UiO : University of Oslo

Harald Thommesen

Observing the CMB Sky with GreenPol, SPIDER and Planck

Thesis submitted for the degree of Philosophiae Doctor

Institute of Theoretical Astrophysics
Faculty of Mathematics and Natural Science
University of Oslo
Oslo, Norway
June 2019

2019

Series of dissertations submitted to the
Faculty of Mathematics and Natural Science
University of Oslo
No. 2164

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

An Kristian, Andrea und Halger

Acknowledgements

This thesis and all the work described within would not have been possible without the help of my fantastic supervisors, Ingunn and Hans Kristian. Thank you for all the support you gave me, your patience, and the motivation. Not only am I grateful for the past, but also for the future, and I'm looking forward to the time ahead.

I started my time as a PhD candidate together with Marie and Håvard. Those poor fellows didn't get a position within the field of CMB, but had to settle with CO emission. I'm so sorry for you. Thank you for the time together and I wish you all the best for the future! Another person that was there right from the beginning was Mattia. You were a great companion for mastering Cosmology II and Ethics, and I'm happy I don't have to do it again. I also want to thank Ben, Max, and Lluis for creating a social atmosphere including hikes and drinks that made settling in Oslo a lot easier and in general very welcoming.

The institute would not be such an enjoyable place without all my team members Ata, Ragnhild, Ranajoy, Kristian, Trygve, Unni, Maksym, Matthew and Eirik. A special thank you goes out to Monica for all her tips and helpful comments, regarding this thesis and otherwise.

During my work at the institute I got the exciting opportunity to participate in a field trip to Greenland to deploy the `GreenPol` telescope. I want to thank Ari Kaplan, Nic Rupert, and Peter Meinhold for making this possible and accepting me in their team. Thank you for the good laughs and the hard work that made this adventure such an unforgettable experience. I also greatly appreciate the help I got for the dedicated chapter in this thesis.

Lastly I want thank my family and my friends at home for their endless support and encouragement. My biggest personal thank you goes to my brother and my parents for always being there for me.

Contents

Prelude

The topic of the thesis in your hands is the *Cosmic Microwave Background Radiation*, for convenience simply abbreviated as the "CMB". Most people outside of physics, or outside of science in general, are for obvious reasons completely unaware of the existence of this radiation. When curious friends or family members make the mistake of asking me what I am working on in Oslo, my answer usually leaves puzzled expressions on their faces. My standard approach to make the CMB more tangible for laymen is trendy popular science explanations like "the CMB is the relic radiation from the Big Bang and it makes up about 1 % of the static fuzz in a vintage television set". Explanations like these usually only make matters worse. In Part I of this thesis, I therefore want to take the opportunity to demystify the CMB by giving an easy to understand introduction. This will lead us from the prediction of the CMB in the early 20th century to its accidental discovery in 1964, and later on to the experimental endeavors that were necessary to measure its temperature and several other properties. The CMB can tell us a lot about the history of our universe, but it is by far not the only microwave component out there, so I will also introduce the techniques required to disentangle this precious signal from a multitude of contaminators. These textbook-like chapters will make it clear why the CMB is a backbone of modern cosmology and why globe-spanning teams of scientists spend years and fortunes on building instruments and analyzing the data.

I will use Part II of this thesis to tell the story of my contribution to the scientific community. Contrary to popular opinion, the daily life of a cosmologist does not (only) comprise of reading textbooks and staring at the sky through a telescope, but rather requires computational expertise, like developing complicated computer codes that wrestle terabytes of observational data into the right shape. During my time at the *Institute of Theoretical Astrophysics* in Oslo, I have participated in three CMB experiments: `GreenPol`, `SPIDER` and `Planck`. These experiments differ in a great number of ways, such as their available budget, their deployment location, the used instrumentation, or the observed frequencies, just to name a few. What they do have in common, however, is their rough underlying concept: Map the intensity of the CMB on the sky and infer cosmological parameters from it. This task can be split into several logical components, which as a whole is called a *pipeline*. The components of the pipeline are typically the same for all CMB experiments. Below I list the most important ones:

1. Data acquisition: The process of scanning the sky with the instrument, typically resulting in terabytes of time-ordered data points (timestreams)

2. Data cleaning: The process of removing contamination caused by the instrument from the data

3. Mapmaking: The process of converting the timestreams of the scan to a map on the sky

4. Component separation: The process of disentangling the CMB from contaminating radiation from the Milky Way
5. Cosmological parameter estimation: Infer cosmological parameters from a clean CMB map.

The list above is far from complete, but it demonstrates the structuredness of the workflow in CMB analysis. During my time as a PhD candidate, I got the opportunity to work on several CMB experiments, which I have joined at different stages of their lifetimes. This had the effect that I worked simultaneously on many different parts of the generic pipeline, albeit related to different CMB missions. I will use my gained experience to present all components of a typical CMB analysis pipeline, and illustrate them by showing real-life examples. I will start out in Chapter 4 by describing my expedition to Greenland, where I participated in the data acquisition process with `GreenPol`. In Chapter 5 I will explain the process of cleaning the data and mapping them onto the celestial sphere, by using `SPIDER` as an example. Finally I will show in Chapter 6, by the example of `Planck`, how to analyze sky maps and gain cosmological information from them.

Part I

Background

Chapter 1

The Cosmic Microwave Background

If you were an astronaut in outer space and were able to observe light at millimeter wavelengths instead of visible light in the nanometer range, you would be surprised by an unusual sight: Instead of gazing at a myriad of stars as you are used to, you would be bathed from all sides in extremely uniform radiation that at first glance doesn't show any structure at all. This uniform glow of light is known as the *Cosmic Microwave Background Radiation* (CMB) and is the main topic of this thesis.

In this Chapter I will guide you through the basic building blocks that you need to know about CMB cosmology. This tour will take you from its underlying theory, to high precision experiments and won't stop before learning about its mathematical description. Here is a little teaser of what awaits you:

The CMB radiation is a relic stemming from a time when the Universe was an extremely dense and hot environment, similar to the surface of the Sun. It was released when the Universe was barely some 300 000 years old. Compare this to the age of the Universe today, which is about 14 billion years, and you'll realize how old this remnant really is. In fact, it is the oldest light we can ever observe. When observing the CMB, we see the Universe at an extremely early stage, long before any stars or galaxies had time to form, which makes it such an important discovery for the understanding of our origins. The theory that describes how the Universe evolved from this dense and hot environment into the structures that we see today is called the *Hot Big Bang model* and we will learn more about it later on in Section 1.1. Nowadays, the CMB is considered as the most powerful evidence for the Hot Big Bang theory. Even though the CMB was theoretically predicted as early as 1948, it wasn't experimentally observed until 1964. During the 50+ years that followed, several dozens of high precision CMB experiments ventured out to measure the CMB on the sky, which has lead to an explosion in cosmological data. I will give a more detailed account on some of the most important experiments in Section 1.2. A mathematical tool to analyze the CMB sky is the *CMB power spectrum.* This mathematical representation contains all the statistical information that we need to estimate cosmological parameters from the measured maps. The power spectrum is a key feature for the understanding of CMB analysis and I will give a detailed description of the mathematical theory behind it in Section 1.3.

The discovery of the CMB has completely transformed the understanding of our Universe. Prior to the discovery of the CMB, the scientific community was in conflict on whether our space is static or evolving with time. The CMB has tipped the scales and today the Big Bang model is universally accepted. I hope that this teaser has awakened your interest in the CMB and I promise it's worth while reading on.

1.1 The Hot Big Bang Scenario

At the beginning of the 20th century, cosmology was a rather small and neglected branch of physics, almost bordering on philosophy. The reason for this was the lack of observational data that made most theoretical ideas untestable. The majority of people were in the belief that we live in an unchanging Universe that has an infinite age and lifespan. This notion of an ever-present and unchanging Universe is called the *Perfect Cosmological Principle*, which states that "the properties of the Universe are the same in all of space and time". This picture started to change when Edwin Hubble discovered a linear relationship between the distance to galaxies and their radial velocities. In a paper in 1929 [1], he describes his work at Mount Wilson Observatory, where he measured the distance to "extra-galactic nebulae"[1] using Cepheid variables and correlated them to measurements of their radial velocities. Surprisingly, it turned out that the recession speed v of galaxies was not random but increased linearly with the distance D to Earth. This is today known as the famous *Hubble Law*,

$$v = H_0 D, \tag{1.1}$$

where H_0 is the Hubble constant. The realization that the Universe is expanding had far-reaching consequences and split the scientific community into two groups. One side was supporters of the so called "Big Bang model", which is based on the idea that since the Universe has been expanding for a long time, all the matter must have been infinitely dense in the very early past. The second group was favoring the "Steady State model" [2], that persists on an unchanging Universe by assuming a mechanism that constantly fills up the expanding space by creation of matter. This debate was going on for almost two decades, until another piece of the puzzle fell into place: The origin of helium.

In the late 40s, the production of helium in the interior of stars through nuclear fusion of hydrogen atoms was well-understood. By comparing spectroscopic observations with calculations however, it turned out that the production of helium in stars could only explain a small fraction of the amount that was actually observed. This excess gave birth to the idea that the majority of helium wasn't created inside stars, but rather "baked" shortly after the Big Bang when the Universe was still extremely dense and billions of degrees hot. In 1948, physicist Roger Gamow along with his student Ralph Alpher studied the nuclear processes that were expected to have happened in the early Universe and published a paper explaining the origin of the chemical elements, where they were

---

[1] These turned out to be galaxies.

able to successfully reproduce the abundance of helium [3]. This was another point in favor of the Big Bang model. In the same year, Ralph Alpher together with Robert Herman predicted that the relic radiation from the hot early Universe should still be around today and have a temperature of approximately 5 K [4]. It took almost 20 years until the Steady State theory of the Universe received its final blow. In 1964, the radio astronomers Arno Penzias and Robert Wilson were struggling with a low-level noise in their radio antenna. Mysteriously, this noise with a temperature of roughly 3.5 K was slightly stronger when pointing towards zenith than towards the ground, and in addition free from seasonal variations [5]. They were put in touch with Robert Dicke, a scientist from Princeton University, who also came to the conclusion that the glow from the Big Bang should still be detectable in the form of blackbody radiation in the microwave range [6]. Later on, Penzias and Wilson were rewarded the Nobel Prize in Physics for the discovery of the cosmic microwave background radiation.

The discovery of the expansion of the Universe together with the discovery of the CMB and the explanation for the abundance of helium are today considered as the most important evidence in favor of the Big Bang. In the next Chapter, we will dive into the hot and dense early universe and investigate the physical conditions that resulted in the emission of the cosmic microwave background. These are crucial to appreciate the significance of the CMB for cosmology.

1.1.1 The hot early Universe

The CMB photons that Wilson and Penzias discovered in 1964 are extremely numerous, in fact far more numerous than the photons we receive from stars and from the Sun. However, their typical wavelength lies in the millimeter range, so despite their superior number, the CMB doesn't carry a lot of energy. Combining this fact with the knowledge that the Universe is expanding leads to the conclusion that these photons must have been far more energetic in early times, since their wavelengths have been stretched by cosmic expansion. This increase of wavelength due to cosmic expansion is called *cosmological redshift* and is also the reason why far away galaxies appear redshifted. If we think backwards in time and reverse the course of cosmic expansion and structure formation, all matter and radiation gets squished together to form what's colloquially called the "primordial fire ball". While matter simply gets condensed to a smaller volume of space, radiation gets both condensed and in addition more energetic by shrinking its wavelength. In order to understand the properties of this primordial fire ball, let's first discuss the most abundant element in the Universe, namely hydrogen.

The electron orbiting the hydrogen atom is relatively weakly bound and can easily be knocked out by ultraviolet radiation. At the very beginning of the Universe, hydrogen atoms were extremely short-lived, since they got ionized virtually instantly by high-energy photons. This state of matter, where electrons are separated from their nuclei, is called plasma and has two interesting properties. First, the mean free path of any photon is extremely short because of *Thomson scattering* off of free electrons. This has the effect that photons move in complicated zig-zag paths, constantly colliding with electrons, which makes the plasma *opaque*. Second, since photons, protons, and electrons are in constant interaction with each other, the matter and radiation field have the same temperature. Expressed in other words, the matter and radiation field are in *thermal equilibrium*, which means the radiation exhibits a *blackbody spectrum* that is *only* characterized by its temperature. As cosmic expansion continues, the peak wavelength λ_{max} of the blackbody spectrum grows proportionally, while its temperature drops. At some point in cosmic history, the energy stored in the photons drops below the ionization energy of hydrogen and the electrons recombine with the protons to form neutral hydrogen atoms. This transition from an ionized to a neutral Universe is called “era of recombination” and happens in a very confined amount of time. Neutral hydrogen atoms do not scatter photons, and the hitherto opaque space suddenly became transparent. From this moment on, photons could stream freely across the expanding space with (almost) no interactions at all. Since the primordial plasma was in thermal equilibrium, the released radiation features a blackbody spectrum. These photons constitute the cosmic microwave background today and have been traveling since the Universe was approximately 370 000 years old and about a thousandth of today’s size. The moment photons started to move freely is called the “epoch of photon decoupling”, acknowledging the fact that radiation and matter ceased to interact. This was the starting signal for structure formation because from that moment on, density perturbations in the matter field were no longer damped by the pressure exerted by radiation, and therefore could grow into the structures we see today.

1.2 Observational successes

Penzias and Wilson’s radio antenna was followed by a long series of experiments that attempted to map and examine the CMB further. Just to get an idea of the vast number of experiments performed over the years, the *list of cosmic microwave background experiments* on Wikipedia currently contains an astounding number of 58 entries [7]. While going through all of these missions is not the goal of this section, there are a few that are worth of further inspection because they stand out in terms of their impact on the understanding of the microwave sky. These “VIPs” among the CMB experiments are `COBE`, `WMAP` and `Planck`. They produced some of the most valuable cosmological data by observing with an unobstructed view from space.

The Cosmic Background Explorer (COBE) was launched by NASA in 1989 and operated until 1993. For more information on `COBE`, please refer to [8]. One of `COBE's` main results was achieved by FIRAS, the Far Infrared Absolute Spectrophotometer onboard the spacecraft. It mapped out the spectral behavior of the CMB and showed that it indeed exhibits a near-perfect blackbody spectrum

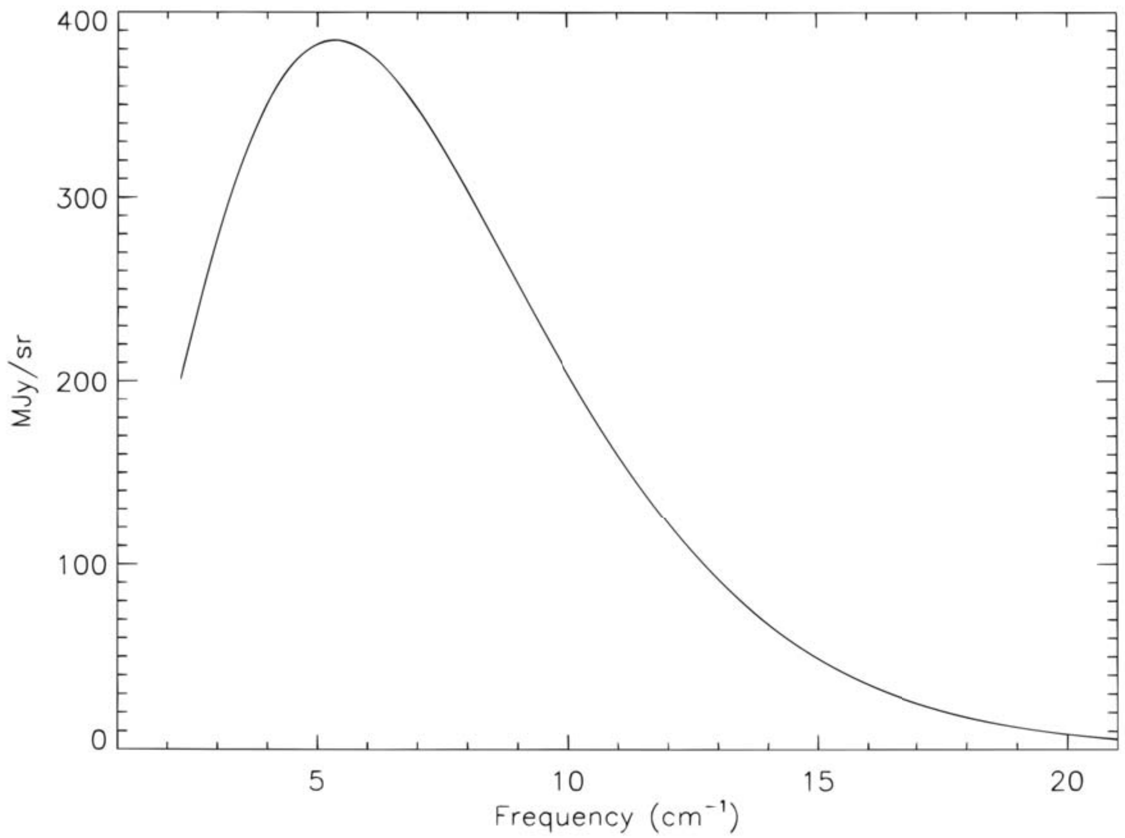


Figure 1.1: The spectrum of the CMB radiation as measured by `COBE`. The line represents a Planck blackbody curve fitted to the data points, which, along with their error bars, are hidden by the thickness of the line. Figure courtesy of [9].

with an approximate temperature of 2.7 K [9], as it was theoretically predicted. Figure 1.1 shows a Planck blackbody curve fitted to the measurement. The agreement between the measurement and the theoretical curve is striking, considering that the thickness of the curve is bigger than the uncertainties of the underlying data points. This result was again a huge success for the Big Bang theory.

Another important instrument onboard `COBE` was the Differential Microwave Radiometer (DMR), which scanned the sky for the duration of four years and delivered the first full-sky multi-frequency CMB maps [10]. These maps revealed that the CMB temperature is subject to tiny variations. According to the Big Bang theory, these temperature fluctuations correspond to minuscule density fluctuations in the primordial plasma, which collapsed and grew into the structures we see today. Their confirmation by `COBE` was therefore another huge success.

The CMB sky in temperature is only one side of a coin. Besides the temperature, one can also analyze the polarization of the CMB photons. The main goal

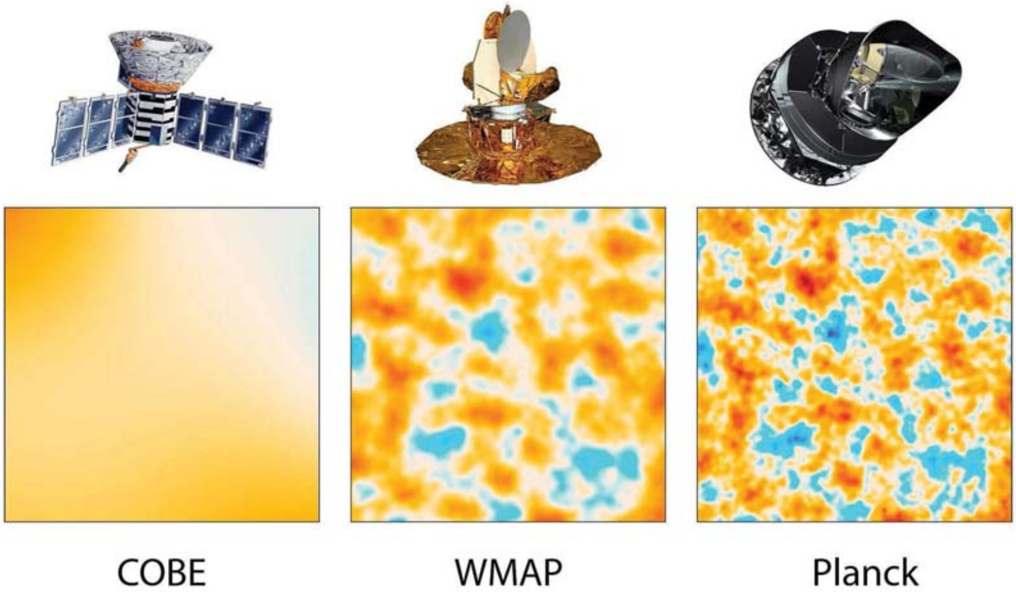


Figure 1.2: Improvement of resolution in CMB temperature maps. The three panels show the identical spot on the sky, as measured by `COBE` (1989-1993), `WMAP` (2001-2010) and `Planck` (2009-2013) [Image Credit: NASA/JPL-Caltech/ESA]

of today's CMB experiments is the search for a polarization pattern called *B-modes*, which if detected, would prove *inflation* (see Sec. 1.3.3). For this reason, `WMAP` and later on `Planck's` goals were not only to achieve better resolved maps in temperature, but in addition to do the same for polarization. Further, they complemented the frequency range of the measurements, which is necessary for the removal of galactic emission from the resulting sky maps (see Sec. 2 and 3 for details).

