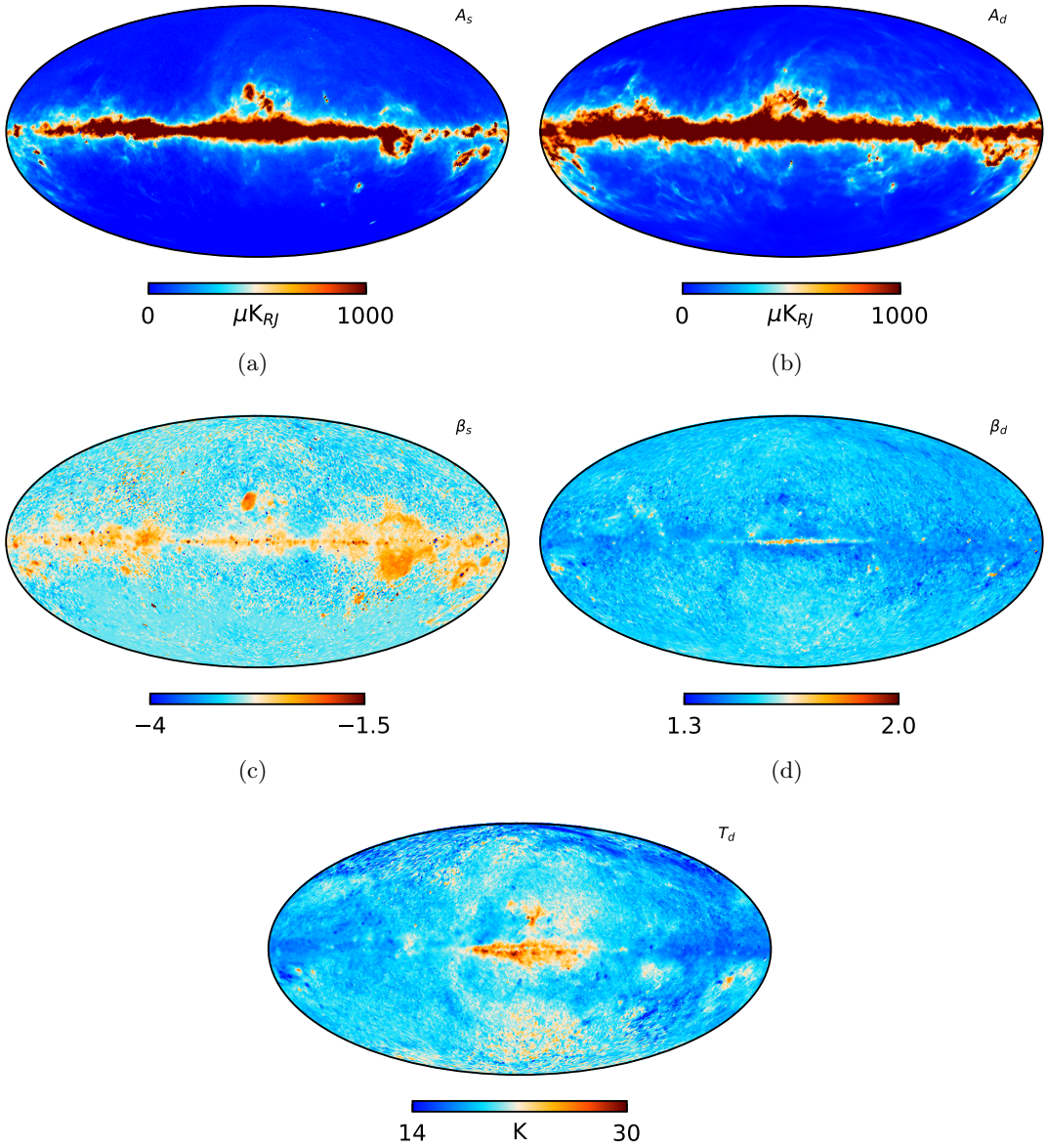


Figure 8.4: Component maps not including the CMB derived from milestone 6. Maps a) and b) show the amplitude of the low-frequency, and the thermal dust component, respectively. Map c) shows the β_s parameter of the low-frequency component. Maps d) and e) show β_d and T_d parameters of the thermal dust component.