The Wilkinson Microwave Anisotropy Probe (WMAP) operated from 2001 until 2010 and observed the temperature and polarization of the CMB at five frequencies ranging from 23 to 94 GHz. Compared to its predecessor `COBE`, its sensitivity was 45 times greater, while its angular resolution was improved by a factor of 33 [11]. In 2009, `Planck` started observing at even higher resolution and delivered the most detailed full-sky maps to date. For a full overview of `Planck's` scientific successes, please refer to [12]. A more thorough introduction on the `Planck` instrument will be given in Sec. 6. Figure 1.2 illustrates the evolution of the CMB maps recorded in temperature over the years. The three panels show the identical spot on the sky, but observed by `COBE` on the left, `WMAP` in the middle and `Planck` on the right.

The reason why `COBE`, `WMAP` and `Planck` have such a high status among the CMB experiments is their ideal operating environment, namely in space. This puts them far away from artificial microwave sources and eludes atmospheric absorption. It is however possible to reduce the costs of CMB experiments by observing from balloons or even from high-altitude locations on the ground. Apart from the reduced cost, this also has the advantage of basically unlimited observation time and easier instrumental maintenance. The price for this is however a limited frequency range, higher atmospheric absorption and also a reduced observation area on the sky. Later on in this thesis I will present both a ground-based as well as a balloon-borne mission that I have been involved with, namely `GreenPol` in Sec. 4 and `SPIDER` in Sec. 5.

1.3 Statistical description of the CMB

The combined experimental efforts described in the previous section have lead to the understanding that the CMB sky consists of three main components. The first is the blackbody temperature of the CMB photons averaged across the celestial sphere. The current best estimate for this temperature is 2.725 K [13]. This constant term is also called the CMB *monopole* because of its correspondence to the lowest multipole in spherical harmonic space (see Sec. 1.3.1). As discovered by `COBE`, the CMB temperature is not completely constant across the sky, but subject to minute deviation from this constant blackbody term. The most dominant structure among these CMB fluctuations is the CMB *dipole*, which occurs because of the motion of the observer with respect to the last scattering surface. The underlying process that generates the dipole is therefore a Doppler boost, which blueshifts the CMB photons in the direction of motion, while redshifting the photons in the direction we are receding from. This has the effect that we observe a dipolar structure on the sky with an amplitude of about 3 mK [14]. While the main component of the CMB dipole is not an intrinsic CMB fluctuation, it is a very valuable tool for calibration of CMB experiments. For this reason, more details about the CMB dipole will be given in Sec. 6.3.

Subtracting the monopole and the dipole from a CMB map reveals the third and main component of the CMB, namely higher-order temperature fluctuations at the 100 μK level. After this long introduction, Figure 1.3 finally shows a complete picture of the CMB, as observed by the `Planck` instrument. These fluctuations are the main interest for cosmologists, and they are analyzed by means of their underlying source distribution, called the *CMB power spectrum*.

1.3.1 Decomposition of the CMB

When looking at Figure 1.3, the density fluctuations look extremely random to the eye. That's because the CMB, according to theory, *is* in fact random, namely an isotropic and Gaussian random field. The interest for the CMB of cosmologists therefore does not lie in the exact position of hot and cold spots on the celestial sphere, but only in the statistical properties of the size and amplitude of the structures we observe.

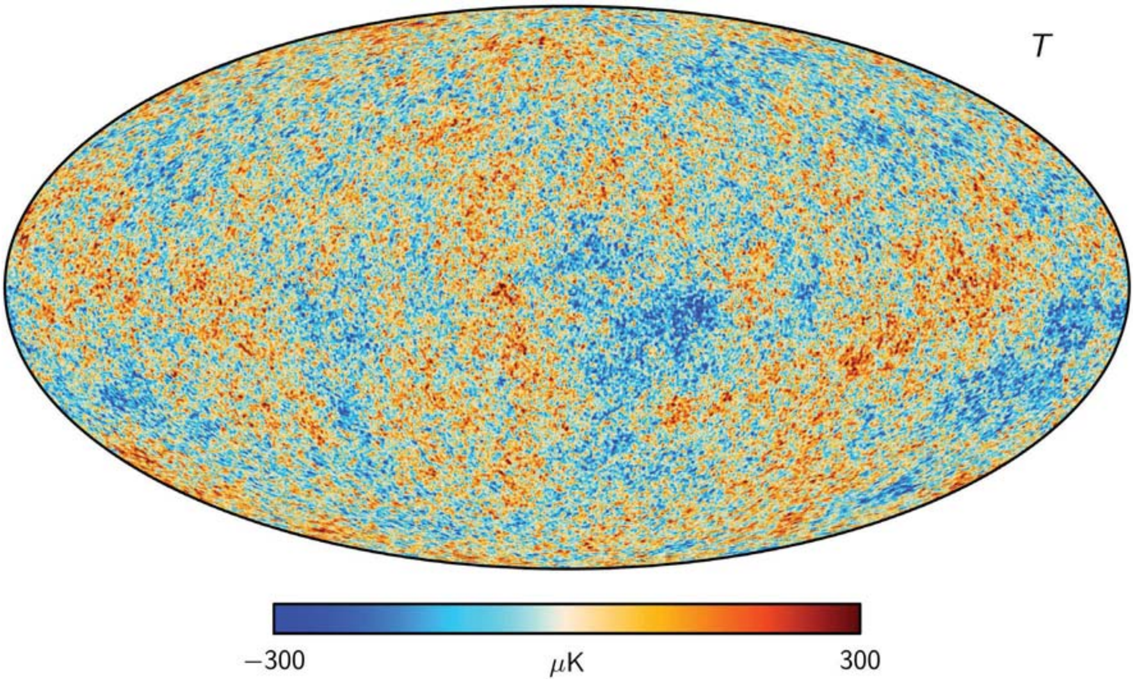


Figure 1.3: CMB higher-order fluctuations. Figure courtesy of [15].

Just as we can decompose a wave into its unique Fourier modes with corresponding weights (i.e. fourier transformation), we can treat the CMB sky $T(\hat{n})$ as a periodic function on the sphere and decompose it into the sum of spherical harmonic base functions $Y_{\ell m}$ with respective weights $a_{\ell m}$. The entire CMB sky can therefore be expressed in spherical harmonic space, which reads as

$$T(\hat{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n}) . \tag{1.2}$$

Here, ℓ is the *multipole*, which can be interpreted as the angular scale we are considering. The lowest multipole, $\ell = 0$, represents the monopole, $\ell = 1$ represents the dipole, and so forth. The higher the multipole, the smaller the scale on the sky, and therefore the higher is the resolution required to observe it. It is also important to note that for each multipole ℓ, we need $m = 2\ell + 1$ azimuthal, i.e. directional dependent modes, to fully describe it.

Using this expansion, we can identify the three main CMB components from Sec. 1.3 as

$$T(\hat{n}) = T_0 + D + \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} a_{lm} Y_{lm}(\hat{n}) , \tag{1.3}$$

where T_0 ($\ell = 0$) stands for the monopole, D ($\ell = 1$) for the dipole, and the sum ($\ell \geq 2$) are the higher-order fluctuations shown in Fig. 1.3. This breakdown makes sense, because the higher-order fluctuations happened at the *surface of last scattering*, and are therefore the signal we are interested in. While the

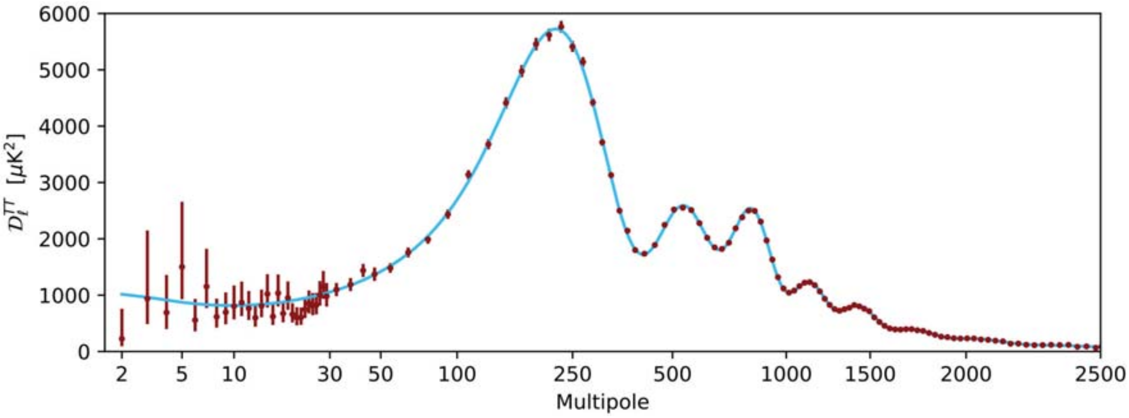


Figure 1.4: The CMB power spectrum in temperature as measured by `Planck`. The red dots represent the data, while the cyan line is the best-fit curve from theory. Figure courtesy of [12].

CMB might have an intrinsic dipole component, the dominating contribution is caused by *today's* motion, and therefore irrelevant in the context of cosmology. Lastly, the monopole can simply be treated as an offset in temperature, and is usually subtracted from the maps.

1.3.2 The temperature power spectrum

According to theory, the hot and cold structures in the CMB were sourced by quantum fluctuations in the early Universe, and are therefore homogeneous, isotropic, and Gaussian distributed. Thus, we are not interested in the exact position of these structures in the sky, but rather in the underlying statistics from which the observed sky is one single realization. Conveniently, all the statistical information of a CMB map (e.g. Fig. 1.3) is losslessly compressed in its power spectrum C_ℓ, which can be computed from the coefficients $a_{\ell m}$ of its spherical harmonic expansion (Eq. 1.2). The assumption of Gaussianity means that each coefficient $a_{\ell m}$ is a Gaussian deviate from the mean

$$\langle a_{\ell m} \rangle = 0, \tag{1.4}$$

with variance

$$C_\ell = \langle |a_{\ell m}|^2 \rangle. \tag{1.5}$$

The power spectrum C_ℓ is independent from the azimuthal mode m due to the assumption of isotropy. This allows us to average out the dependence on m and simply express it as a function of angular scale (i.e. multipole ℓ). Figure 1.4 shows the power spectrum published by the `Planck` collaboration [12]. The red dots represent the power spectrum calculated from the observed map, while the cyan line is the best-fit curve calculated from the theory. The agreement between the data and the theory is astonishing. The cause for the large error bars in Fig. 1.4 at large scales is of statistical nature. Since we measure the power at a particular scale ℓ by computing the average

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2, \tag{1.6}$$

we have to make do with a very small number of statistical samples when working in the low-ℓ regime. This means that the observed average will deviate from the true underlying average and therefore the precision of the measurement is limited by fundamental statistics. This limitation is called *cosmic variance*.

The theoretical curve in Fig. 1.4 is extremely sensitive to cosmological models and their parameters. The power spectrum therefore serves as the link between the observed Universe and the theory. It is outside the scope of this thesis to describe in detail what physical processes are responsible for the individual bumps in the spectrum. However, modeling different kinds of Universes, and comparing the computed and the observed power spectra allows us to test fundamental cosmological properties. For instance, the ℓ-position of the first peak of the spectrum is a probe for the spatial curvature and the total mass density of the Universe. The height ratio of the second and the third peak probe the baryon density, while the low-ℓ slope provides information about inflation.

1.3.3 The polarization of the CMB

So far we have only talked about the temperature, i.e. the energy distribution of the CMB. Photons however are not solely characterized by their energy, but also by their state of *polarization*. This hands us an alternative window through which we can observe the sky and test our models. Not only do we gain new insight into the CMB but it also allows us to break parameter degeneracies that emerge when being limited to temperature only. Let's first look into why we expect the CMB to be polarized in the first place.

The cause of polarization

Polarization describes the orientation of the electric and magnetic field vectors of an ensemble of photons. Incident light can either be non-polarized in the case when the field vectors of the photons are randomly oriented, or it can be linearly polarized, when the field vectors preferentially align with a particular direction. The third possibility is that the electric vector orbits in a circular (or elliptical) motion around the direction of propagation, in which case we have circularly (or elliptically) polarized light. There is however no process known that would introduce circular or elliptical polarization into the CMB, which is why we will ignore this effect in the following.

The physical process that is responsible for the linear polarization of the CMB is *Thomson scattering*. Thomson scattering describes the elastic scattering event of a photon from a free electron, where the wavelength remains unchanged

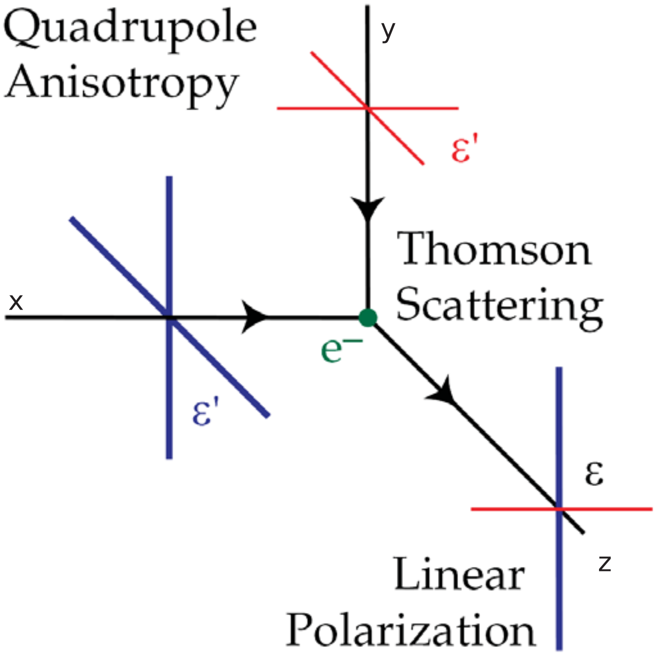


Figure 1.5: Schematic of generation of linear polarization through Thomson scattering. High (low) intensity radiation is originating from the left (top) and scattered into the direction normal to both. Figure courtesy of [16].

while the path is deflected. When a photon interacts with an electron, the electron will oscillate both parallel to, and synchronous with, the electric field of the incident light. Since an oscillating electron is an accelerated charge, this motion will result in the emission of another photon in the plane normal to the electric field that caused the oscillation. The polarization of the created photon is aligned with the oscillation direction of the electron, which is turn is aligned with the electric field vector of the sourcing photon. This has the effect, that if unpolarized light from a localized source hits an electron, polarized light will be scattered into the plane normal to the incidence direction. However, if the electron is hit by isotropic radiation, the linearly polarized light from directions 90° apart would blend, and the observed light would therefore remain unpolarized. The situation is different, however, if the incident light exhibits a *quadruople anisotropy*, i.e. the light separated by 90° differs in intensity. This situation is plotted in Fig. 1.5. An electron interacts with high-intensity light incident from the left (x-axis) and with low-intensity light from the top (y-axis). The light is scattered and we observe it along the z-axis. In this situation, the scattered light originating from the top will be observed as polarized in the x-direction, while the light coming from the left will be observed as polarized in the y-direction. Since the intensity incident from the left is dominating, the net polarization measured by the observer will be in the y-direction. It is important to note that because of the orthogonality of the spherical harmonic functions, the quadrupole is the only anisotropy which can create polarization from Thomson scattering [16].

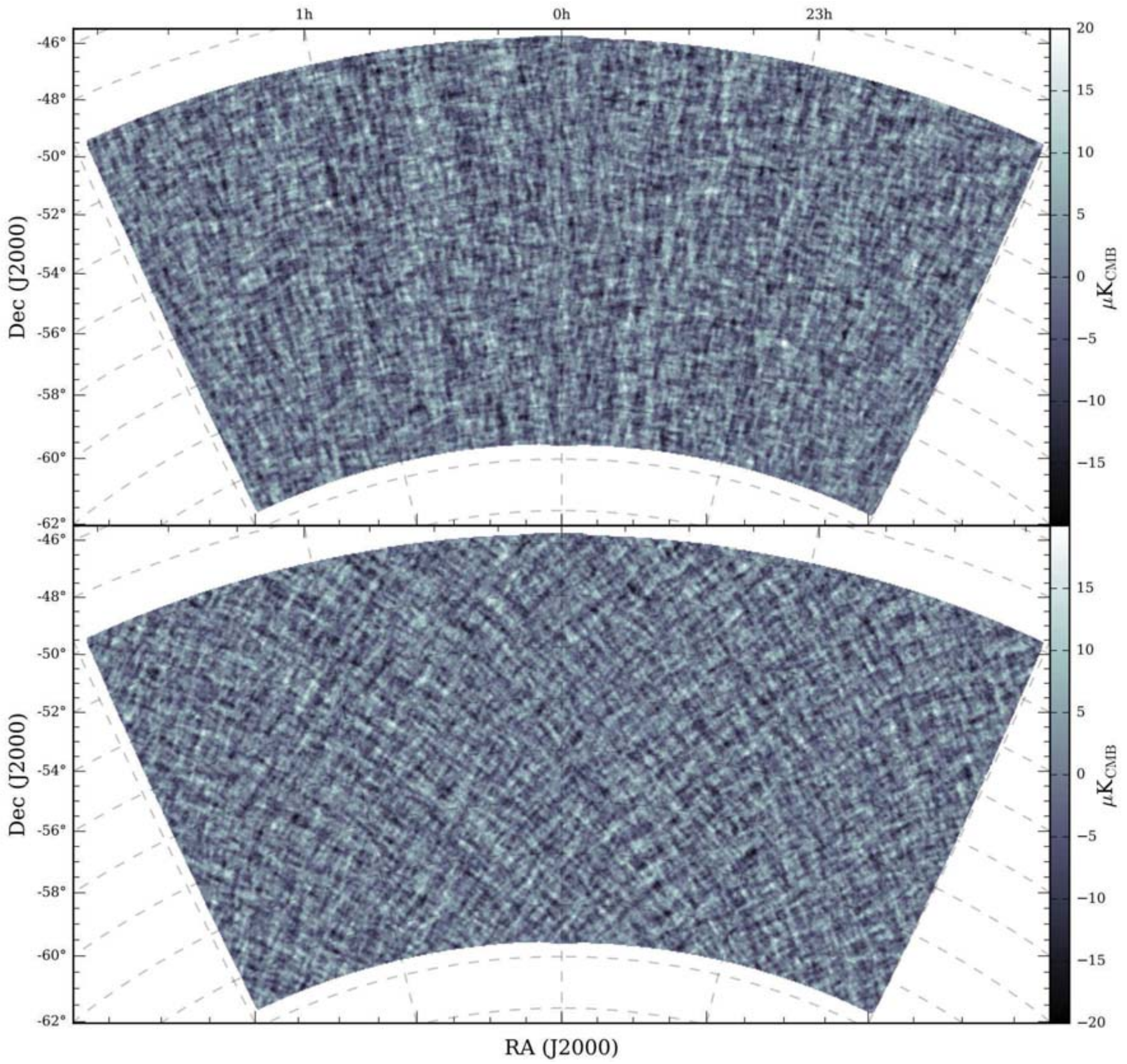


Figure 1.6: Q and U polarization maps as seen by the *South Pole Telescope*. Figure courtesy of [17].

Description of polarized light

Light, both polarized and unpolarized, is commonly described by the *Stokes parameters* (I, Q, U, V). These can be computed from the time averages of the electric field vectors given by

$$I \equiv \langle E_x^2 \rangle + \langle E_y^2 \rangle \tag{1.7}$$

$$Q \equiv \langle E_x^2 \rangle - \langle E_y^2 \rangle \tag{1.8}$$

$$U \equiv \langle E_{x'}^2 \rangle - \langle E_{y'}^2 \rangle . \tag{1.9}$$

I gives the intensity of light (i.e. temperature), which is always a positive entity. Q and U describe linear polarization as seen from two cartesian coordinate systems (x, y) and (x', y') rotated by 45°. V describes circular polarization, which we again ignore since it is irrelevant for the CMB. While I is a scalar quantity, and therefore independent from the choice of coordinate system, Q and U rely on angles measured with respect to some arbitrary axis, and are therefore coordinate dependent. We have already seen a map of the CMB Stokes parameter I on the sky, namely the temperature map in Fig. 1.3. Figure 1.6 shows a smaller patch of the sky measured by the South Pole Telescope, observed through the polarization components Q and U. Note that the polarization signal is about a factor of 100 weaker than the temperature signal. The Q map clearly shows a stripe pattern along lines of constant declination and right ascension, while the U map shows a similar pattern but rotated by 45°.

Stokes Q and U maps are useful tools to map the polarization state across the sky. In terms of cosmological analysis however, it is convenient to look for certain patterns in these maps that are independent from the choice of coordinate system. The generation and interpretation of polarization patterns in the CMB is a very complex topic, and for a more in-depth review I therefore refer the reader to [16]. In a nutshell, Thomson scattering caused by a quadrupole anisotropy allows for the creation of two distinct patterns in the polarized CMB, the so-called *E-modes* and *B-modes*. These are illustrated in Fig. 1.7 and owe their names to the electric and magnetic field, with which they share the properties of being curl free (E-mode) and divergence free (B-mode). The reason why these patterns are so important to cosmologists only becomes clear when considering what kinds of quadrupolar anisotropies exist, and how these are created. The quadrupole corresponds to the multipole moment $\ell = 2$. Hence, there are five azimuthal modes, namely $m = 0, \pm 1, \pm 2$. As it turns out, the $m = 0$, $m = \pm 1$ and $m = \pm 2$ quadrupole moments lead to distinct classes of perturbations, which are *scalar*, *vectorial* or *tensorial* respectively. Scalar perturbations are ordinary compressions of the primordial plasma, and are the only ones that lead to gravitational collapse of matter. Vector perturbations on the other hand manifest themselves as vortexes in the plasma. Lastly, tensor perturbations are created through fluctuations of the metric, which can be regarded as the stretching and compressing of space itself. This effect leads to distortions in such a way that circularly arranged test particles get twisted into ellipses. Figure 1.7 shows the quadrupole moments for scalar and tensor perturbations. The most interesting candidate in the context of the polarized CMB is the tensor perturbation, as distortions of space are nothing else than *gravitational waves*, which are postulated by *inflation*. Inflation is a major theory of cosmology originally proposed by Alan Guth in the 80s in order to resolve fundamental problems of the Big Bang model [18]. According to the theory of inflation, the Universe started out with a period of exponential growth before settling down to normal expansion. This growth process only lasted for a fraction of a second, and it predicts the excitement of tensor perturbations in the form of gravitational waves. The crucial point here is that of the three classes of perturbations, *all* are able to generate E-modes, while *only* tensor perturbations are responsible for B-modes. A detection of B-modes in the CMB would therefore be a huge success for both the inflationary paradigm and the Big Bang theory as a whole. Hence, the entire CMB community is on the hunt for this elusive signature, which is often referred to as “the smoking gun of inflation”.

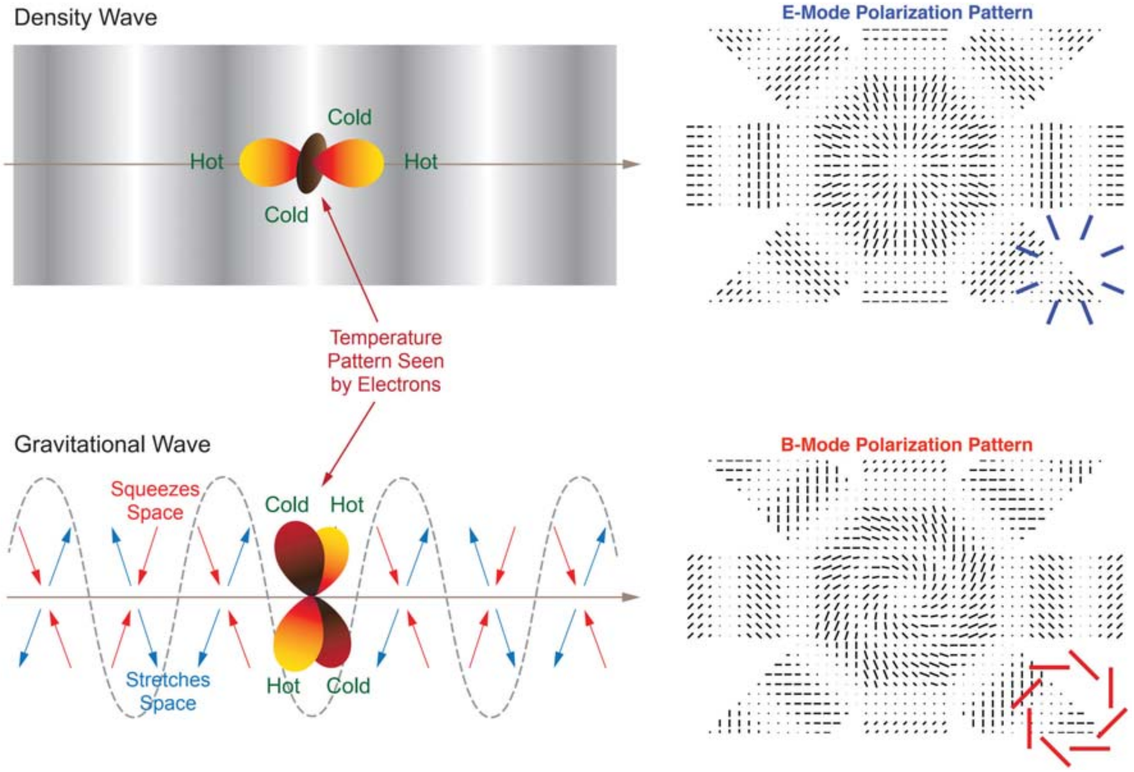


Figure 1.7: Top: The $m = 0$ quadrupole moment leads to scalar perturbations (density waves) and results in an E-mode pattern. The E-mode is a curl-free pattern where the polarization is either parallel or normal to all radial lines. Bottom: The $m = 2$ quadrupole moment leads to tensor pertubations (gravitational waves) and result in both E- and B-mode patterns. The B-mode pattern is divergence-free with the polarization direction aligning at ± 45° angles to radial lines. Figure courtesy of [19].

Polarization power spectrum

Analogous to the way we expand the temperature anisotropy field in spherical harmonic space, as given by Eq. 1.2, we can express the polarization fields in terms of their *spin ± 2 spherical harmonics*. For more information on E-B decomposition and a rigorous mathematical description, I refer the reader to [20]. This results in two further sets of spherical harmonic coefficients, from which we can compute power spectra. In total we then have three sets of observables, namely $a^{\mathrm{T}}_{\ell m}$, $a^{\mathrm{E}}_{\ell m}$ and $a^{\mathrm{B}}_{\ell m}$. Given these , the full set of power spectra using temperature, polarization and their cross-terms, is

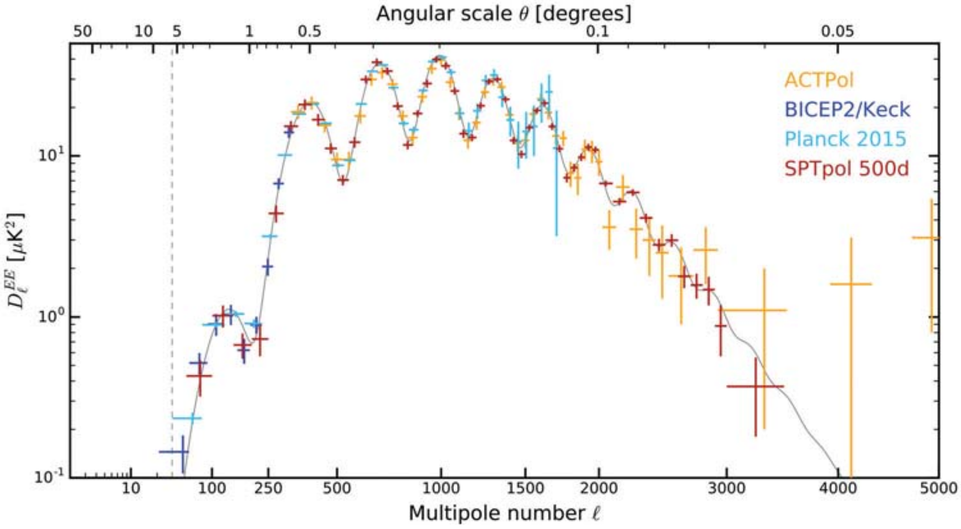
$$\begin{array}{lll} C^{\mathrm{TT}}_{\ell m} = \langle a^{\mathrm{T}*}_{\ell m} a^{\mathrm{T}}_{\ell' m'} \rangle & C^{\mathrm{EE}}_{\ell m} = \langle a^{\mathrm{E}*}_{\ell m} a^{\mathrm{E}}_{\ell' m'} \rangle & C^{\mathrm{BB}}_{\ell m} = \langle a^{\mathrm{B}*}_{\ell m} a^{\mathrm{B}}_{\ell' m'} \rangle \\ C^{\mathrm{TE}}_{\ell m} = \langle a^{\mathrm{T}*}_{\ell m} a^{\mathrm{E}}_{\ell' m'} \rangle & C^{\mathrm{TB}}_{\ell m} = \langle a^{\mathrm{T}*}_{\ell m} a^{\mathrm{B}}_{\ell' m'} \rangle & C^{\mathrm{EB}}_{\ell m} = \langle a^{\mathrm{E}*}_{\ell m} a^{\mathrm{B}}_{\ell' m'} \rangle, \end{array} \tag{1.10}$$

where $C^{\mathrm{TB}}_{\ell m} = C^{\mathrm{EB}}_{\ell m} = 0$ [20]. These can be used to further compare theory and observation, as demonstrated in Fig. 1.8. The top panel shows the EE power spectrum as measured by various CMB experiments, along with the best-fit model given by the solid line. Data and theory are in fairly good agreement.

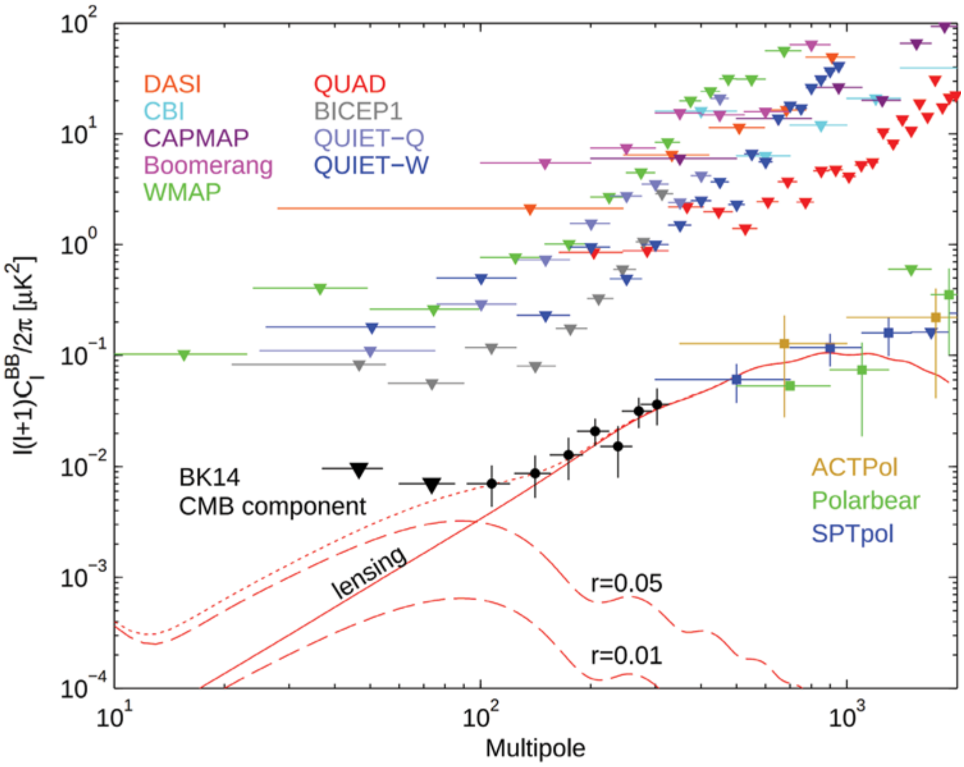
The situation is more difficult in the case of the BB power spectrum shown in the bottom panel, since the B-mode signal is notably weaker than the E-mode signal. The dashed red lines show the theoretical predictions for the gravitational wave signal for two values of the *tensor-to-scalar ratio* r. The tensor-to-scalar ratio represents the amplitude of the tensor perturbations at recombination relative to the scalar perturbations and can be computed from the EE and BB power spectra. It is commonly defined as

$$r \equiv \frac{P_{\mathrm{T}}(k_0)}{P_{\mathrm{S}}(k_0)}, \tag{1.11}$$

where P_{T} and P_{S} are the primordial power spectra for tensor and scalar perturbations respectively, and k_0 is the pivot scale, usually chosen to be $0.05\ \mathrm{Mpc}^{-1}$. At present, the collective goal of the CMB community is to measure the primordial B-mode signal. Problematic is not only the faintness of this signal, in addition one has to correct for the effect of *weak gravitational lensing*, a process which transforms E-modes into spurious B-modes. Another complicating factor is galactic dust particles that polarize the CMB, and if not accounted for, can lead to false B-mode detections as it was the case with BICEP2 in 2014 [21]. To this day, the search for B-modes has only lead to a constraint of the upper boundary of the tensor-to-scalar ratio, $r < 0.07$ [22]. The proof for inflation however is still outstanding.



(a) E-mode power spectrum. Figure courtesy of [17].



(b) B-mode power spectrum. Figure courtesy of [19].

Figure 1.8: Top: Compilation of E-mode measurements compared to theoretical model. Bottom: Compilation of B-mode measurements compared to theoretical predictions of gravitational wave signal for two values of the tensor-to-scalar ratio r (dashed) and the lensing signal (solid).

Chapter 2

Galactic Foregrounds

CMB analysis is complicated by the fact that there sits an entire galaxy between us and the last scattering surface, which is the origin of the CMB photons. The galaxy in question is not any galaxy, but our home, the Milky Way. While it looks rather sparse and transparent in visible light, the view from longer wavelengths looks rather grim. The truth is that the CMB is only one of several components of the microwave sky, and the map presented earlier in Fig. 1.3 is the result of painstaking foreground cleaning. Figure 2.1 reveals the true and unembellished picture, as observed through `Planck's` nine frequency channels. Two characteristics immediately strike the eye: First, the CMB is heavily polluted. The highest impact is visible along the equator of the maps and manifests itself as huge diffuse structures that trace the morphology of the Milky Way. Second, the undesired foreground emission varies with frequency. The sky looks notably cleaner at intermediate frequencies than at the low or high end of the frequency range. For example, the 70 GHz channel visually appears fairly tidy, especially at high latitudes, while the high frequency channels are completely dominated by foreground contamination.

The complex nature of galactic foregrounds is summarized in Fig. 2.2. These plots serve to illustrate the frequency behavior of the various foregrounds, which will be described in greater detail later. Note that the complexity of foreground emission differs greatly between temperature and polarization measurements, as only some foregrounds are capable of polarizing light. This figure also introduces the prime candidates responsible for CMB pollution:

- Thermal dust emission (Sec. 2.1)
- Synchrotron radiation (Sec. 2.2)
- Free-free emission (Sec. 2.3)
- Spinning dust emission (Sec. 2.4)

Saving the detailed description of these foregrounds for later, we can start by making sense of their spectral behavior shown in Fig. 2.2. In the temperature domain, the low-end of the frequency range is dominated by synchrotron, free-free, and spinning dust emission, while thermal dust prevails at the opposite end of the frequency spectrum. Conveniently, there remains a narrow frequency band around 100 GHz in which the CMB is still the dominating signal. In polarization the situation looks grimmer. While only synchrotron and thermal dust emission carry a polarized component, their signals dominate the *entire* frequency range and leave no loophole for the CMB to peak through. Recalling that the polarization power spectrum is the prime target to confirm the theory of

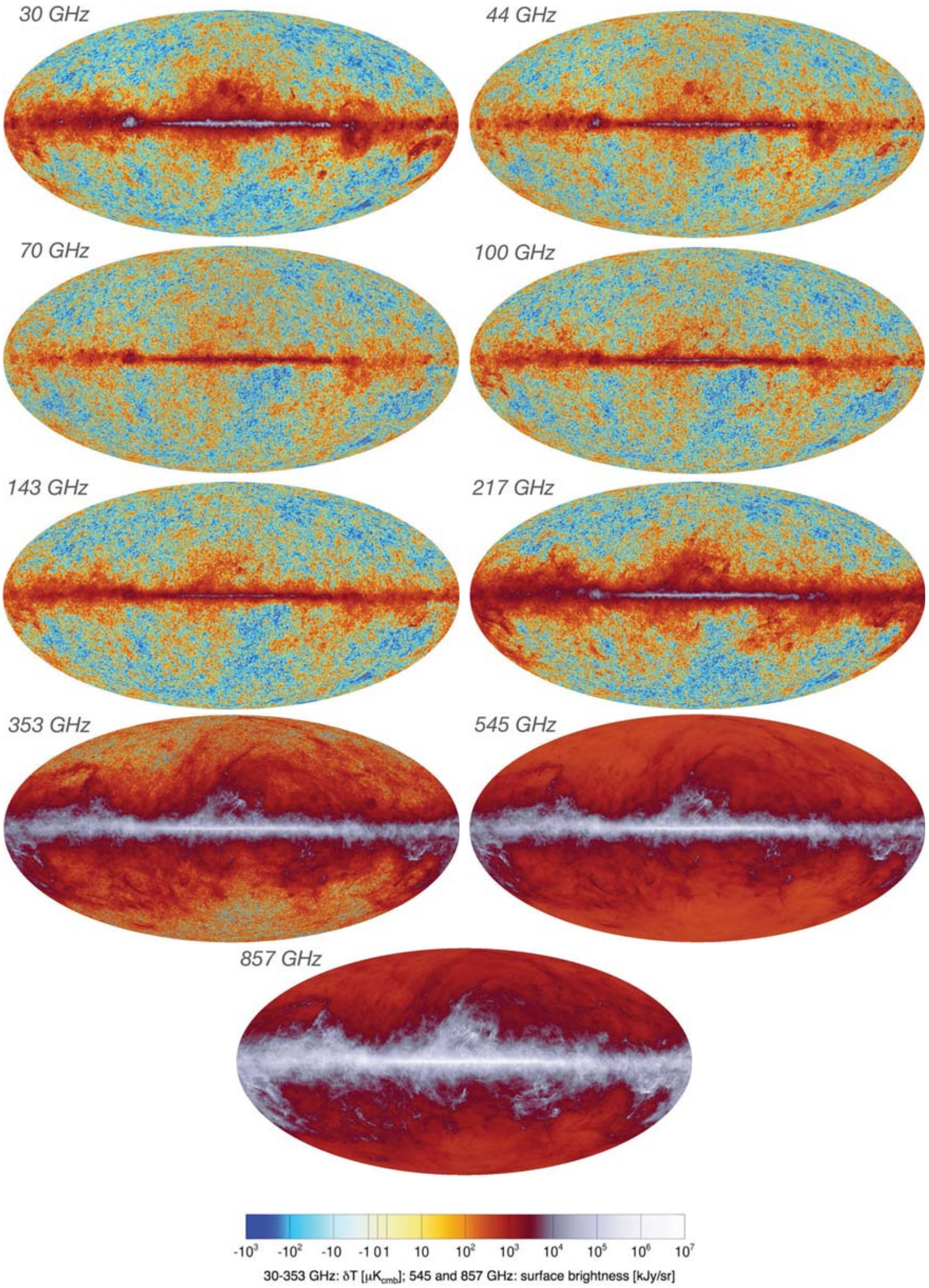
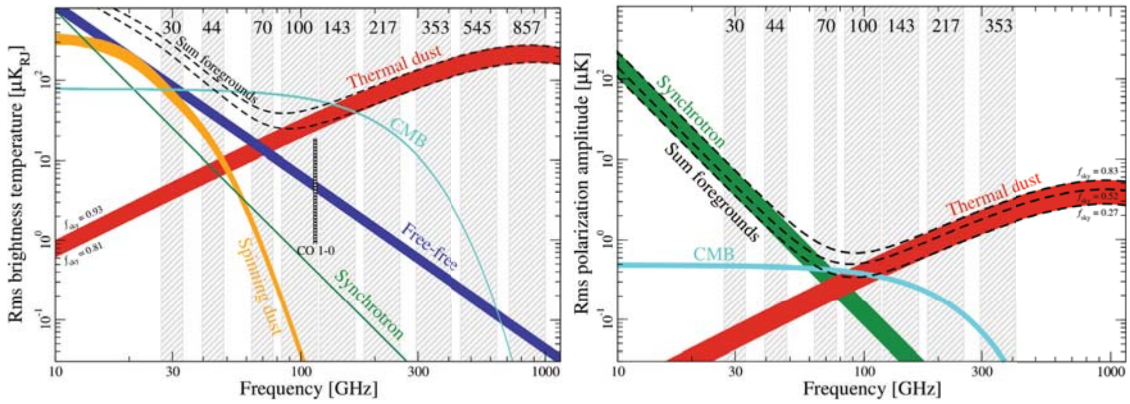


Figure 2.1: The microwave sky in temperature as seen through `Planck's` nine frequency channels. Figure courtesy of [15].