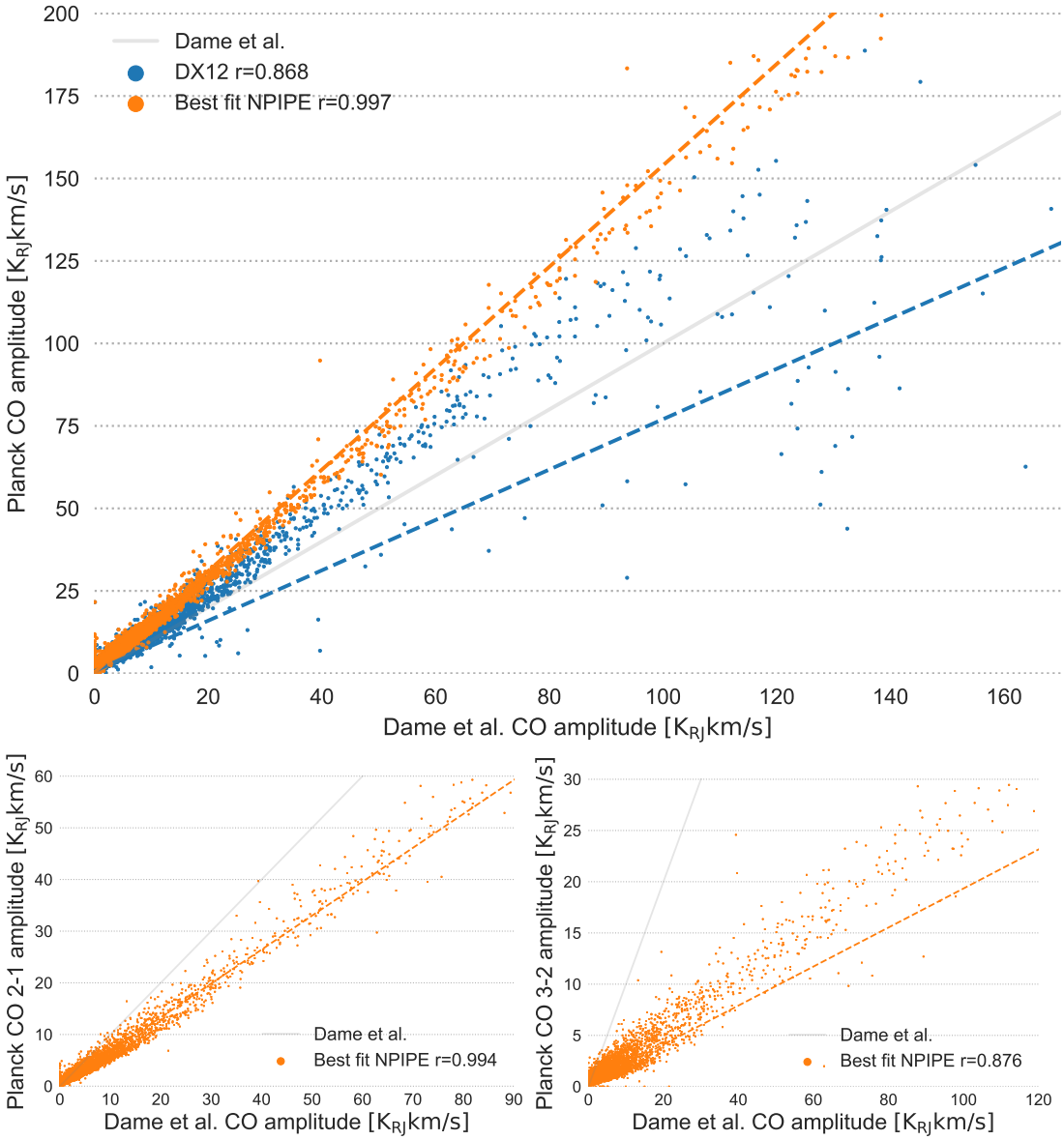


Figure 8.5: T-T scatter plot comparison between the full sky CO survey of Dame et al. [66] and each CO emission line as sampled in milestone 6 (M6). Top: CO $1 \rightarrow 0$ for DX12 and M6. Bottom left: CO $2 \rightarrow 1$ for M6. Bottom right: CO $3 \rightarrow 2$. Represented with their corresponding Pearson correlation coefficient r.

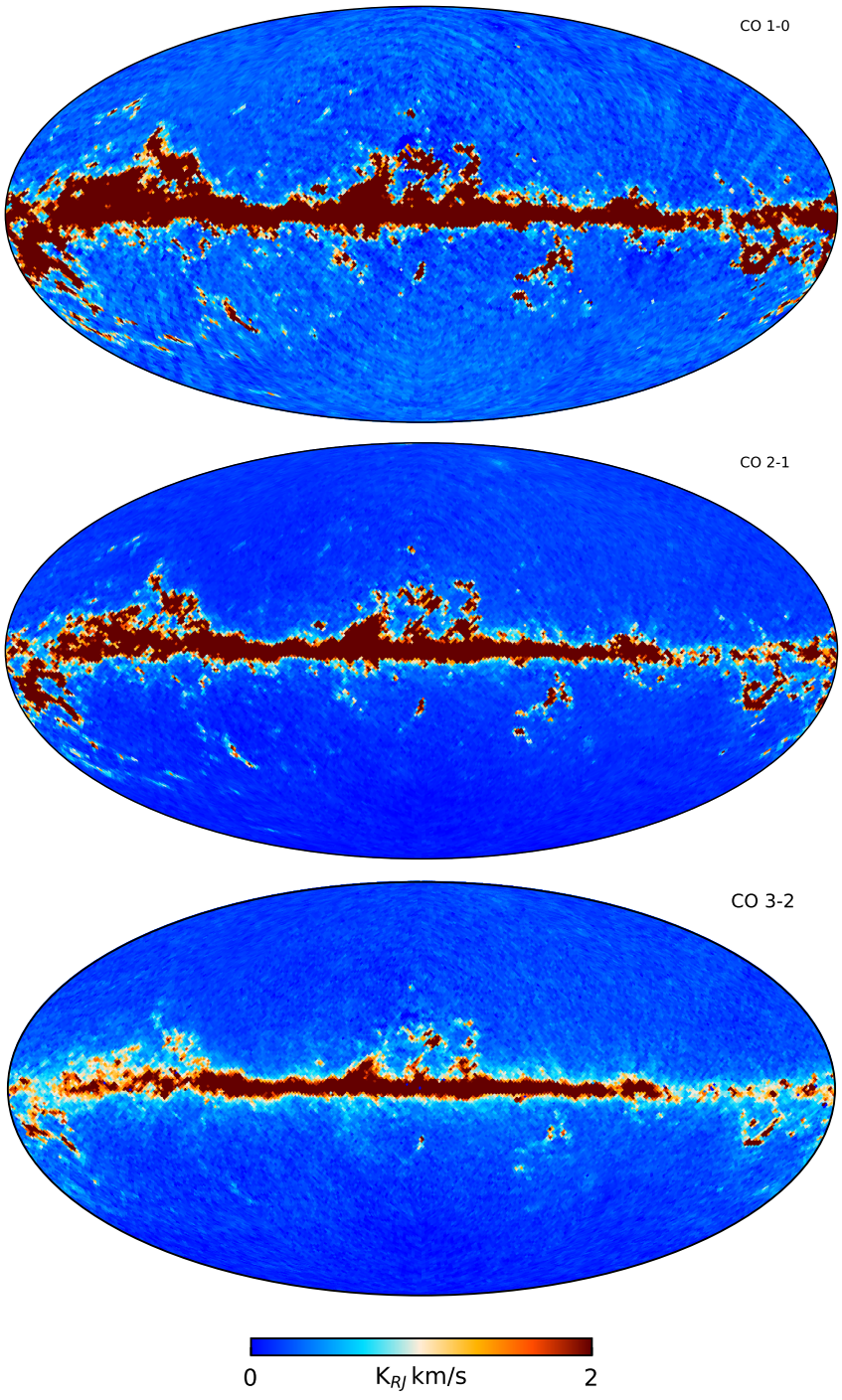


Figure 8.6: Milestone 6 maximum posterior CO components. We note lack of dust morphology compared to the CO 3 $\rightarrow$ 2 component in figure 8.4, which is a result of more data points.

Chapter 9

Conclusion and future prospects

"Predictions are difficult. Especially about the future."

— Yogi Berra

In this thesis, we have observed how the main issues with the CMB sky data now lie in the interplay between instrumental effects and foreground modeling. We have applied a new set of sky maps and attempted to apply component separation in an effort to reveal systematic effects subsequently corrected for in map-making. While we have reached the end of the road for this thesis, and the results appear very promising, there are still improvements to be made. However, the future of NPIPE maps looks bright.