(a) Temperature

(b) Polarization

Figure 2.2: Spectral energy distribution of various foreground components for temperature (a) and polarization (b). Figure courtesy of [12].

inflation should make it clear that foreground removal is of utmost importance in order to draw any reliable conclusions from CMB measurements.

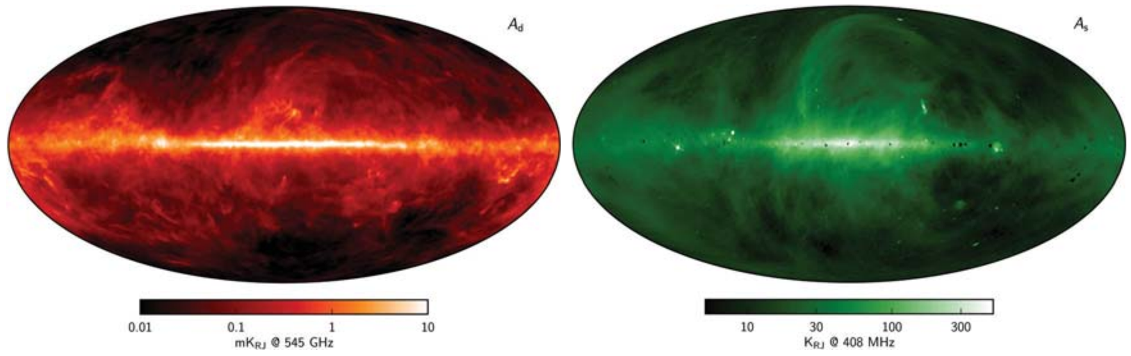
Component separation

While the CMB instrument has no way of distinguishing between a CMB and a foreground photon, one can combine the knowledge of the foreground spectral distributions with the redundance of frequency channels in order to disentangle the various foreground components in a process called *component separation*. The reason why CMB telescopes employ multiple frequency bands[1] is to make this technique possible in the first place. While `WMAP` covered the low-end of the spectrum, and therefore mainly probed the amplitude of synchrotron emission, `Planck` extended the measurements to very high frequencies in order to get a handle on thermal dust emission. The vertical gray bands in Fig. 2.2 mark the frequencies measured by `Planck`. As we will see shortly, some component separation techniques rely on accurate foreground models, i.e. parameterized model spectra as shown in Fig. 2.2, which is why it is often advantageous to combine the data from several CMB experiments and thereby overcome the parameter degeneracies associated with these models.

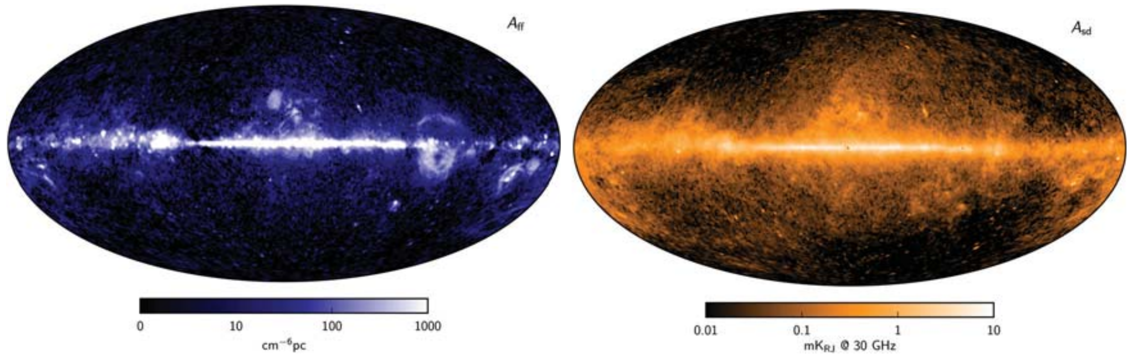
Component separation is a rather complex task and there is a large variety of methods on the market that either separate the astrophysical components or filter out the CMB from a set of data maps. In order to not depend on one single technique, `Planck` even employs four independent component separation pipelines that do just that [15]. In Chapter 3, we will concentrate on one particular method that is implemented in one of these pipelines, namely *Bayesian component separation*.

---

[1] `Planck`: 9 bands (30-857 GHz); `WMAP`: 5 bands (22-90 GHz); `COBE`: 3 bands (31.5-90 GHz)



(a) Thermal dust (b) Synchrotron



(c) Free-free (d) Spinning dust

Figure 2.3: Amplitudes of diffuse foreground components as the result of Planck component separation. Figure courtesy of [23].

2.1 Thermal dust

As shown in Fig. 2.1-2.2, thermal dust radiation is the dominating contaminant from ∼100 GHz upwards, and completely overpowers the CMB in the highest `Planck` channels. Its amplitude increases further into the infrared regime, where it has been mapped by the `COBE-DIRBE` instrument [24]. While the spectral energy distribution (SED) peaks around 2000 GHz, it is not negligible at lower frequencies where CMB anisotropies are measured [25]. Thermal dust emission is caused by minute dust grains that inhabit the interstellar space. The full variety of possible compositions is still unclear, but several compounds have been identified using spectral features of absorption/emission spectra, such as carbon in the form of graphite and silicates [26]. These dust grains heat up by absorbing starlight and subsequently reradiate the absorbed energy, thereby blending with the CMB photons. Due to its thermal origin, the radiation spectrum resembles that of a blackbody. The exact spectrum depends mainly on three factors [25, and references therein]: First, the equilibrium temperature of an ensemble of dust particles depends on the grain size distribution, as larger grains have a lower equilibrium temperature than smaller ones; second, the composition of dust particles determines their absorption and emission characteristics; third, the properties of dust particles may vary along the line of sight. Fortunately, knowing the exact dust properties is not necessary to satisfyingly model the thermal dust emission by means of phenomenological models. In practice, the thermal spectrum is commonly modeled empirically as a modified blackbody, also referred to as "graybody". Such a model is based on an ordinary blackbody emission spectrum which is reshaped by a frequency-dependent modification term. The model parameters are then fitted to match the observations. Within the `Planck` analysis, the modified blackbody emission s_d is usually modeled as

$$s_\mathrm{d} = A_\mathrm{d} \cdot \left(\frac{\nu}{\nu_0}\right)^{\beta_\mathrm{d}+1} \frac{\exp\left(\gamma\nu_0\right)-1}{\exp\left(\gamma\nu\right)-1}, \tag{2.1}$$

where $\gamma = \frac{h}{k_\mathrm{B}T_\mathrm{d}}$ and ν_0 is the pivot frequency at 545 GHz [23]. The free parameters are the spectral index (emissivity) β_d that specifies the slope of the spectrum, T_d is the equilibrium temperature of the dust, and A_d is the emission amplitude at the pivot frequency. Usually, these parameters are free to vary across the sky and are fitted per pixel. In one `Planck` analysis, these parameters were found to be distributed according to $T_\mathrm{d} \sim \mathcal{N}(23{\pm}3)$ and $\beta_\mathrm{d} \sim \mathcal{N}(1.55{\pm}0.1)$ [23]. Fig. 2.3a shows the intensity of thermal dust at 545 GHz. The majority of the emission originates from the galactic equator, but it extends well into high latitudes.

As illustrated in Fig. 2.2, confusion from thermal dust not only affects the CMB in temperature, but also in polarization. A considerable fraction of the thermal dust emission has been found to be linearly polarized, which arises from the interaction of asymmetric grain shapes with the interstellar magnetic field (IMF). A portion of the dust grains happen to be oblate, which makes their absorption/emission behavior dependent on the grain orientation. Linear polarization then arises when these grains align with the IMF, leading to a polarization fraction of up to ~ 20 %[27].

2.2 Synchrotron emission

We now turn our focus on cosmic synchrotron radiation, which hampers CMB analysis at frequencies below ~50 GHz. This is the case both for temperature and polarization analysis, as illustrated by the foreground SED in Fig. 2.2. Unlike thermal dust radiation, synchrotron radiation is non-thermal radiation generated by relativistic electrons that gyrate along galactic magnetic field lines [28]. Its characteristics therefore depend on the magnetic field strength and the energy distribution of cosmic-ray electrons that populate the interstellar medium. The energy distribution of cosmic-ray electrons can be well approximated by a power law [e.g. 29]. A single electron gyrating about magnetic field lines will continuously lose energy, and therefore radiate across a wide frequency spectrum. However, the superposition of SEDs of a large ensemble of electrons combined forms again an SED that can be described by a power law, as illustrated in Fig. 2.4a[e.g. 30, 29]. The SED can therefore be modeled as

$$s_\mathrm{s}(\nu) = A_\mathrm{s} \left(\frac{\nu}{\nu_0}\right)^{\beta_\mathrm{s}}, \tag{2.2}$$

where ν_0 is some reference frequency and β_s the power law index defining the steepness of the SED. Theory and observations have shown that above 20 GHz, $\beta_\mathrm{s} \approx -3$, with significant flattening at lower frequencies [23].

The best synchrotron survey available dates back to 1982, and is the 408 MHz all-sky continuum survey by Haslam et. al. [31]. The resulting sky map, shown in Fig. 2.4b, is the combination of maps from four different observation sites. It has been widely used to constrain the synchrotron SED in the low-frequency regime and is therefore a valuable data set for component separation techniques. While the synchrotron emission is most intense in the galactic plane, there are large structures that basically cover the entire sky. The hallmark of the synchrotron sky is a long arch-like structure that extends from the equator far up to the northern galactic pole. This feature is called the *North Polar Spur* and is assumed to be a super nova remnant [33].

Single electrons that gyrate along magnetic field lines emit either circularly or elliptically polarized light, depending on the viewing direction. However, it can be shown that an ensemble of electrons with a distribution of pitch angles[2] emits linearly polarized light, since the elliptical polarization states cancel [29]. Consequently, synchrotron radiation is found to be linearly polarized with a polarization fraction up to 75 % [23].

2.3 Free-Free emission

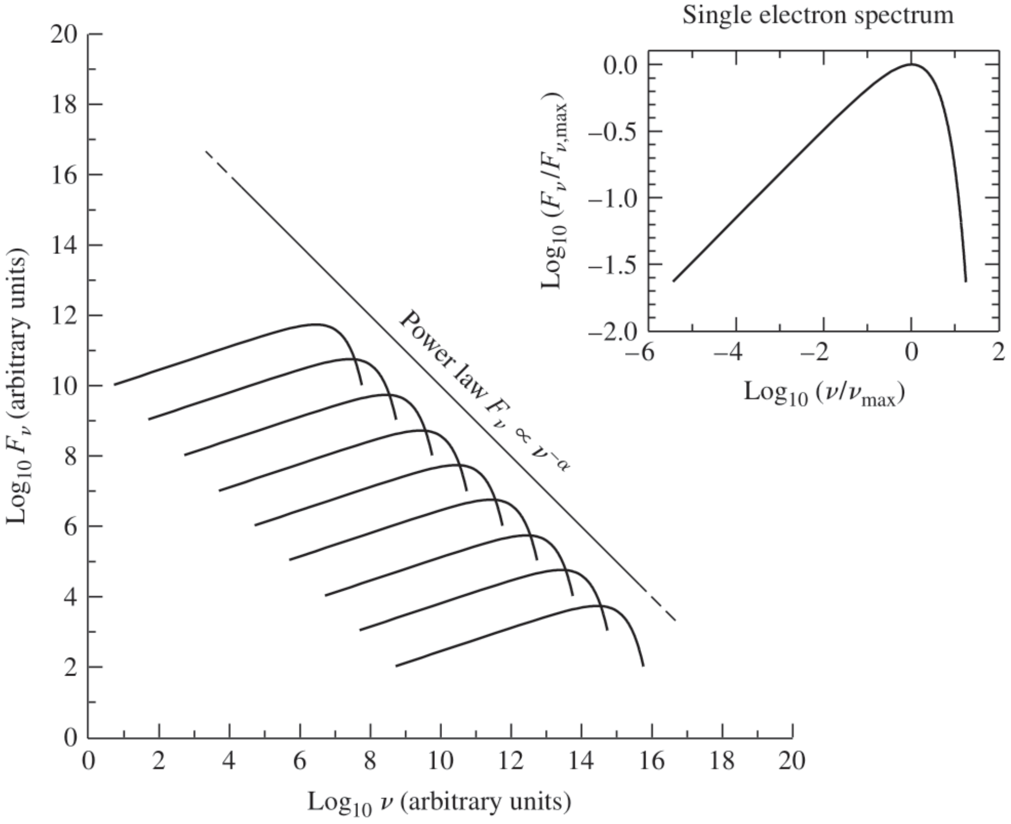
Similar to synchrotron radiation, free-free emission mainly contributes to the low frequency range of the foreground SED, as illustrated in Fig. 2.2. However, this component can be ignored in polarization. Free-free radiation is thermal *bremsstrahlung* that arises when fast charged particles (mainly electrons) interact with ions, atoms, or molecules. When these particles encounter the Coulomb field of the impactor, they are decelerated and emit radiation that can be calculated from physical grounds [34]. The term “free-free” reflects the fact that the particles both start and end in free (unbound) states. Although the physics of bremsstrahlung is well-understood, free-free radiation is problematic since it is not easily traceable. While synchrotron and thermal dust emission can be efficiently mapped at very low or high frequencies, e.g. the Haslam survey at 408 MHz or `Planck` at 857 GHz, free-free emission is only dominant in a very narrow frequency interval and therefore more difficult to handle.

The SED of free-free emission can be expressed using a two-parameter model as described in [e.g. 23, 34, and references therein]. The SED reads as

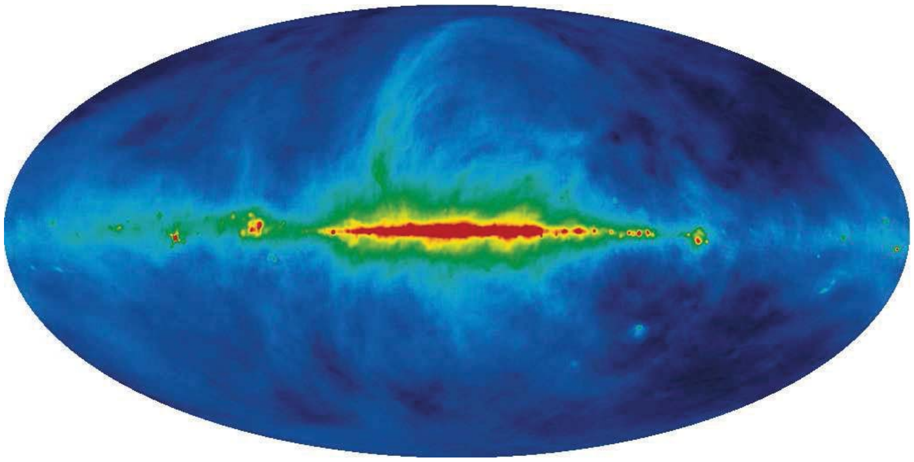
$$s_{\mathrm{ff}} = T_e \left(1 - e^{-\tau_\nu}\right) , \tag{2.3}$$

---

[2] Angle between the velocity vector $\vec{v}$ and the magnetic field line $\vec{B}$.



(a) Synchrotron SED



(b) Synchrotron sky

Figure 2.4: Top: A single electron emits a peaked spectrum, as shown in the top-right panel. The superposition of single-electron spectra from an ensemble of electrons whose energies are distributed according to a power law, leads again to an SED that is described by a power law. Figure courtesy of [30]; Bottom: Haslam map [31] reprocessed by Remazeilles et. al. [32].

where T_e is the electron temperature and τ_ν the optical depth, which is proportional to

$$\tau_\nu \propto T_e^{-1.5}\ \nu^{-2}\ \mathrm{EM}\ g_{\mathrm{ff}}. \tag{2.4}$$

Here, g_{ff} is the Gaunt factor[3]. The free parameters of the model are the electron temperature T_e and the emission measure EM, which is related to the integrated electron density along the line of sight. The resulting SED is close to a power law with a power law index around $\beta_{\mathrm{ff}} \approx -2$ [23].

Figure 2.3c shows the morphology of the free-free emission on the sky, as the result of `Planck` component separation. Its structure is correlated with those of Hα regions, since both Hα and free-free emission depend on integrated electron densities and both are emitted by the same ionized medium. However, contamination from Earth's geocoronal emission in Hα maps makes it difficult to create templates for free-free emission [34].

2.4 Spinning dust

Spinning dust emission is a contaminant that was discovered in several data sets (e.g. `COBE` [35], OVRO [36], whose morphology is strongly correlated with thermal dust emission, while its spectrum is similar to that of synchrotron radiation [35]. At first, this component was simply referred to as *anomalous microwave emission (AME)*, since the underlying physical process was unclear. The best explanation for this anomaly today is that it is another manifestation of cosmic dust particles, caused not by their thermal motion but rather by rotation about their axis. If a dust grain has a non-zero electric dipole moment, rotation will necessarily lead to emission of electromagnetic radiation. According to models, thermal dust radiation is preferentially caused by larger dust grains, while AME is caused by smaller grains [37]. AME has not been observed in polarization, which suggests that ultra-small dust particles are unlikely to align with the galactic magnetic field [38].

In practice, spinning dust emission is described in terms of phenomenological models. To address AME in the `Planck` analysis, a spectral template is rigidly shifted to find the best fit to the data [23]. Figure 2.3d shows the result from the `Planck` analysis. The spatial correlation with thermal dust is apparent.

2.5 Other foregrounds

Apart from these main foreground components, there are a few that play a less central role, but still deserve to be mentioned. Among them is *molecular line emission*, mainly caused by carbon monoxide (CO) that is concentrated in molecular clouds in our galaxy. Carbon monoxide emits light at very sharp

---

[3] Correction factor for quantum mechanical effects.

frequencies, which leads to contamination when the instrument's bandpass happens to overlap with them. Another foreground is *Zodiacal light* that originates from inside the Solar system and is caused by Sun light reflected of dust particles [39]. In practice, Zodiacal light can be modeled by fitting spatial templates to the sky maps. Apart from foregrounds of the Milky Way, there are a few that have their origin in extra galactic space. One of them is the *thermal Zeldovich effect*, which is caused by CMB photons that scatter off of high-energy electrons in galaxy clusters. The energy distribution of these photons is thereby modified, leading to a distortion of the blackbody spectrum. This effect is very weak and only plays a role around large clusters like Coma and Virgo [23]. Lastly, *cosmic infrared backround (CIB) radiation* can be detected across the entire sky, and is the result of redshifted thermal dust emission from very distant galaxies. The CIB can be described in a similar way as the CMB, namely as the sum of a monopole and fluctuation term. In practice, the fluctuation term is often neglected, while the monopole is subtracted from the sky maps [23].

Chapter 3

Theory of Astrophysical Component Separation

The previous Chapter highlighted the importance of astrophysical component separation in the context of CMB analysis. Because of the diffuse foreground emission that envelops the microwave sky, data maps are essentially worthless without an adequate treatment of these contaminants. Fortunately, the wealth of multi-frequency data sets after `WMAP` and `Planck` gave rise to the development of several independent methods. These can be roughly categorized into two groups, depending on the basic objective one has in mind. One group aims to unravel the composition of microwave sources by assuming parametric models of the foreground emission spectra. The other aims for the mere extraction of the CMB signal by considering the sum of foregrounds as a whole without aiming to separate them. Applications of the latter group employ different approaches, such as fitting of foreground templates (e.g. SEVEM [40]), filtering in both pixel and harmonic space using "wavelets" (e.g. NILC [41]), or a pure filter algorithm in spherical harmonic space (e.g. SMICA [42]). In this section we will concentrate on the working principle of a component separation method that belongs to the former group, namely one that allows for the full separation of foreground components. This method is implemented in the `Commander` algorithm that is widely used in `Planck` analysis.