9.1 Conclusion

Through comprehensive analysis, we have provided a set of Planck sky maps that exhibit a dramatic decrease in systematic errors as compared to the official Planck 2015 analysis. Our new set of sky maps allows for a better sky model fit, resulting in a significant χ^2 value improvement. When applying fewer sky maps to our model, we see that the low-frequency component serves as a good fit for low-frequency degenerate foregrounds. To fully remove the residual signal, however, we require a refined model. Furthermore, our full analysis as presented in milestone 6, employs a great number of newly cleaned, high-frequency maps which seemingly results in an improved thermal dust fit. While the introduction of more high-frequency sky maps negatively affects the goodness-of-fit value, we observe significant improvement over the similar analysis of the previous milestone, which applies an old set of maps. Overall, component maps describe a physical sky, to the extent our simplified model is able. Furthermore, our resulting CO maps show high correlation with the community standard of Dame et al. [66]. Although the iterative NPIPE map-making and component separation process has seemingly converged, there is still room for improvement. An ideal result for component separation would be residual maps void of any systematic and foreground morphology, but as previously mentioned, the LFI maps still exhibit clear signs of free-free emission

due to a flawed approximation. There are also problems directly related to our component separation efforts. As mentioned in the last chapter, a mistake was made during the calibration of the high-frequency detector maps probing thermal dust. Because of the frequency nature of these maps, bandpass and gain calibration must be carried out separately. While the likely ramifications of these gain calibration are small, it would be beneficial to carry out a proper monopole, bandpass and gain calibration. As a result of incomplete analysis of DX11 and DX12, we do not have sufficient evidence to measure the extent of NPIPE improvement directly. We do, however, note that systematic effects established in previous data releases have now been corrected. Although the component maps are insufficient to serve as a comparison set, the analysis served as a useful playground for familiarizing oneself with the component separation framework.

9.2 The future of component separation

A natural next step in the component separation of NPIPE and in general is adding further complexity to the sky model. For such a model to be well constrained, additional data is critical. From this thesis, we have learned how the χ^2 map describes the model fit for our system, however, if we increase the data volume, the model fit internally becomes worse. This is mostly an effect of a sub-optimal sky model. However, with our data, we are constrained to a simplified version of the sky due to degeneracies between certain types of emission. Specifically, we face the challenge of decomposing the low-frequency sky signals, as well as modeling the complexity of thermal dust radiation.

9.2.1 Resolving low-frequency degenerate foregrounds

We previously discussed how the low-frequency component is an approximation to an ensemble of separate emission components. The reason for the approximation is that their associated frequency spectra are profoundly degenerate. This simplification has now come back to haunt us, as free-free morphology is appearing in all LFI residuals as illustrated in figure 9.1. There is, however, a solution to this. With the inclusion of more data points, we may break these degeneracies. For example, by probing synchrotron and free-free emission at low frequencies, and spinning dust around the 30 GHz, we will be able to disentangle them. To break these degeneracies, we may employ independently derived data sets such as C-BASS[41], S-PASS[42] and QUIJOTE [43], measuring at 5 GHz, 2.3 GHz, and 10-20 GHz respectively. In addition to these, there are presently available data from the Haslam survey from 1982 [67], a 408 MHz measurement of the full sky, as well as all the WMAP frequency bands from 23-94 GHz. The application of independent data sets is not foreign to the CMB component separation community and was a vital part of the 2015 Planck release, in which both the Haslam and WMAP data were included. It would be quite interesting to redo this type of analysis in this improved set of sky maps, as the method yielded positive results in regards to separating the low-frequency component. In an ideal world, we would employ all possible datasets, and get a better estimate of the frequency spectrum of synchrotron, free-free, and spinning dust radiation, as well as a probe of the Sunyaev-Zel'dovich effect.

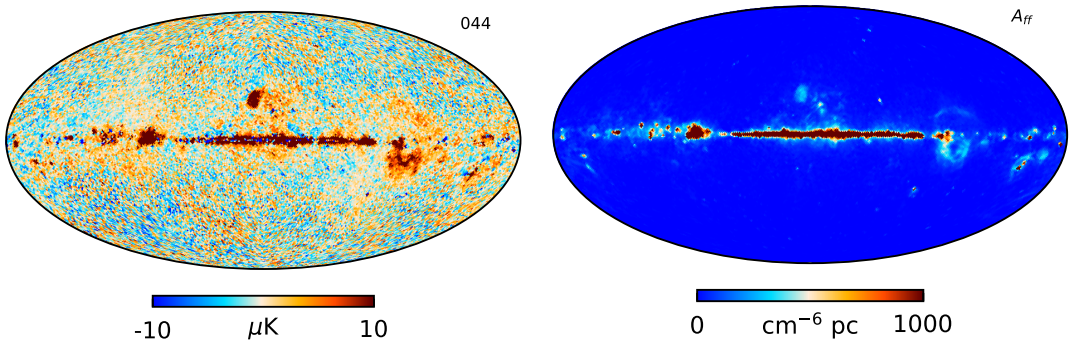


Figure 9.1: The map on the left shows the 44 GHz LFI band residual map exhibiting clear signs of free-free contamination, compared to a reference map of free-free shown on the right.