3.1 Bayesian Component Separation by Gibbs Sampling

In this section we will review the theory of Bayesian component separation that is the backbone of `Commander`. The problem that we are confronted with can be briefly summarized as follows: We have a set of frequency maps $\mathbf{d}_\nu$ that we know is contaminated with several foregrounds. What we are interested in are the cosmological CMB signal $\mathbf{s}$, the power spectrum C_ℓ, and the foregrounds $\mathbf{f}_\nu = \{\mathbf{f}_\mathrm{d}, \mathbf{f}_\mathrm{synch}, \mathbf{f}_\mathrm{ff}, \ldots\}_\nu$. The method that we describe here is a parametric one, which means it requires the foreground SEDs to be expressible by means of parametric models, such as it was described in Sec. 2. For this method to work, it is not important if the foreground SED is based on physical grounds, or if it is an entirely phenomenological model. What is important is that the foreground models only depend on a handful parameters θ, such as amplitudes and spectral indices, and that they don't exceed the number of frequency bands. The ultimate goal in the context of component separation is to maximize or map the *posterior distribution*

$$P(C_\ell, \mathbf{s}, \mathbf{f}_\nu \mid \mathbf{d}_\nu). \tag{3.1}$$

Expressed in words, we jointly want to find the CMB signal, the power spectrum, and the foregrounds that best fit the data. The problem presented in Eq. 3.1 has been addressed by the scientific community in two separate steps, and therefore we will follow the same approach. First, remember that the fundamental goal of cosmologists is to estimate the CMB power spectrum, since it contains all the information we can learn. The CMB map and the foregrounds themselves are merely byproducts that are needed on the stony path to C_ℓ. The first part of the problem in Eq. 3.1 is called *power spectrum estimation*, which is unrelated to foregrounds and simply stands for the process of estimating the power spectrum from a noisy CMB map. Expressed in mathematical terms, we want to maximize the probability $P(C_\ell|\mathbf{d})$. This problem will be addressed in Sec. 3.1.2. The second problem is *foreground parameter estimation* that entails fitting the foreground parameters θ to the frequency maps $\mathbf{d}_\nu$, which requires computing the posterior distribution $P(\theta|\mathbf{d}_\nu)$ per pixel. This problem will be addressed in Sec. 3.1.1. Finally, these two problems are merged in Sec. 3.1.3. Generally, this review is based on the description given by Eriksen et. al. [43, 44, 45], which is highly recommended to the interested reader.

3.1.1 Foreground sampler

The fundamental idea of foreground fitting is that the CMB follows a perfect blackbody spectrum, while all the other components show a different spectral behavior. This makes it possible to unravel the mix of microwave components collected by the instrument. The basic data model we have to deal with reads as

$$\mathbf{d}_\nu = \mathbf{A}\mathbf{s}_\nu + \mathbf{n}_\nu, \tag{3.2}$$

where $\mathbf{d}_\nu$ is a set of frequency maps as seen by the instrument; $\mathbf{s}_\nu$ is the true sky signal, which is the sum of CMB and all foregrounds; $\mathbf{A}$ stands for the convolution with the instrumental beam; lastly, $\mathbf{n}_\nu$ is the instrumental noise, which is assumed to be Gaussian distributed with a vanishing mean and furthermore uncorrelated between pixels and frequencies. The variance per pixel p is therefore given by $\sigma_\nu^2(p)$. In the next step we define a parametric frequency model for the true sky signal, that describes the intensity measured at pixel p in terms of the modeled foregrounds. For example, a model including only CMB and synchrotron emission would look like

$$S_\nu(\theta) = S_\nu\left(T_{\mathrm{CMB}}, A_{\mathrm{s}}, \beta_{\mathrm{s}}\right) = B_\nu(T_{\mathrm{CMB}}) + A_{\mathrm{s}}\left(\frac{\nu}{\nu_0}\right)^{\beta_{\mathrm{s}}}. \tag{3.3}$$

Here, the synchrotron component is given by a power law, which depends on an amplitude A_{s} with respect to some reference frequency ν_0, and a spectral index β_{s}. B_ν describes the blackbody term of the CMB, which is only characterized by the CMB temperature T_{CMB}. In order to find the best-fit parameters per pixel, we need to compute the posterior distribution $P(\theta|\mathbf{d}_\nu)$, generally given by $P(\theta|\mathbf{d}_\nu) \propto \mathcal{L}(\theta)P(\theta)$, where $P(\theta)$ is a prior on θ. Since the noise is assumed

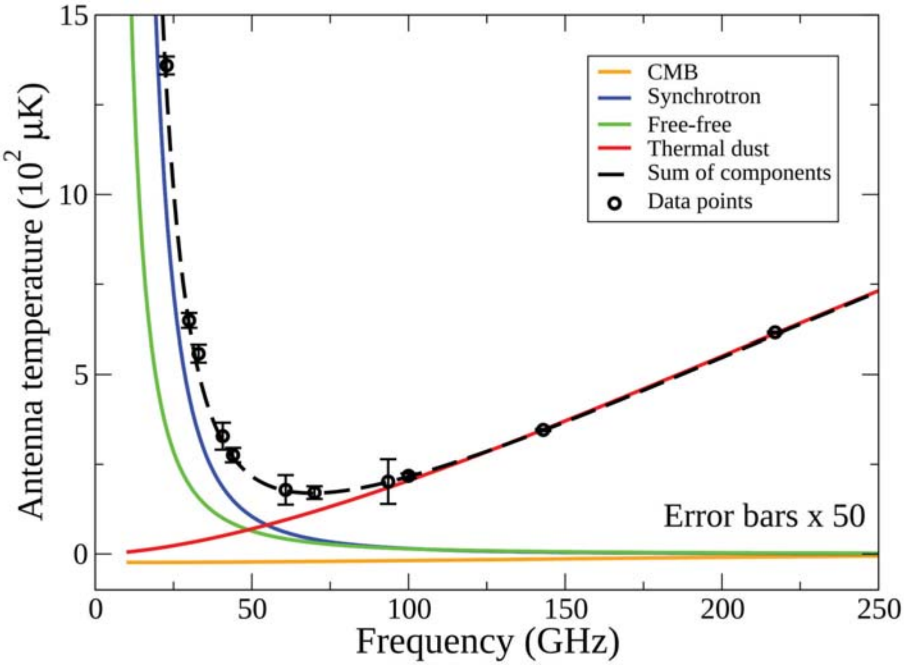


Figure 3.1: Parametric fit of foreground parameters θ to the observed data points, performed individually per pixel. Figure courtesy of [44].

to be Gaussian, the likelihood $\mathcal{L}$ may be reduced to a simple χ^2 given by

$$\chi^2(\theta) = \sum_{\nu} \left(\frac{d_\nu - S_\nu(\theta)}{\sigma_\nu} \right)^2 = -2 \ln \mathcal{L}(\theta). \tag{3.4}$$

The objective now is to minimize the χ^2, or equivalently maximize the likelihood, for each single pixel. This is done using standard minimization techniques such as MCMC. The resulting spectral parameter fit for an arbitrary pixel is illustrated in Fig. 3.1. It shows a parametric model that includes the CMB and three foregrounds. The circles show the measured data points d_ν, with error bars given by σ_ν. The parameters θ are fitted in such a way that the sum of the modeled foregrounds indicated by the dashed line matches the data points. The individual foreground SEDs resulting from the parametric fit are shown as colored curves. Given the foreground parameters θ, it is now possible to evaluate the foreground amplitude at any desired frequency in the valid range.

3.1.2 Power spectrum estimation

This section describes the problem of estimating the power spectrum C_ℓ from noisy CMB maps. The task that will be described here is closely related to that in Paper I, and I therefore invite the reader to view Part III for a real-world application of the sketched algorithms. For the purpose of power spectrum estimation, we adopt the simple model

$$\mathbf{d} = \mathbf{s} + \mathbf{n}, \tag{3.5}$$

where $\mathbf{d}$ is the observed data map, $\mathbf{s}$ is the CMB signal, and $\mathbf{n}$ is instrumental noise. Further, we assume both the signal and the noise to be Gaussian random fields with vanishing mean and covariance matrices $\mathbf{S}$ and $\mathbf{N}$ respectively. The overall goal is to estimate the probability density $P(C_\ell|\mathbf{d})$. While direct sampling from this distribution is not a trivial task, it can be shown that it is feasible to sample from the joint distribution $P(C_\ell, \mathbf{s}|\mathbf{d})$, and then marginalize over the signal $\mathbf{s}$. The reason why this is convenient lies in the theory of *Gibbs sampling*. Assume $P(x, y)$ is the probability of interest, while only the conditional distributions $P(x|y)$ and $P(y|x)$ are known. Then, according to Gibbs theory, the consecutive sampling steps

$$x^{i+1} \leftarrow P(x|y^i) \tag{3.6}$$

$$y^{i+1} \leftarrow P(y|x^{i+1}) \tag{3.7}$$

from the conditional densities will after some burn-in period converge to being samples of the multivariate density $P(x, y)$. This idea translated to the problem at hand, $P(C_\ell, \mathbf{s}|\mathbf{d})$, leads to the sampling steps

$$\mathbf{s}^{i+1} \leftarrow P(\mathbf{s}|C_\ell^i, \mathbf{d}) \tag{3.8}$$

$$C_\ell^{i+1} \leftarrow P(C_\ell|\mathbf{s}^{i+1}, \mathbf{d}), \tag{3.9}$$

where we note that $P(C_\ell|\mathbf{s}^{i+1}, \mathbf{d}) = P(C_\ell|\mathbf{s}^{i+1})$, since knowing the noisy data $\mathbf{d}$ while already knowing the signal $\mathbf{s}$ does not lead to any more information. This means that the theory of Gibbs sampling allows us to reduce the problem of sampling from the joint density $P(C_\ell, \mathbf{s}|\mathbf{d})$ to that of sampling from two conditional densities $P(\mathbf{s}|C_\ell, \mathbf{d})$ and $P(C_\ell|\mathbf{s})$. Sampling from these probability densities is still no trivial task, but there exist textbook algorithms to do that. For the sake of completeness, we will briefly sketch the sampling procedures while following the description by [46, 43].

C_ℓ-sampling

C_ℓ-sampling is the problem of drawing samples from the probability distribution $P(C_\ell|\mathbf{s})$, which makes up one of two Gibbs sampling steps (Eq. 3.9). It can be shown, which I will omit, that $P(C_\ell|\mathbf{s})$ follows the *inverse Gamma distribution*. To draw a sample C_ℓ^{i+1} given the signal map $\mathbf{s}^{i+1}$ from the inverse Gamma distribution, let's define $s_{\ell m}^{i+1}$ as the spherical harmonic coefficients of $\mathbf{s}^{i+1}$. In the next step, compute $\sigma_\ell = \sum_{m=-\ell}^{+\ell} \left|s_{\ell m}^{i+1}\right|^2$. Finally, build a vector ρ_ℓ of length $(2\ell - 1)$ with Gaussian random variates with zero mean and unit variance. The power spectrum sample is then

$$C_\ell^{i+1} = \frac{\sigma_\ell}{|\rho_\ell|^2}. \tag{3.10}$$

Map sampling

Now we will turn to the problem of drawing samples from $P(\mathbf{s}|C_\ell^i, \mathbf{d})$, as given by Eq. 3.8. To get a more intuitive picture of the sketched approach, I again invite the reader to consult Paper I in Part III. The map sampling procedure is performed in two steps. First, we have to compute the so-called *mean field map* by solving the following mapmaking equation for $\hat{\mathbf{x}}$,

$$\left(S^{-1} + Y^t \mathbf{N}^{-1} Y\right) \hat{\mathbf{x}} = Y^t \mathbf{N}^{-1} \mathbf{d}. \tag{3.11}$$

Here, S is a prior matrix that contains a guessed CMB power spectrum in its diagonal, while Y symbolizes spherical harmonic transforms. The resulting mean field map $\hat{\mathbf{x}}$ is a *Wiener filtered map*, which is biased. To generate an unbiased sample, we therefore have to add a fluctuation term $\hat{\mathbf{y}}$ by solving the mapmaking equation

$$\left(S^{-1} + Y^t \mathbf{N}^{-1} Y\right) \hat{\mathbf{y}} = S^{-1/2} \omega_1 + Y^t \mathbf{N}^{-1/2} \omega_2, \tag{3.12}$$

where the ω are Gaussian white noise maps with zero mean and unit variance. The full sample $\mathbf{s}^{i+1}$ is then the sum of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$.

3.1.3 Joint sampling

The final goal is now to join the foreground sampler from Sec. 3.1.1 and the power spectrum sampler from Sec. 3.1.2 into one single Gibbs sampling algorithm. We therefore aim to find a maximum to the joint probability $P(C_\ell, \mathbf{s}, \theta|\mathbf{d}_\nu)$. Let's return to the parametric model from Sec. 3.1.1 that only contains the CMB and the synchrotron component. The parameter vector therefore only contains two elements, namely $\theta = \{A_\mathrm{s}, \beta_\mathrm{s}\}$. The joint algorithm then consists of the following Gibbs sampling steps,

$$\begin{align}
\beta_\mathrm{s}^{i+1} &\leftarrow P(\beta_\mathrm{s}|C_\ell^i, A_\mathrm{s}^i, \mathbf{s}_\mathrm{CMB}, \mathbf{d}) \tag{3.13}\\
A_\mathrm{s}^{i+1} &\leftarrow P(A_\mathrm{s}|C_\ell^i, \beta_\mathrm{s}^{i+1}, \mathbf{s}_\mathrm{CMB}^i, \mathbf{d}) \tag{3.14}\\
\mathbf{s}_\mathrm{CMB}^{i+1} &\leftarrow P(\mathbf{s}_\mathrm{CMB}|C_\ell^i, A_\mathrm{s}^{i+1}, \beta_\mathrm{s}^{i+1}, \mathbf{d}) \tag{3.15}\\
C_\ell^{i+1} &\leftarrow P(C_\ell|\mathbf{s}_\mathrm{CMB}^{i+1}). \tag{3.16}
\end{align}$$

Note that the first two steps are performed for each individual pixel, while the map and power spectrum sampling steps are performed in spherical harmonic space. Furthermore, using this approach, the CMB temperature T_CMB is no longer an element of the parameter vector θ, since the CMB component is sampled by means of solving the mapmaking equations 3.11 and 3.12, as described above.

Part II

Analysis

The generic analysis pipeline

`GreenPol`, `SPIDER`, and `Planck` are three vastly different CMB experiments, not only considering their technical specifications, but also in their operating environment and the involved time scales. `Planck` was a billion-euro satellite mission by the European Space Agency, owing its name to famous scientist *Max Planck* who derived the blackbody formula. Launched into space in 2009, it observed the full sky for the considerable duration of four years until retiring in 2013. The collected data occupied hundreds of scientists for years and still offered work for me when I joined in 2016. `SPIDER` on the other hand is a balloon-borne experiment that took off in 2015 from McMurdo station in Antarctica. The observation time was limited to approximately 16 days, and for this reason and its lower vantage point from Earth's atmosphere, it concentrated on a small patch in the southern hemisphere. Its lifetime however is not over yet, since another flight is planned for the near future. I joined `SPIDER` in 2017, and its data kept me busy for the better part of my remaining time as a PhD candidate. Last but not least, `GreenPol` is a ground-based telescope. It's compactly constructed to fit into a standard shipping container, which makes it a mobile unit deployable at basically any desired observation site. As the name suggests, the dedicated observing location for `GreenPol`l was the icy center of Greenland. In stark contrast to `SPIDER` and `Planck`, I didn't join the project at the data analysis stage, but rather on-site when the first scientific data were taken. This meant going to Greenland to spend a month on the ice, an unforgettable experience.

In general, CMB experiments are undertakings that take many years of effort. First, the instrument has to be designed, built, and tested. Then, during the critical operational phase, data are taken for periods ranging from weeks such as `SPIDER` and `GreenPol`, to even years in the case of `Planck`. While my time as a PhD candidate was too short to accommodate an entire CMB mission, I got to taste many different sides of CMB analysis by joining experiments at early, intermediate, and late stages of their respective lifetimes. In the following I will briefly sketch the generic CMB analysis pipeline and emphasize the parts I was involved in.

The very first step of the generic CMB pipeline is the act of data taking, i.e. the measurement. I experienced on-site data acquisition and telescope maintenance through `GreenPol`. Telescopes scan the sky pixel-by-pixel, resulting in long rows of time-ordered data (TOD). These are typically messy, since CMB instruments are subject to nonidealities, such as systematic effects caused by the instrument itself, or interference from outside. Consequently, raw TOD have to be processed in order to not propagate these nonidealities into the scientific products. One of my main challenges of the previous years was the implementation of a *filter algorithm* that frees the notoriously messy `SPIDER` data from such effects.

Clean TOD are an absolute must for the next processing step, namely *mapmaking.* Mapmaking is the process of converting the time-ordered data, i.e.

consecutive measurements of sky temperature combined with directional information, into one or several maps on the sky. Any systematics left in the TOD will either visibly mess up the maps, or bias the inferred cosmological parameters. Mapmaking for `SPIDER` was one of my main responsibilities during my second year.

Sky maps free from systematic effects are still far from being ready for cosmological parameter estimation. As we have discussed in Chapter 2, the microwave sky is a composite of many different microwave sources. Therefore it is necessary to disentangle the mix of components, in other words perform *component separation*, as described earlier in Chapter 3. I have performed this task on several occasions related to the `Planck` data set.

Once we have obtained a clean CMB map, we can progress with the final two steps of the generic pipeline. The first is to *estimate the power spectrum from* the CMB map, as described in Sec. 3.1.2, and the second is to *estimate the cosmological parameters* that give the best fit to the observed spectrum. However, before performing these last two steps, it is common practice to perform null tests, a technique that verifies the absence of systematic effects in the CMB map. The general idea is to split the full mission data into two, and propagate these two parts through the analysis pipeline. The power spectrum of the difference of these two maps should then only show the noise, since the CMB signal cancels. Comparing this to simulations indicates whether you have all systematic effects under control, or whether the maps might be biased. This is a task that I was involved with in the `Planck` analysis, see Paper II and III in Part III of this thesis.

I will use the remaining chapters of this thesis to present the individual pieces of the generic pipeline by using `GreenPol`, `SPIDER`, and `Planck` as my model examples. The rest of this thesis is structured as follows: Chapter 4 describes the process of data acquisition by the example of `GreenPol`. In Chapter 5 we progress to data cleaning and mapmaking by the example of `SPIDER` data. Lastly, in Chapter 6 we will show how component separation is performed in practice.

Chapter 4

GreenPol: On-site Data Acquisition

4.1 Mission goal

`GreenPol` is a ground-based CMB telescope aiming to map the polarized galactic foreground emission at very low frequencies. Since the polarized CMB signal is dominated by thermal dust and synchrotron emission, sensitivity alone is not the only key player for a successful B-mode detection. The present-day obstacles are foreground characterization, and while dust emission was pretty well mapped by `Planck`, there exist few measurements at low frequencies to constrain the synchrotron foreground. Low-frequency observations are therefore in great demand, and `GreenPol's` objective is to take on this task.

There is also the question of what sky area to target, since ground-based telescopes evidently only map parts of the sky. While most ground-based CMB telescopes observe from high-altitude locations on the southern hemisphere, such as the Atacama desert[1] or the South Pole[2], the northern hemisphere has mostly been avoided. Another scientific goal is therefore to target the northern sky and thereby assess viable long-term research sites in the northern hemisphere. Good observing sites are characterized by high altitudes and low water vapor in the air. Figure 4.1 shows a global satellite measurement of the *mean precipitable water vapor*[3] in the atmosphere. Inaccessible locations like the Himalaya aside, there seems to be only one suitable candidate qualified for high-quality CMB observations, namely the polar desert in the center of Greenland.

4.2 Summit Camp

The selected observation site is the *Summit Camp*, situated on a 3 kilometer-thick slab of ice in the center of Greenland. The exact location is shown in Fig. 4.2, along with an aerial overview of the camp. In order to avoid RFI[4], the telescope was positioned outside of the town proper, about 1 kilometer away from the main hub. The telescope itself was constructed in such a way that it fits inside a standard shipping container, as shown in Fig. 4.3. The entire structure was fitted with rails, on which the instrument would roll out and observe when the weather conditions were suitable, or otherwise fold up and take shelter inside the container. The mission was scheduled for six weeks in summer of 2018. Figure 4.10 shows the loading process onto a transport plane

---

[1] e.g. *Atacama Cosmology Telescope (ACT)* or *Simons Observatory*

[2] e.g. *South Pole Telescope*, *BICEP*, *Keck Array.*

[3] Depth of water in a column of the atmosphere, if it was precipitated as rain.

[4] Radio Frequency Interference.

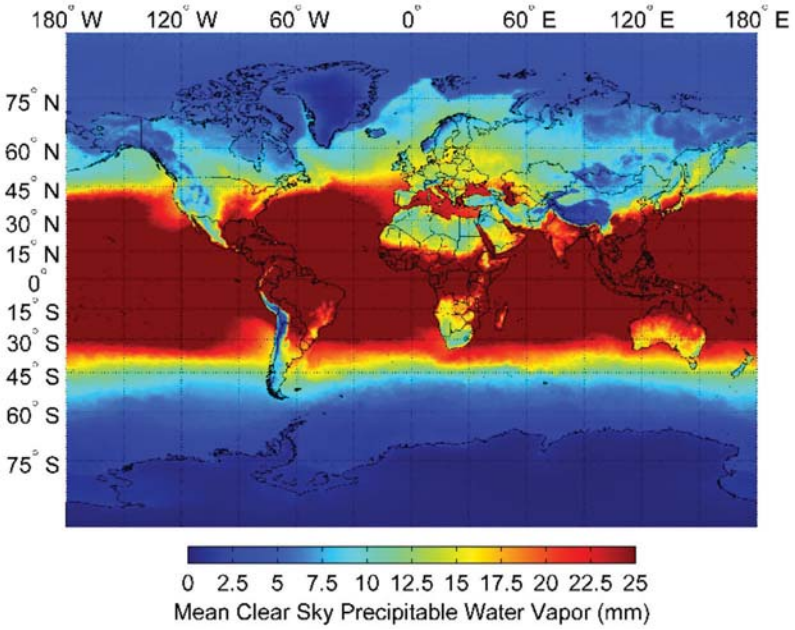


Figure 4.1: World map of mean precipitable water vapor. In the northern hemisphere, only Greenland seems to be a viable option. Figure courtesy of [47].

after completing the mission, along with miscellaneous impressions from the site.