9.2.2 Decomposing thermal dust

On our hunt to clean the CMB map, we are recurrently plagued by either blue or red line stretching across the galactic plane. This aberration is most likely due to a bad model fit of thermal dust. As discussed in section 3.1.2, our thermal dust modified blackbody curve encapsulates the frequency spectrum of both silicate and carbonaceous dust grains, as well as CIB. Ideally, we would split these and model them as three separate components, this is, however, proving hard to accomplish. With the frequency spectra of each component being extremely similar, we are unable to tell them apart with the available data sets. However, through the use of external data such as DIRBE [68], FIRAS [69], WISE [70] and Hershel [71], we may probe the frequency spectrum at between 60 GHz and 2,4 THz, pinpointing varieties of thermal dust, CIB and even zodiacal light.

9.2.3 Time-domain Gibbs sampling

New improvements to component separation do not only come in the form of additional observational data. Our analysis tools are still evolving, and one of the most promising tools is *time-domain Gibbs sampling*. In the future, we hope to construct a new version of `Commander`, where component separation is done directly in the TOD. In a nutshell, this means implementing map-making into `Commander`. Such an algorithm would render the work carried out in this thesis redundant, as there would be no need for manual iteration between map-making and component separation; everything is done within a single code. Analysis directly on the TOD is necessary in order to uncover the weak signals of the elusive B-modes in polarization analysis, shrouded by systematic effects and foreground contamination.

Part IV

Appendix

Bibliography

[1] HK Eriksen, JB Jewell, C Dickinson, AJ Banday, KM Górski, and CR Lawrence. Joint bayesian component separation and cmb power spectrum estimation. *The Astrophysical Journal*, 676(1):10, 2008.

[2] Arno A Penzias and Robert Woodrow Wilson. A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, 142:419–421, 1965.

[3] Planck Collaboration. Planck 2015 results-xiii. cosmological parameters. *Astronomy & Astrophysics*, 594:A13, 2016.

[4] Scott Dodelson. *Modern cosmology*. Academic press, 2003.

[5] Sean M Carroll. *Spacetime and geometry. An introduction to general relativity*. Pearson Education Limited, 2014.

[6] Isaac Newton. *Philosophiae naturalis principia mathematica*. Londini, Jussu Societatis Regiae ac Typis Josephi Streater. Prostat apud plures Bibliopolas., 1687.

[7] Albert Einstein. Die grundlage der allgemeinen relativitätstheorie. *Annalen der Physik*, 354(7):769–822, 1916.

[8] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, 1929.

[9] Rhodri Evans. The cosmic microwave background. In *The Cosmic Microwave Background*, pages 55–89. Springer, 2015.

[10] Georges Lemaître. Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. In *Annales de la Société scientifique de Bruxelles*, volume 47, pages 49–59, 1927.

[11] Georges Lemaître. The beginning of the world from the point of view of quantum theory. *Nature*, 127(3210):706, 1931.

[12] Hermann Bondi and Thomas Gold. The steady-state theory of the expanding universe. *Monthly Notices of the Royal Astronomical Society*, 108(3):252–270, 1948.

[13] Fred Hoyle. A new model for the expanding universe. *Monthly Notices of the Royal Astronomical Society*, 108:372, 1948.

[14] Ralph A Alpher, Hans Bethe, and George Gamow. The origin of chemical elements. *Physical Review*, 73(7):803, 1948.

[15] Ralph A Alpher and Robert C Herman. Remarks on the evolution of the expanding universe. *Physical Review*, 75(7):1089, 1949.

[16] Andrew McKellar. Molecular lines from the lowest states of diatomic molecules composed of atoms probably present in interstellar space. *Publications of the Dominion Astrophysical Observatory Victoria*, 7, 1941.

[17] DW Sciama. Peculiar velocity of the sun and the cosmic microwave background. *Physical Review Letters*, 18(24):1065, 1967.

[18] MV Gorenstein, GF Smoot, and RA Muller. Anisotropy of the cosmic microwave background radiation. In *Bulletin of the American Astronomical Society*, volume 9, page 431, 1977.

[19] Fritz Zwicky. Die rotverschiebung von extragalaktischen nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.

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

[21] John C Mather, ES Cheng, RE Eplee Jr, RB Isaacman, SS Meyer, RA Shafer, R Weiss, EL Wright, CL Bennett, NW Boggess, et al. A preliminary measurement of the cosmic microwave background spectrum by the cosmic background explorer (cobe) satellite. *The Astrophysical Journal*, 354:L37–L40, 1990.

[22] NASA. Legacy Archive for Microwave Background Data Analysis (LAMBDA). `https://lambda.gsfc.nasa.gov/`. Accessed: 2018-05-25.

[23] Lawrence Berkeley Laboratories. COBE DMR Images of CMB. `http://aether.lbl.gov/www/projects/cobe/COBE_Home/DMR_Images.html`. Accessed: 2018-05-19.

[24] Mark J Devlin, Angelica de Oliveira-Costa, Tom Herbig, Amber D Miller, C Barth Netterfield, Lyman A Page, and Max Tegmark. Mapping the cosmic microwave background anisotropy: The first flight of the qmap experiment. *The Astrophysical Journal Letters*, 509(2):L69, 1998.

[25] Carlo R Contaldi, JR Bond, D Pogosyan, BS Mason, ST Myers, TJ Pearson, UL Pen, S Prunet, AC Readhead, MI Ruetalo, et al. Cmb observations with the cosmic background imager (cbi) interferometer. *arXiv preprint astro-ph/0210303*, 2002.

[26] Tom Montroy, Peter AR Ade, A Balbi, JJ Bock, JR Bond, J Borrill, A Boscaleri, P Cabella, CR Contaldi, BP Crill, et al. Measuring cmb polarization with boomerang. *New Astronomy Reviews*, 47(11-12):1057–1065, 2003.

[27] NW Halverson, EM Leitch, C Pryke, J Kovac, JE Carlstrom, WL Holzapfel, M Dragovan, JK Cartwright, BS Mason, S Padin, et al. Degree angular scale interferometer first results: a measurement of the cosmic microwave background angular power spectrum. *The Astrophysical Journal*, 568(1):38, 2002.

[28] Pea de Bernardis, Peter AR Ade, JJ Bock, JR Bond, J Borrill, A Boscaleri, K Coble, BP Crill, G De Gasperis, PC Farese, et al. A flat universe from high-resolution maps of the cosmic microwave background radiation. *Nature*, 404(6781):955, 2000.

[29] Adam G Riess, Alexei V Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M Garnavich, Ron L Gilliland, Craig J Hogan, Saurabh Jha, Robert P Kirshner, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009, 1998.

[30] Gary Hinshaw, D Larson, Eiichiro Komatsu, DN Spergel, CLaa Bennett, J Dunkley, MR Nolta, M Halpern, RS Hill, N Odegard, et al. Nine-year wilkinson microwave anisotropy probe (wmap) observations: cosmological parameter results. *The Astrophysical Journal Supplement Series*, 208(2):19, 2013.

[31] NASA/JPL-Caltech/ESA. PIA16874: The Universe Comes into Sharper Focus. `https://photojournal.jpl.nasa.gov/catalog/PIA16874`, 2013. Accessed: 2018-04-17.

[32] Benjamin P Abbott, Richard Abbott, TD Abbott, MR Abernathy, Fausto Acernese, Kendall Ackley, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikari, et al. Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102, 2016.

[33] Peter AR Ade, N Aghanim, Z Ahmed, RW Aikin, Kate Denham Alexander, M Arnaud, J Aumont, C Baccigalupi, Anthony J Banday, D Barkats, et al. Joint analysis of bicep2/keck array and planck data. *Physical review letters*, 114(10):101301, 2015.

[34] Peter AR Ade, Randol W Aikin, Denis Barkats, Steven J Benton, Colin A Bischoff, Jamie J Bock, Justus A Brevik, Immanuel Buder, Eric Bullock, C Darren Dowell, et al. Detection of b-mode polarization at degree angular scales by bicep2. *Physical Review Letters*, 112(24):241101, 2014.

[35] S&T: Leah Tiscione. "Proof of Inflationary Universe" to be Announced - Sky Telescope. `http://www.skyandtelescope.com/astronomy-news/proof-of-inflationary-universe-to-be-announced-today/`. Accessed: 2018-04-17.