4.3 The instrument

`GreenPol` is an off-axis carbon fiber telescope of Gregorian type. Its main components are a 2.2 m primary parabolic reflector, a 0.9 m secondary reflector, a rotating polarization modulator[5] and a dewar. The dewar is a cooled vacuum vessel that houses three feed horns at 10 GHz, an amplifier, and the detector unit.

Optical path

The optical path of the instrument is sketched in Fig. 4.4. The incoming microwave radiation from the sky is collected by the primary mirror and guided to the feed horns via the rotating polarization modulator and the secondary mirror. The rotating polarization modulator sits at the focal point of both the primary and the secondary reflector dish and serves to rotate the polarization plane of the incoming radiation. The secondary reflector projects the light onto the dewar window and into the feed horns. The window is made from an infrared filtering material and seals the vacuum chamber. The feed horns and the front-end amplifier (see Fig. 4.5) are cryogenically cooled down to a tempera-

---

[5] i.e. spinning half wave plate

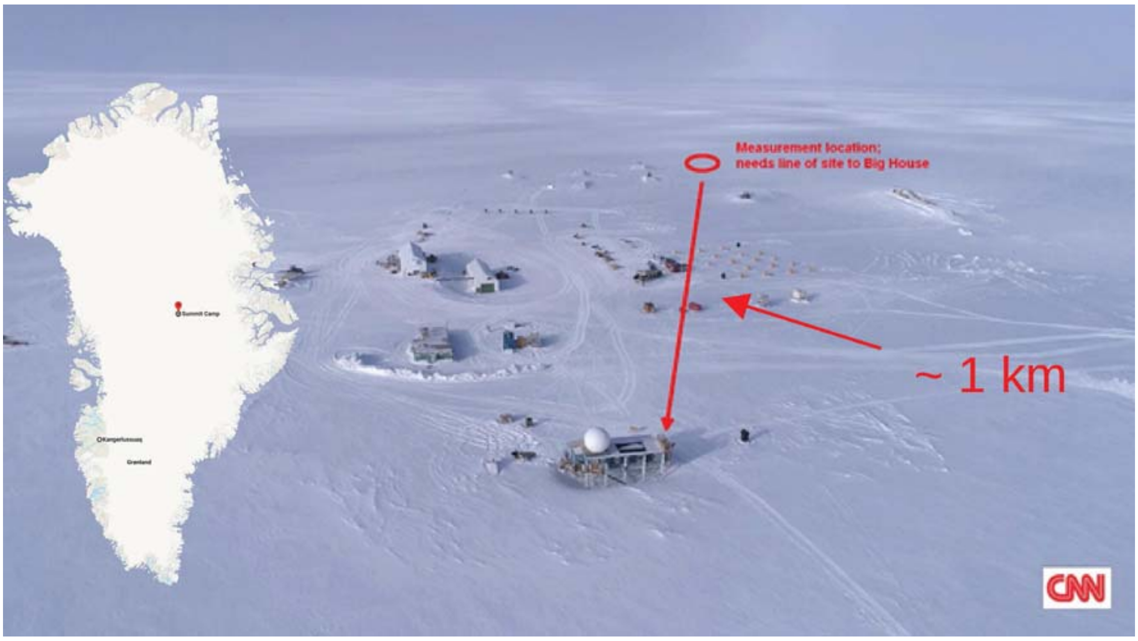


Figure 4.2: Map of Summit Camp and its geographical location on Greenland's ice sheet.

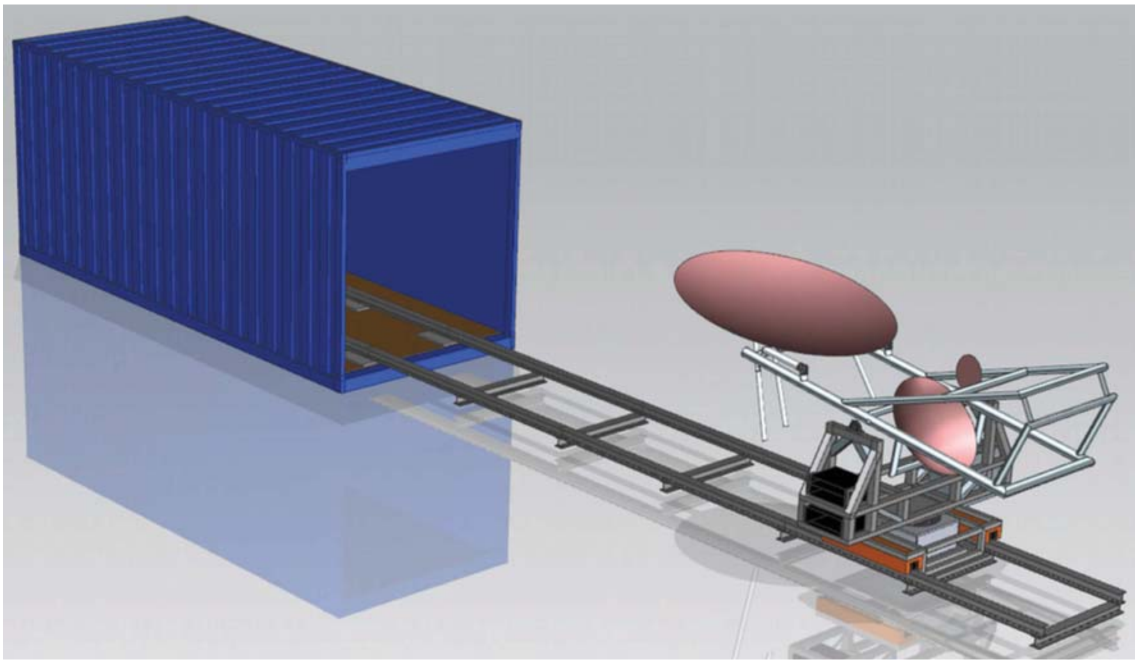


Figure 4.3: "Telescope in a box"; the instrument is foldable and fitted on rails, which makes it easily storable in case of poor weather conditions. Figure courtesy of *UCSB*.

ture of ∼20 K. When leaving the vacuum chamber, the microwave radiation is further amplified and finally absorbed by the detector diode.

Figure 4.4: Sketch of optical path with baffles removed.

Polarization modulator

The rotating polarization modulator is a key element in the optical path and allows for a better separation of the polarized signal from instrumental noise, thereby increasing the sensitivity of the telescope. The basic idea behind this approach is to modulate the polarization angle of the incoming radiation at a specific frequency f, which is outside the range of typical $1/f$-noise. The modulator has to be placed at the very beginning of the optical path, in order to ensure that only the incoming signal is modulated, and not the noise or systematics whose suppression is the objective of the technique.

`GreenPol's` modulator is a spinning half-wave plate (see Fig. 4.10)[6], which consists of a polarizing wire-grid mounted on top of a reflective material. The entire

---

[6] This picture demonstrates the persistent fight against freezing fog.

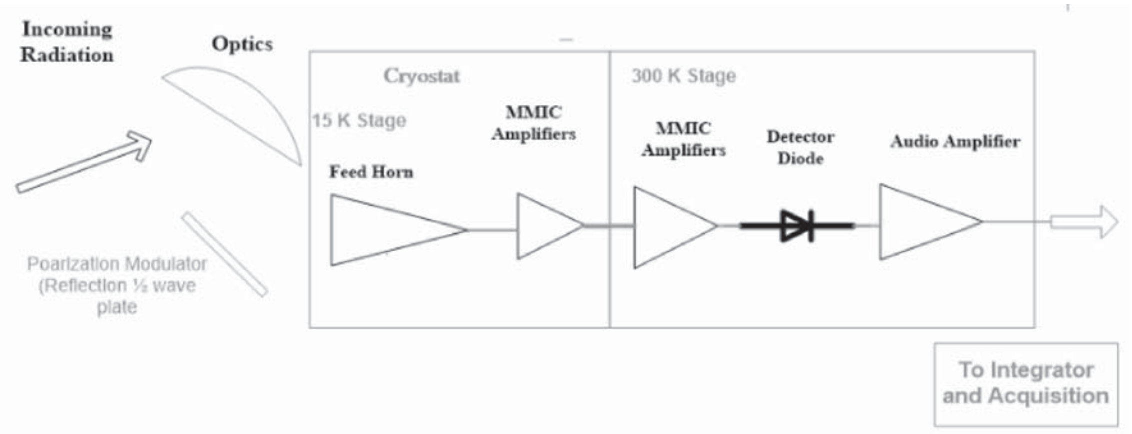


Figure 4.5: Schematics of dewar components. Figure courtesy of *UCSB*.

plate spins at a frequency of 20 Hz. Light that hits the grid is decomposed into its two polarization states, parallel and perpendicular to the wires. The parallel component is reflected off of the surface, while the perpendicular component passes through the wires in order to be reflected by the material underneath and thereby uniting with the parallel component. The space between the grid and the reflective plate creates a phase shift between the two states, which has the effect of rotating the polarization plane of the incoming radiation. The combination of the rotation of the half-wave plate and the rotation caused by the phase shift results in the rotation of the polarization vector at twice the mechanical speed. Consequently, a detector sensitive to only one polarization direction will observe a linearly polarized signal as a sinusoidal modulated intensity signal. The Stokes parameters from Eq. 1.7-1.9 are then computed from time-separated intensities of the sinusoidal signal.

Scanning strategy

The *scanning strategy* of a CMB telescope defines the pattern that is used to map the sky. `GreenPol` performs *constant elevation scans*, which are nothing else than repeated circular scans on the sky. The telescope is tilted to a user-defined elevation, for instance 70° above the horizon, and then repeatedly rotated along the azimuth direction while keeping the elevation *constant*, therefore the expression. As the Earth rotates, the projected circle will slowly swipe across the sky, as in Fig. 4.6. The chosen elevation angle determines the covered sky fraction and the depth of the map. A larger circle will cover a larger fraction of the sky, but since the measurements are spread out on a larger surface, the signal-to-noise ratio will consequently be lower. The reverse is true for small circles. We opted for a varying elevation between 65° and 70°.

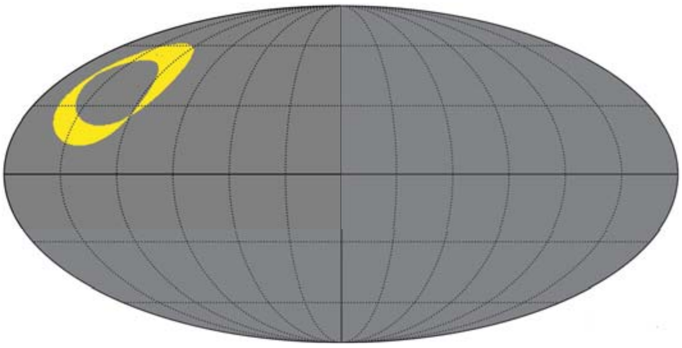


Figure 4.6: Illustration of scanning strategy. Circles at constant elevation will swipe the sky as Earth rotates.

4.4 Working at Summit

Running and maintaining a telescope amounts to more than just sitting in a shack and looking at a computer screen. Being in an extreme climate entails constant weather monitoring to ensure the safety of the telescope. While cloud cover only impairs the quality of the collected data, other “natural disasters” like freezing fog on the instrument might lead to electronic shortcuts when the ice melts later on, or strong winds might turn the telescope into a very expensive kite. The not-so-stable weather at Summit often called for several stow and deploy procedures per day, which isn’t automated but depends on manpower.

Other tasks amount to basic maintenance work, such as helium refills for the cryogenic cooling system, fuel refills for the power generator, and of course never-ending snow shoveling. Maintenance work also implies repair work on the instrument itself. During my stay at Summit, several system failures cropped up that had to be taken care of before continuing to observe. The most severe ones were a temperature leak in the cryogenically cooled dewar, which meant dismounting the heavy vacuum chamber from the telescope structure, and opening it up to fix the problem. There was also a jumping bit in the electronics, probably caused by vibrations during the transportation phase, that corrupted the time stamp on the science data and required fixing. The largest of all problems however was a slipping screw at the azimuth drive, which rendered large amounts of pointing data unusable.

When the basic needs of the telescope are taken care of, one can roll out the telescope and start observing. There are several essential system checks that are performed on a regular basis, like standard calibration procedures and test measurements, which will be outlined in the following.

4.5 System tests

The following calibration and test procedures ensure the validity of the measured data and verify the functionality of the instrument.

4.5.1 Gain calibration

Since the quantity of interest is the temperature T on the sky, while the diode readout is a voltage V, we need a relation between the two. This conversion factor is the *gain* G $[K/V]$ given by

$$G = \frac{T_{\text{warm}} - T_{\text{cold}}}{V_{\text{warm}} - V_{\text{cold}}}, \tag{4.1}$$

where T_{X} are known temperatures, while V_{X} are the voltage responses of the instrument. One way to compute G is to place two loads with known temperatures into the optical path of the telescopes and then note the response of the detector by reading of the voltage. An ambient load[7] is usually used as the warm temperature (T_{warm} = ambient temperature), while the cold load is optimally liquid nitrogen (LN_2) with a known temperature of $T_{\text{cold}} \sim 77$ K, depending on the altitude. The calculation of G from Eq. 4.1 using the known temperatures and the corresponding voltages is then straightforward.
This ideal technique poses two problems. Firstly, the detector is likely to not behave linearly across this wide range of temperatures. When observing the warm load, the detector will be in compression regime, resulting in a voltage that indicates a lower signal than should be measured. Secondly, liquid nitrogen is hard to get by. Even though it was available at Summit, it is impractical to handle since the telescope has to be tilted into a very awkward position in order to safely place the container holding the nitrogen. One way to get around this problem is to guess a sky temperature (around 10-20 K) and use that as the cold load. That will give a rough estimate of the gain. A more practical way to avoid these two problems is to perform *skydips*.

Skydip method

The skydip method uses a *partial calibrator*, which is a warm load that is partially transmissive. The temperature of the partial calibrator is given by

$$T_{\text{partial}} = \beta T_{\text{ambient}} + (1-\beta) T_{\text{sky}} + T_{\text{sys}}, \tag{4.2}$$

where β is a known transmission factor (≈ 0.1) and T_{sys} the intrinsic system temperature. The system temperature is a constant and simply produces an offset in the measurement. The use of the partial calibrator helps with the above-mentioned compression issue. Since T_{sky} and T_{sys} are unknown, we can use the fact that the former has a known dependence on the elevation θ, while

---

[7] A spongy material that resembles plastic foam.

the latter is constant. Let's define the temperature on the sky as seen by the telescope as

$$T'_{\text{sky}} = T_{\text{sys}} + T_{\text{sky}} = GV_{\text{sky}}, \tag{4.3}$$

where V_{sky} is the response of the detector. Now we will perform a skydip and collect many pairs of (V_{sky}, θ). The measured signal on the sky is then

$$V_{\text{sky}}(\theta) = \frac{T'_{\text{sky}}(\theta)}{G} = \frac{T_{\text{sys}} + T_{\text{sky}}(\theta)}{G}. \tag{4.4}$$

Using the known elevation dependence of the real sky temperature

$$T_{\text{sky}} = \frac{T_{\text{z}}}{\sin(\theta)}, \tag{4.5}$$

and defining T_{z} as the temperature in the direction of zenith, we find

$$V_{\text{sky}}(\theta) = V_{\text{sys}} + \frac{V_{\text{z}}}{\sin\theta}. \tag{4.6}$$

Now we can compute V_z and V_{sys} by performing a least square fit to the measured data. In the next step we move our telescope to an arbitrary elevation θ_{cal} and measure the detector response with and without partial calibrator, which gives us V_{partial} and V_{sky} respectively. Now we are ready to compute the gain G. We start out with Eq- 4.1,

$$G = \frac{T_{\text{partial}} - T'_{\text{sky}}}{V_{\text{partial}} - V_{\text{sky}}}. \tag{4.7}$$

Defining $\Delta V \equiv V_{\text{partial}} - V_{\text{sky}}$ and plugging in equations 4.2 and 4.3 gives us

$$\begin{aligned}
G\ \Delta V &= T_{\text{partial}} - T'_{\text{sky}} && (4.8)\\
&= \beta T_{\text{ambient}} + (1-\beta)T_{\text{sky}} + T_{\text{sys}} - (T_{\text{sys}} + T_{\text{sky}}) && (4.9)\\
&= \beta(T_{\text{ambient}} + T_{\text{sky}}) && (4.10)\\
&= \beta\left(T_{\text{ambient}} + \frac{GV_{\text{z}}}{\sin\theta_{\text{cal}}}\right). && (4.11)
\end{aligned}$$

Solving for G finally leads to

$$\boxed{G = \beta T_{\text{amp}}\left(\Delta V + \frac{\beta V_{\text{z}}}{\sin\theta_{\text{cal}}}\right)^{-1}} \tag{4.12}$$

This technique gives a good estimate of the gain while not relying on liquid nitrogen. In practice, the compression issue can be taken into account to some extent, so liquid nitrogen calibrations should still be performed every other day, while the skydip method is performed several times per day.

4.5.2 Side lobe measurement

The realistic beam of a telescope contains *side lobes.* These are unwanted structures of the beam that are not part of the main beam profile and result in the collection of photons from directions that deviate from the pointing direction. These are instrument-specific and have to be mapped out. This is done by placing an artificial source in front of the telescope and scanning the field of view around it, as shown in Fig. 4.7. If side lobes are present, structures will appear that do not coincide with the direction of the source, which is clearly the case in our measurement. Note that because of the limited lower elevation angle of the telescope, Fig. 4.7 doesn't show the actual source because it's out of view below the lower edge of the plot. Due to this unfortunate placement *only* the side lobes are visible.