[36] T Matsumura, Y Akiba, J Borrill, Y Chinone, M Dobbs, H Fuke, A Ghribi, M Hasegawa, K Hattori, M Hattori, et al. Mission design of litebird. *Journal of Low Temperature Physics*, 176(5-6):733–740, 2014.

[37] JP Filippini, Peter AR Ade, M Amiri, SJ Benton, R Bihary, JJ Bock, JR Bond, JA Bonetti, SA Bryan, B Burger, et al. Spider: a balloon-borne cmb polarimeter for large angular scales. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy V*, volume 7741, page 77411N. International Society for Optics and Photonics, 2010.

[38] S Aiola, G Amico, P Battaglia, E Battistelli, A Bau, P De Bernardis, M Bersanelli, A Boscaleri, F Cavaliere, A Coppolecchia, et al. The large-scale polarization explorer (lspe). In *Ground-Based and Airborne Instrumentation for Astronomy IV*, volume 8446, page 84467A. International Society for Optics and Photonics, 2012.

[39] E Battistelli, A Baú, D Bennett, L Bergé, J-Ph Bernard, P De Bernardis, G Bordier, A Bounab, É Bréelle, EF Bunn, et al. Qubic: The qu bolometric interferometer for cosmology. *Astroparticle Physics*, 34(9):705–716, 2011.

[40] Kevork N Abazajian, Peter Adshead, Zeeshan Ahmed, Steven W Allen, David Alonso, Kam S Arnold, Carlo Baccigalupi, James G Bartlett, Nicholas Battaglia, Bradford A Benson, et al. Cmb-s4 science book. *arXiv preprint arXiv:1610.02743*, 2016.

[41] Oliver G King, Charles Copley, Rod Davies, Richard Davis, Clive Dickinson, Yaser A Hafez, Christian Holler, Jaya John John, Justin L Jonas, Michael E Jones, et al. The c-band all-sky survey: instrument design, status, and first-look data. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy V*, volume 7741, page 77411I. International Society for Optics and Photonics, 2010.

[42] Ettore Carretti, Bryan Gaensler, Lister Staveley-Smith, Marijke Haverkorn, Michael Kesteven, Stefano Cortiglioni, Gianni Bernardi, and Sergio Poppi. S-band polarization all sky survey (s-pass). *ATNF Proposal*, 2007.

[43] F Poidevin, JA Rubino-Martin, R Genova-Santos, R Rebolo, M Aguiar, F Gomez-Renasco, C Gutierrez, RJ Hoyland, C Lopez-Caraballo, A Oria Carreras, et al. The quijote experiment: Prospects for cmb b-mode polarization detection and foregrounds characterization. *arXiv preprint arXiv:1802.04594*, 2018.

[44] Giuseppe D Racca, René Laureijs, Luca Stagnaro, Jean-Christophe Salvignol, José Lorenzo Alvarez, Gonzalo Saavedra Criado, Luis Gaspar Venancio, Alex Short, Paolo Strada, Tobias Bönke, et al. The euclid mission design. In *Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave*, volume 9904, page 99040O. International Society for Optics and Photonics, 2016.

[45] Planck Collaboration. Planck 2015 results-i. overview of products and scientific results. *Astronomy & Astrophysics*, 594:A1, 2016.

[46] Planck Collaboration. Planck 2015 results-x. diffuse component separation: Foreground maps. *Astronomy & Astrophysics*, 594:A10, 2016.

[47] Planck Collaboration. Planck 2015 results-vii. high frequency instrument data processing: Time-ordered information and beams. *Astronomy & Astrophysics*, 594:A7, 2016.

[48] Mathieu Remazeilles et al. Planck intermediate results-xlviii. disentangling galactic dust emission and cosmic infrared background anisotropies. *Astronomy & Astrophysics*, 596:A109, 2016.

[49] C Dickinson, RD Davies, and RJ Davis. Towards a free–free template for cmb foregrounds. *Monthly Notices of the Royal Astronomical Society*, 341(2):369–384, 2003.

[50] EM Leitch, ACS Readhead, TJ Pearson, and ST Myers. An anomalous component of galactic emission. *The Astrophysical Journal Letters*, 486(1):L23, 1997.

[51] Planck Collaboration. Planck 2013 results. ix. hfi spectral response. *Astronomy & Astrophysics*, 571:A9, 2014.

[52] ESA/AOES Medialab. Planck Focal Plane. `http://sci.esa.int/planck/40389-planck-focal-plane/`. Accessed: 2018-04-03.