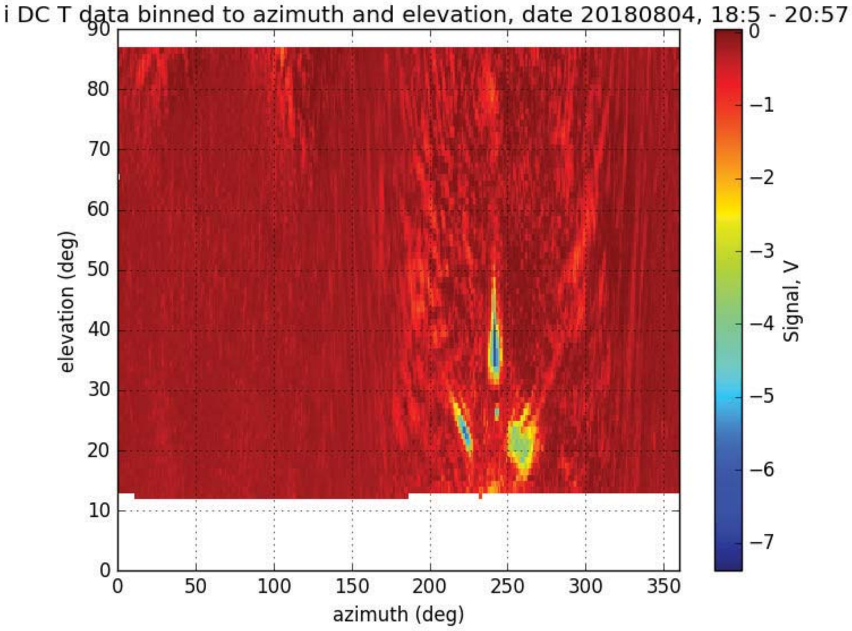
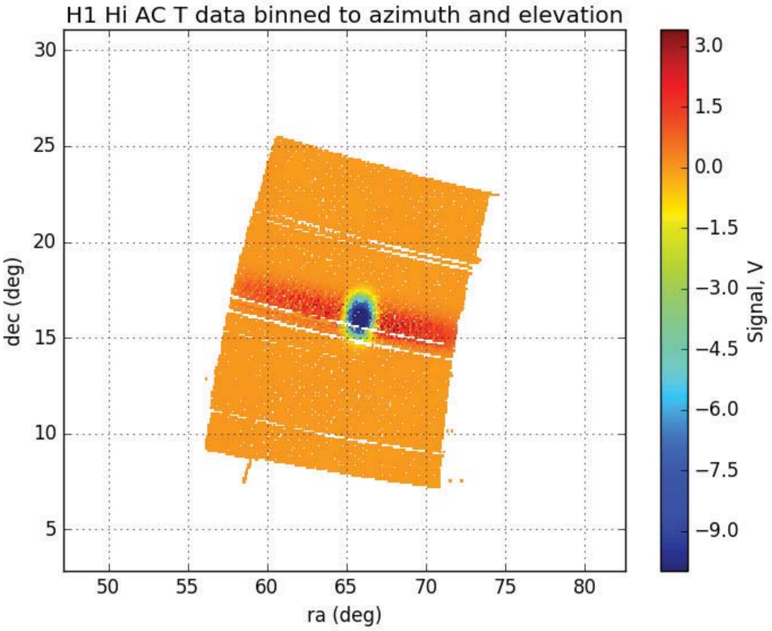


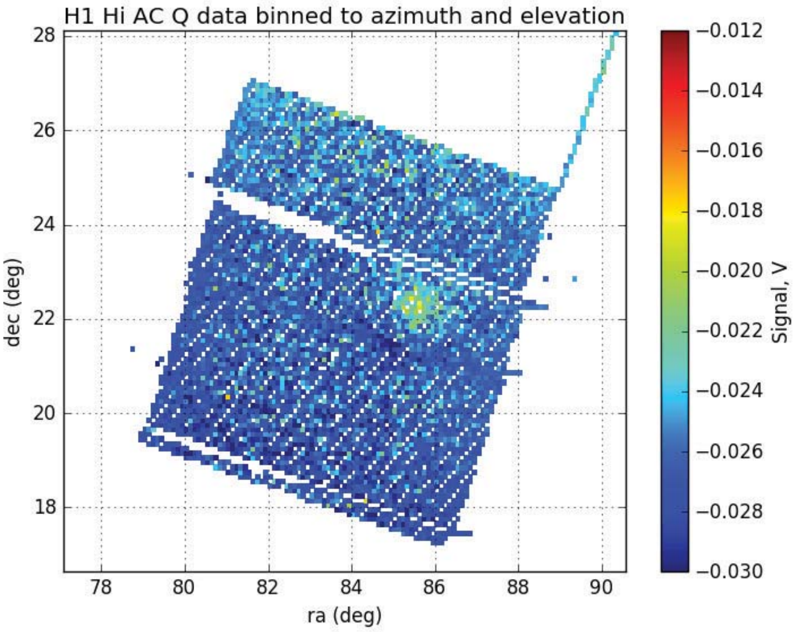
Figure 4.7: View at artificial microwave source reveals sidelobes.

4.5.3 Test measurements and pointing calibration

In order to test the functionality of the telescope, it is common practice to look for known sources on the sky and check if these actually appear in the measurement. This is done by consulting online catalogs that state the time-dependent position of certain celestial objects on the sky. By pointing the telescope into the given direction and scanning the field around it, we can test both whether the pointing is well calibrated and how sensitive the instrument is. Figure 4.8b shows a successful detection of the moon in temperature. The Crab Nebula (Tau A) is a super nova remnant that emits synchrotron radiation. It is one of the few known polarized microwave sources. Figure 4.8a shows a successful detection seen through the Stokes parameter Q.



(a) Moon observed in temperature.



(b) Crab nebula (Tau A) as seen in polarization.

Figure 4.8: Test measurements on known celestial objects.

4.6 Summary and outlook

Observing from Greenland has proven to be challenging but not impossible. While the polar desert seemed promising for long-term observations, strong winds and freezing fog hampered these attempts. Also the number of cloudless days were surprisingly few and rather the exception than the rule. As it turned out, the entire summer of 2018 was rather the exception than the rule. Weather statistics showed later that both wind speeds and cloud coverage were exceptionally high that year.

Not only the weather, but also the telescope turned out to be a challenge. While it was working smoothly towards the end, there were several little but essential things that took time to fix. Probably, all these faults were the result of vibrations during transportation. The combination of these challenges resulted in a disappointingly small volume of scientific data. However, the telescope has not retired yet. At the time of writing, it is back in California and has only recently received a new primary mirror to make it ready for future observations at White Mountain.

The objective of this chapter was to present the challenges of data acquisition. When done right[8], the result is a large data volume of time-ordered information (i.e. timestreams) that have to be analyzed. Figure 4.9 illustrates random raw time-ordered data from the `GreenPol` mission. The top panel shows the Stokes parameter I, which looks rather well behaved apart from a few spikes[9]. The pointing information, i.e. the galactic latitude θ and longitude ϕ are shown in the two panels below. In the next chapter we will progress in our generic pipeline and see how timestreams are filtered and then mapped onto the celestial sphere. For this demonstration we will turn our attention to a new CMB mission, namely `SPIDER`.

---

[8] And with a little bit of luck.

[9] Pobably due to radio frequency interference.

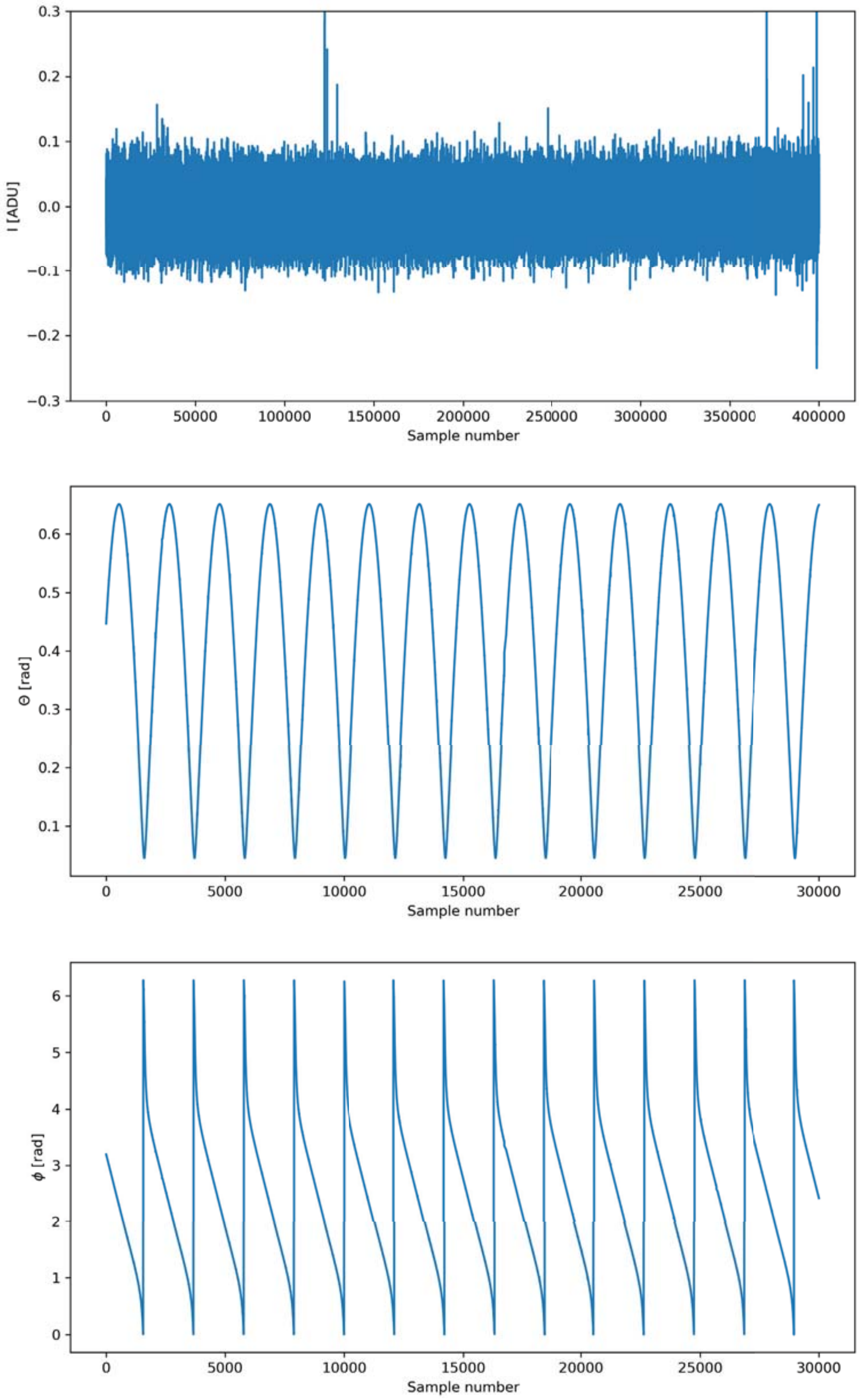


Figure 4.9: Random raw data from `GreenPol` mission. Top: Stokes parameter I. Middle and bottom: Pointing information given by galactic latitude θ and longitude ϕ

(a) Deployed telescope (b) Telescope on rails

(c) Transport procedure (d) Freezing fog on half-wave plate

(e) Office container (f) Main hub and common room

Figure 4.10: On-site impressions

Chapter 5

SPIDER: From Timestreams to Maps

`SPIDER` is a CMB telescope designed to target the polarized microwave sky on the hunt for the inflationary B-mode signature. As opposed to `GreenPol` presented in the previous Chapter, `SPIDER` concentrates on the southern celestial hemisphere and observes at an even higher altitude, namely from a balloon. Two flights were planned, of which the first was successfully performed in January 2015. It took off on 1 January 2015 from McMurdo station (see Fig. 5.2b) and set out on a 16-day circumpolar orbit at an average altitude of 36 km. The flight resulted in a covered sky fraction of ∼10 % and produced ∼1.6 TB of data which are currently being analyzed. The second flight is still in the planning phase. Figure 5.1 shows the expected E and B-mode spectra for $r = 0.03$, along with the expected sensitivities of `SPIDER's` frequency channels, based on two successful flights. Both flights combined are expected to achieve a sensitivity high enough to constrain the acoustic peak of the B-mode signal at $\ell \sim 80$ down to $r < 0.03$ with 3σ confidence [48]. For comparison, the current best limit on r is $r < 0.07$, which is the result from the combined efforts of BICEP2 and Keck [22].

5.1 The instrument

A schematics of the `SPIDER` instrument is pictured in Fig. 5.2a. It employs six monochromatic refracting telescopes housed in a liquid helium cryostat. Due to the presence of polarized foregrounds, three of the receivers work at 90 GHz, while the remaining three are designed to work at 150 GHz. Each telescope projects the light onto a focal plane unit (FPU), which consists of a grid of *transition edge sensor (TES) bolometers*. These bolometers are coupled to a network of orthogonal dipole antennas and are therefore sensitive to two polarization orientations. Each FPU is cooled to roughly 300 mK. All FPUs combined amount to a total of 2400 TESs, of which 96 are intentionally isolated from antennas and are meant to monitor other sources of detector response [49]. In order to distinguish the polarized sky signal from any unwanted polarized instrumental effects, each telescope aperture is fitted with a half-wave plate that is periodically stepped in orientation [48].

The follow-up `SPIDER2` flight is currently being prepared. The instrument from the first flight has been recovered and large parts will be reused. There will however be significant changes, such as the addition of three 280 GHz receivers that will map the same area of the sky but between the frequencies of the `Planck`

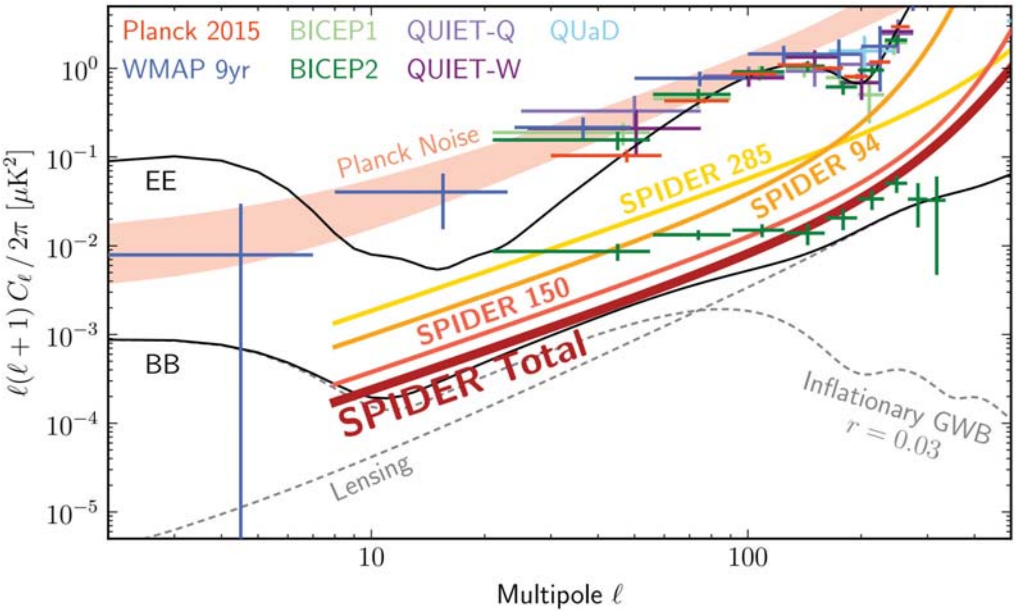
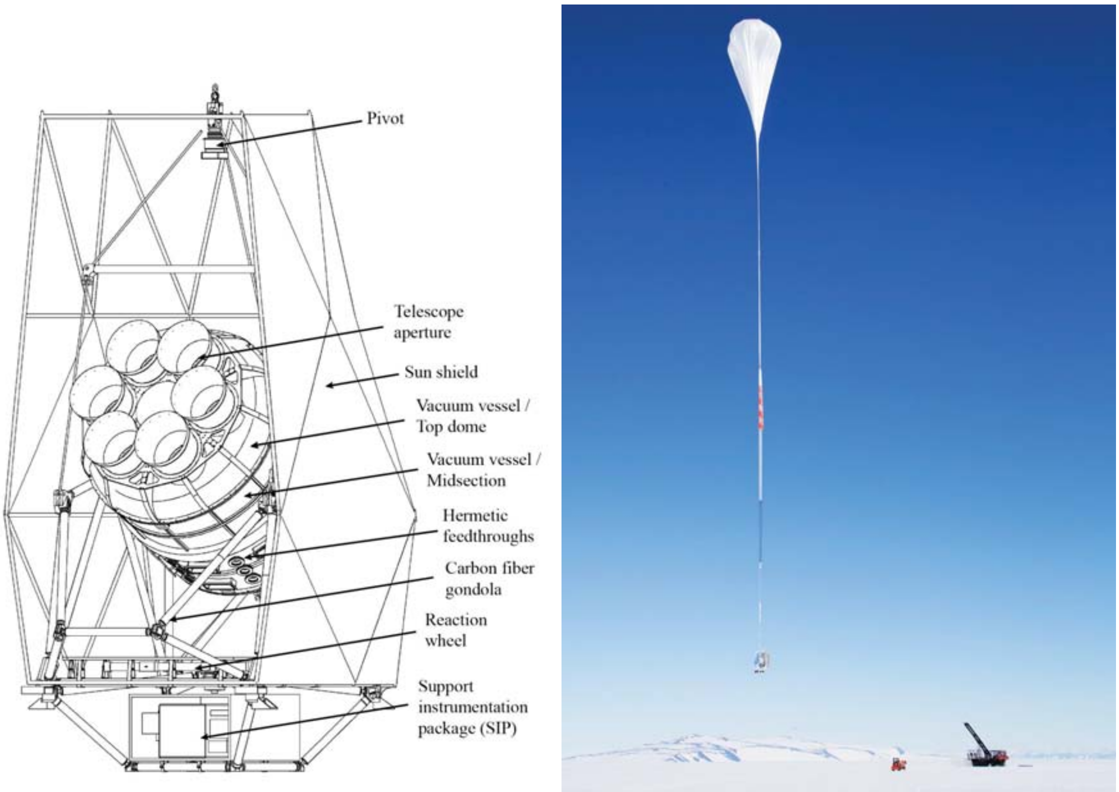


Figure 5.1: Expected E and B-mode spectra for $r = 0.03$, along with sensitivity of `SPIDER` and preceding experiments. Figure courtesy of [49].

channels [48]. This will allow for a better constraint on polarized thermal dust emission.

5.2 Data filtering

The bolometers used in the instrument are highly sensitive detectors that react not only to microwave photons, but also to many other effects whose detection is unintended. This leads to a multitude of different artefacts in the timestreams that have to be corrected or removed before analyzing the data. Figure 5.3 shows a random bolometer readout for a duration of 2 hours (top) and two minutes (bottom). On long time scales, the actual sky signal is dominated by huge *jumps in the offset level*. These have various causes but are all related to large signals that trigger extreme system responses. Another visible artefact are large spikes that are either the result of cosmic rays or communication transmissions when telemetry data are sent to the ground. On short time scales, the data are dominated by the same spikes, but also by a *sinusoidal pattern* caused by the CMB dipole. Another artefact, maybe less obvious in the figure, is a *drift* in the mean offset level which makes the signal slope to the right. Drifts are the result of minute temperature variations of the instrument. The pre-processing of the raw `SPIDER` data by the main analysis team applies a certain amount of flagging, which is nothing else than a data selection process that marks samples that are rejected from the data. Figure 5.3b shows a 2 minute section of the data, with flagged data plotted in green, and the remaining good data in blue. After applying the full flagging, which removes cosmic ray spikes and many other effects, one is left with timestreams that are dominated by the following



(a) Figure courtesy of [50] (b) Figure courtesy of Edward Young

Figure 5.2: (a) Schematics of `SPIDER's` cryostat mounted inside the gondola. (b) Launch from McMurdo station on a NASA balloon

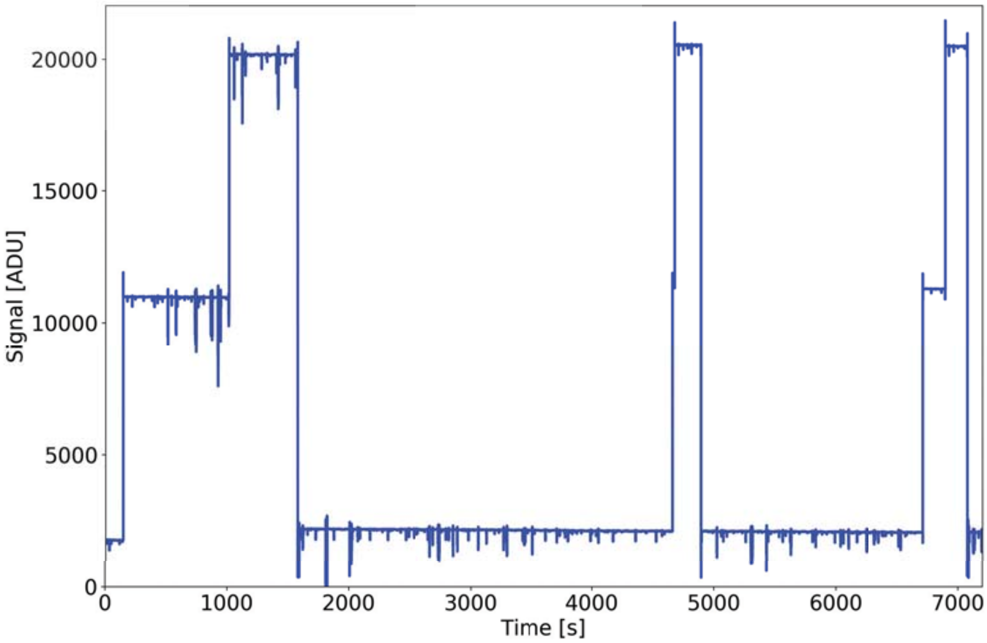
three large scale artefacts:

1. Jumps
2. Drifts
3. Sinusoidal modes

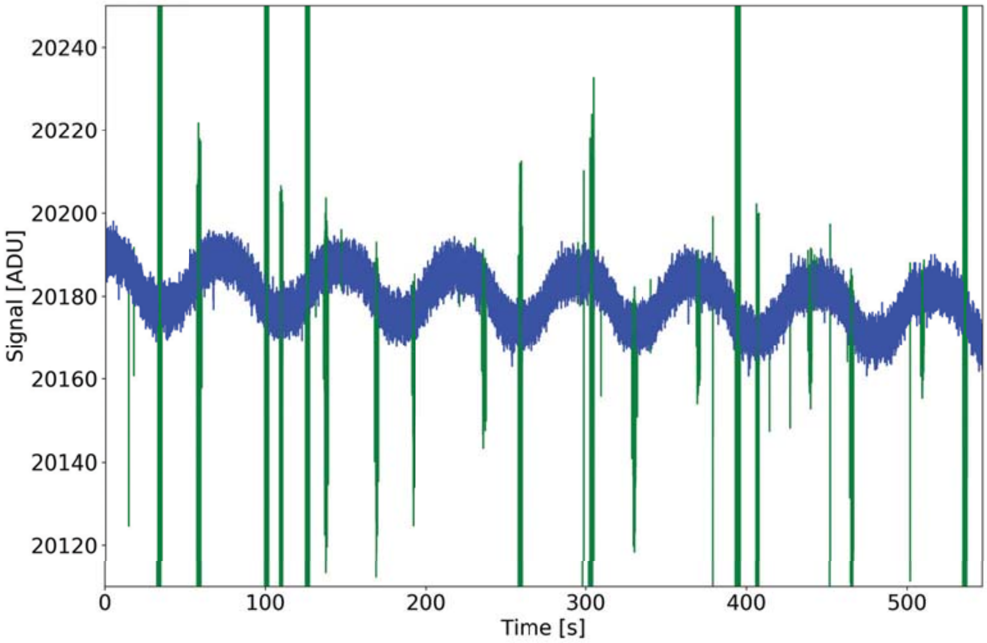
A common solution to problems similar to the one at hand is called *destriping*, a method that was designed to handle correlated low-frequency noise while combining the stages of filtering and mapmaking. The basic idea of destriping is that the correlated component can be approximated by a series of offset functions, which are then subtracted from the timestreams, thereby whitening the noise [51]. The underlying assumption is that the time-ordered data vector $\mathbf{d}$ can then be expressed as

$$\mathbf{d} = \mathbf{Pm} + \mathbf{n}_{\text{corr}} + \mathbf{n}_{\text{w}}, \tag{5.1}$$

where $\mathbf{n}_{\text{w}}$ and $\mathbf{n}_{\text{corr}}$ are vectors containing ordinary white noise and correlated noise respectively. Furthermore, $\mathbf{m}$ is the true sky map and $\mathbf{P}$ the *pointing matrix*, which encapsulates information about the position on the sky for every measured data sample. More information on $\mathbf{P}$ will be given later in Sec. 5.3.