[53] Planck Collaboration. Planck pre-launch status: High frequency instrument polarization calibration. *Astronomy & Astrophysics*, 520:A13, 2010.

[54] G Hinshaw, MR Nolta, CL Bennett, R Bean, O Doré, MR Greason, M Halpern, RS Hill, N Jarosik, A Kogut, et al. Three-year wilkinson microwave anisotropy probe (wmap*) observations: Temperature analysis. *The Astrophysical Journal Supplement Series*, 170(2):288, 2007.

[55] Krzysztof M Gorski, Eric Hivon, AJ Banday, Benjamin D Wandelt, Frode K Hansen, Mstvos Reinecke, and Matthia Bartelmann. Healpix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759, 2005.

[56] Krzysztof M Górski, Eric Hivon, and Benjamin D Wandelt. Analysis issues for large cmb data sets. *arXiv preprint astro-ph/9812350*, 1998.

[57] Planck Collaboration. Planck 2015 results: Viii. high frequency instrument data processing: Calibration and maps. *Astronomy & Astrophysics*, 594:A8, September 2016.

[58] Planck Collaboration. Planck intermediate results-xlvi. reduction of large-scale systematic effects in hfi polarization maps and estimation of the reionization optical depth. *Astronomy & Astrophysics*, 596:A107, 2016.

[59] Reijo Keskitalo. Map-Making and the details of NPIPE. Personal communication.

[60] J Jewell, S Levin, and CH Anderson. Application of monte carlo algorithms to the bayesian analysis of the cosmic microwave background. *The Astrophysical Journal*, 609(1):1, 2004.

[61] Benjamin D Wandelt, David L Larson, and Arun Lakshminarayanan. Global, exact cosmic microwave background data analysis using gibbs sampling. *Physical Review D*, 70(8):083511, 2004.

[62] Jonathan Richard Shewchuk et al. An introduction to the conjugate gradient method without the agonizing pain, 1994.

[63] M Chu, HK Eriksen, L Knox, KM Górski, JB Jewell, DL Larson, IJ O'Dwyer, and BD Wandelt. Cosmological parameter constraints as derived from the wilkinson microwave anisotropy probe data via gibbs sampling and the blackwell-rao estimator. *Physical Review D*, 71(10):103002, 2005.

[64] Alessio Notari and Miguel Quartin. On the proper kinetic quadrupole cmb removal and the quadrupole anomalies. *Journal of Cosmology and Astroparticle Physics*, 2015(06):047, 2015.

[65] Planck Collaboration. Planck 2013 results. xiv. zodiacal emission. *Astronomy & Astrophysics*, 2014.

[66] Thomas M Dame, Dap Hartmann, and P Thaddeus. The milky way in molecular clouds: a new complete co survey. *The Astrophysical Journal*, 547(2):792, 2001.

[67] CGT Haslam, CJ Salter, H Stoffel, and WEz Wilson. A 408 mhz all-sky continuum survey. ii-the atlas of contour maps. *Astronomy and Astrophysics Supplement Series*, 47:1, 1982.

[68] Michael G Hauser, Richard G Arendt, T Kelsall, E Dwek, N Odegard, Janet L Weiland, Henry T Freudenreich, William T Reach, Robert F Silverberg, Samuel Harvey Moseley, et al. The cobe diffuse infrared background experiment search for the cosmic infrared background. i. limits and detections. *The Astrophysical Journal*, 508(1):25, 1998.

[69] John C Mather, ES Cheng, David A Cottingham, RE Eplee Jr, Dale J Fixsen, Tilak Hewagama, RB Isaacman, KA Jensen, Stephan S Meyer, Peter D Noerdlinger, et al. Measurement of the cosmic microwave background spectrum by the cobe firas instrument. *The Astrophysical Journal*, 420:439–444, 1994.

[70] Edward L Wright, Peter RM Eisenhardt, Amy K Mainzer, Michael E Ressler, Roc M Cutri, Thomas Jarrett, J Davy Kirkpatrick, Deborah Padgett, Robert S McMillan, Michael Skrutskie, et al. The wide-field infrared survey explorer (wise): mission description and initial on-orbit performance. *The Astronomical Journal*, 140(6):1868, 2010.

[71] SJ Oliver, J Bock, B Altieri, A Amblard, V Arumugam, H Aussel, T Babbedge, A Beelen, M Béthermin, A Blain, et al. The herschel multi-tiered extragalactic survey: Hermes. *Monthly Notices of the Royal Astronomical Society*, 424(3):1614–1635, 2012.