(a) 2 hours of raw `SPIDER` data



(b) 2 minutes of raw `SPIDER` data

Figure 5.3: Illustration of large scale contamination in `SPIDER's` raw data.

The key idea of destriping is that the correlated noise component can be expressed as a series of offsets given by

$$\mathbf{n}_{\mathrm{corr}} = \mathbf{T}\mathbf{a}, \tag{5.2}$$

where $\mathbf{T}$ is a template matrix that defines the offsets and $\mathbf{a}$ contains the corresponding amplitudes. It is then possible to solve jointly for the offset amplitudes $\mathbf{a}$ and the sky signal $\mathbf{m}$. Unfortunately, the huge jumps in `SPIDER's` raw data make it difficult to directly apply the destriping approach to the data set at hand. Since these jumps happen randomly, they cannot be accounted for by offset templates of static length. To circumvent this problem, we will apply a similar approach, but use variable offset lengths instead. Furthermore, because of the novelty of the problem, we will perform the filter and mapmaking stages completely independently from each other, while postponing a joint solution to the future. In Sec. 5.2.1 we will introduce a template-fitting algorithm specifically tailored to the `SPIDER` data set. The mapmaking problem awaits us in Sec. 5.3.

5.2.1 Large-scale template fitting

Let's consider a raw time-ordered data vector $\mathbf{d}$ of length n_t that contains the sky signal and all of the above-mentioned large-scale systematic contaminations. In order to clean the raw timestream, we assume that it can be expressed as the sum of a large-scale error term $\mathbf{d}_{\mathrm{LS}}$ and the small-scale sky signal $\mathbf{s}$. The simple data model therefore reads as

$$\mathbf{d} = \mathbf{d}_{\mathrm{LS}} + \mathbf{s}. \tag{5.3}$$

Furthermore, we assume that the large-scale error term can be expressed as a linear combination of n_a templates given by

$$\mathbf{d}_{\mathrm{LS}} = \mathbf{T}\mathbf{a}, \tag{5.4}$$

where $\mathbf{T}$ is a $(n_t \times n_a)$-matrix containing the templates in its columns and $\mathbf{a}$ is the coefficient vector for these templates. The small scale sky signal that we are interested in is therefore

$$\mathbf{s} = \mathbf{d} - \mathbf{T}\mathbf{a}. \tag{5.5}$$

Our task now is to build the template matrix $\mathbf{T}$ and to solve for the coefficients $\mathbf{a}$.

Templates

In the following we will need templates whose linear combination lets us describe the above-mentioned large scale contaminations. First, we have to account for changing offset levels, where each offset is separated by a jump. Second, each offset level is characterized by a drift. Third, we have sinusoidal patterns covering the full extent of the timestream.

In order to handle the offsets and the corresponding drifts, we will scan the timestream for jumps and place an offset boundary at each detection. The timestream is therefore split into subsections, each subsection corresponding to an offset level. We consequently have to consider $(n_{\text{jump}} + 1)$ of these offset regions. In the next step we will create two templates for each offset region. The first characterizes the offset level and has a general amplitude of 1 between two jumps and is 0 otherwise. The second template describes the drift within the offset region. It is also 0 outside the considered offset region, but within it has a general slope, starting at -1 at the left boundary and increasing up to 1 at the right boundary.

Lastly, we need to describe any arbitrary pattern in the timestream up to some frequency $\nu_{\max}$. This will allow for the removal of the CMB dipole from the data, but also for any other effect that results in frequency patterns in the timestreams. To facilitate this in the template fit, we create templates for each *Fourier mode* up to the wave number $k_{\max}$ that corresponds to $\nu_{\max}$. Our Fourier templates therefore simply read

$$\mathbf{C}_k = \cos\left(k \cdot 2\pi \frac{t}{T}\right) \tag{5.6}$$

$$\mathbf{S}_k = \sin\left(k \cdot 2\pi \frac{t}{T}\right), \tag{5.7}$$

where t is the time stamp of the data and T is the total duration of the timestream considered. Altogether we are left with

$$n_a = 2(n_{\text{offset}} + n_{\text{harmonic}}) = 2(n_{\text{jump}} + k_{\max} + 1) \tag{5.8}$$

templates that have to be fitted to the raw data. The template fitting method is most easily explained by the aid of an example. Let's assume we have a timestream in which we happen to detect two jumps. The templates that correspond to this model example are shown in Fig. 5.4, where we for simplicity only consider the first two Fourier modes. The lines in black represent the offset templates, and the corresponding drift templates are plotted in blue. The solid (dashed) red lines are the cosine modes (sine modes) for $k = 1$ and $k = 2$, while the vertical lines in orange mark the offset boundaries, i.e. the positions of the jumps. The full linear combination for the problem at hand therefore reads

$$\begin{aligned} \mathbf{d}_{\text{LS}} &= \mathbf{T}\mathbf{a} && (5.9) \\ &= [\mathbf{O}_1 \;\; \mathbf{O}_2 \;\; \mathbf{O}_3 \;\; \mathbf{D}_1 \;\; \mathbf{D}_2 \;\; \mathbf{D}_3 \;\; \mathbf{C}_1 \;\; \mathbf{C}_2 \;\; \mathbf{S}_1 \;\; \mathbf{S}_2]\, \mathbf{a}, && (5.10) \end{aligned}$$

where $\mathbf{O}_i$ and $\mathbf{D}_i$ denote the offset and drift templates respectively and $\mathbf{C}_k$ and $\mathbf{S}_k$ the Fourier modes, which act as a Fourier filter. The coefficient vector $\mathbf{a}$ contains a single number for each template.

The number of Fourier modes to include in the fit is related to the highest frequency component to be filtered, namely $\nu_{\max}$. The higher we choose this frequency, the cleaner the data, but also the higher the loss of cosmological signal. Apart from the dipole in the data, there is a significant amount of scan-synchronous noise at higher frequencies, which is connected to the scanning frequency ν_{scan}. At the current stage of the analysis, we opted for a maximum frequency of $\nu_{\text{max}} = 4\nu_{\text{scan}} \approx 0.06$ Hz, which results in well above 1000 Fourier modes. In order to reduce computation time, we downsample the data by a factor of 100, i.e. we bin the samples in groups of 100, and apply the template fit to these downsampled timestreams. Since we only want to filter large scales, there is no downside to doing that, as long as we subtract the result from the full data.

Jump-finder

The entire template-fitting method crucially relies on correct offset boundary positions given by the jumps. While a false boundary only slightly increases computation time, a missed one will result in poor convergence of the fit. Figure 5.5 demonstrates the basic principle of jump detection. The top panel shows the raw but downsampled timestream, which includes large jumps. These jumps are detected by computing the rolling standard deviation along the timestream, shown by the blue line in the bottom panel. A large jump in the timestream will result in a large spike in the rolling standard deviation. The detection of these spikes is not trivial, since the gain is varying, and therefore, the baseline of the standard deviation is not constant. We generate a baseline by computing the rolling median along the standard deviation, shown by the yellow line. A positive detection of a jump is made when the standard deviation exceeds a certain multiple of the baseline, given by the red line. This multiple is simply a threshold and chosen to be 3 in this analysis. The top panel in Fig. 5.5 shows how well this method performs.

Given a successful jump detection and a subsequent template generation, we are left with the task of solving for the template coefficients $\mathbf{a}$. However, since we need a very large number of templates, depending on the number of jumps in the timestream and the desired maximum frequency ν_{max}, this is a very computationally heavy task.

The conjugate gradient (CG) method

Given the problem at hand,

$$\mathbf{d} = \mathbf{T}\mathbf{a} + \mathbf{s}, \tag{5.11}$$

we need to find the best-fit template coefficients $\mathbf{a}$ before we can solve for the clean data $\mathbf{s}$. Treating $\mathbf{s}$ as a noise term, the general least square estimate of $\mathbf{a}$ reads

$$\mathbf{a} = \left(\mathbf{T}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{T}\right)^{-1} \mathbf{T}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{d}, \tag{5.12}$$

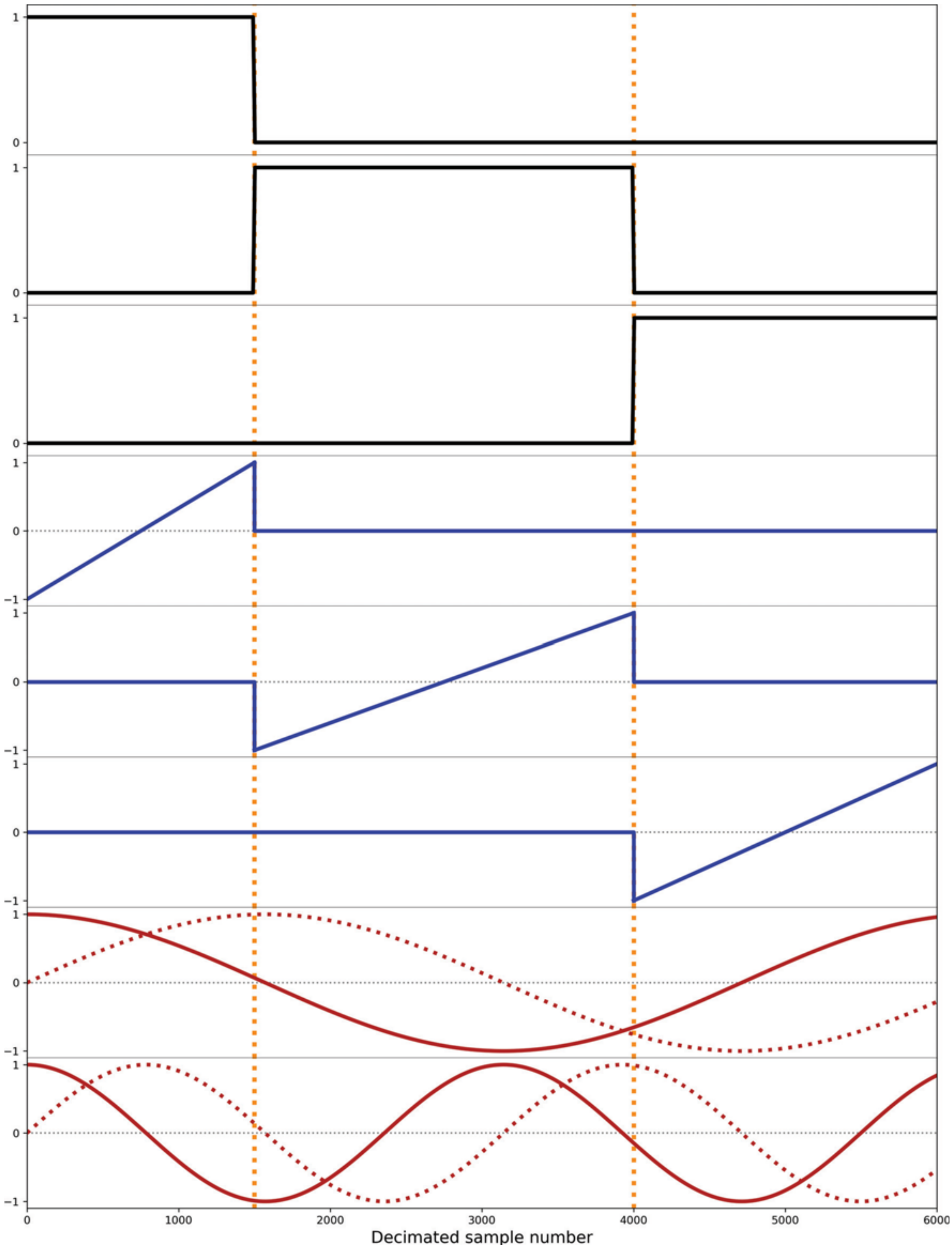


Figure 5.4: Templates for a generic timestream containing two jumps. Black lines represent offset templates. Blue lines represent drift templates. Solid (dashed) red lines represent cosine (sine) modes. Jumps are marked by the vertical dashed lines.

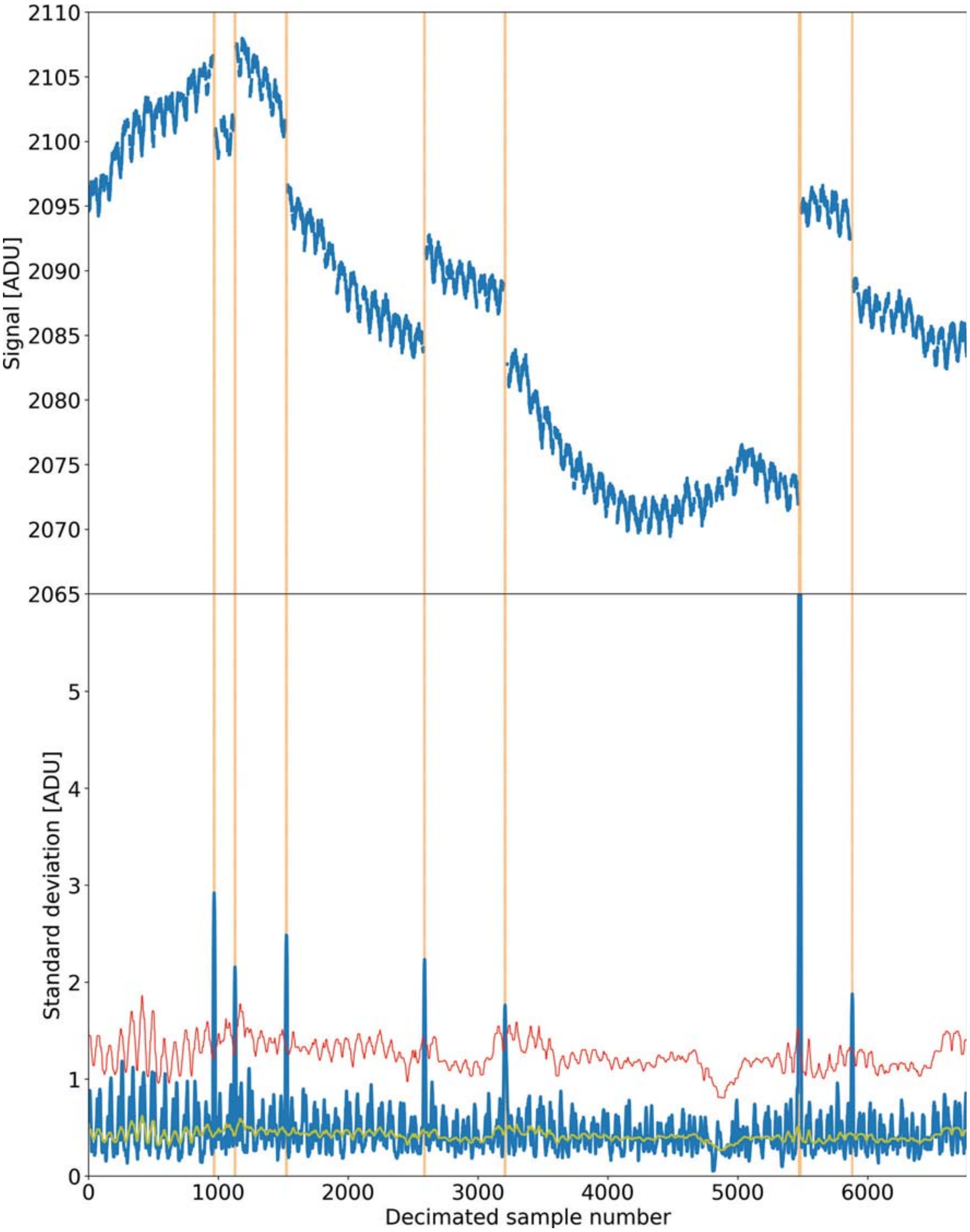


Figure 5.5: Illustration of jump detection. Top: Downsampled raw data. Bottom: Rolling standard deviation (blue), rolling median (yellow) and a multiple of it (red). Detected jumps are marked by vertical lines.

where $\mathbf{N}$ is the noise covariance matrix. At this stage of the analysis, no noise estimation has been performed, which is why we will neglect this term for now and simply treat it as a diagonal matrix. The template matrix $\mathbf{T}$ is a huge object with a dimension of $(n_t \times n_a)$, which makes the computation of the inverse term after the equality sign a challenge. Therefore we rewrite Eq. 5.12 as

$$\left(\mathbf{T}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{T}\right)\mathbf{a} = \mathbf{T}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{d} \tag{5.13}$$

and recognize that we are working with a *linear system of equations* of the form

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{5.14}$$

to which there exist numerical solutions. One method to solve these is the *conjugate gradient method*, which we will sketch by following closely the description by [52].

The conjugate gradient method is a powerful algorithm that allows to solve linear systems of equations as given in Eq. 5.14. Here, $\mathbf{A}$ is a known square matrix[1], $\mathbf{x}$ is an unknown vector and $\mathbf{b}$ is a known vector. There are certain requirements on the properties of $\mathbf{A}$, such as that it has to be postitive-definite. The mathematical background of this method is fully based on quadratic forms, which are scalar quadratic functions of vectors with the form

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} - \mathbf{b}^{\mathrm{T}}\mathbf{x}. \tag{5.15}$$

Mathematicians have shown that the minimum point of $f(\mathbf{x})$ is the solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$. This means that setting the gradient

$$f'(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b} \tag{5.16}$$

to zero solves the linear system of equations. The expression $\mathbf{A}\mathbf{x} - \mathbf{b}$ is also called the “residual” r, and it is the realization that the residual can be used as a direction towards the solution that makes the method of conjugate gradients possible. For a full overview of the method and a pseudo code, I again refer the reader to [52]. The take-home message is that the method of conjugate gradients is an iterative approach, which starts out with a first guess for $\mathbf{x}$, followed by the computation of the residual. This residual will be used to propose a new vector $\mathbf{x}_{i+1}$, which moves the residual r closer to the minimum of Eq. 5.16.

Clean data

When applying the conjugate gradient method to our problem in Eq. 5.13, three things are important to note. First, in order to significantly improve the convergence speed of the algorithm, it is common practice to apply a *preconditioner* $\mathbf{P}$ which transforms $\mathbf{A}$ in such a way that $\mathbf{P}^{-1}\mathbf{A}$ has a smaller condition number. In

---

[1] In our case a $(n_a \times n_a)$-matrix.

our case, we apply a simple diagonal preconditioner $\mathbf{P} = \mathrm{diag}(\mathbf{A}) = \mathrm{diag}(\mathbf{T}^{\mathrm{T}}\mathbf{T})$, which performs effectively. Second, since offsets and Fourier modes with similar dimensions are degenerate, we first fit only for offsets and drifts and use the result as a prior for the full fit. Third, the transition region between jumps makes it necessary to apply a certain amount of additional flagging on the timestreams. Furthermore, the solution for some sections of the timestreams does not converge for reasons unknown at this point. This has the effect that a very small percentage of the data is rejected by the filter, which is a total of approximately 5 %. The top panel in Fig. 5.6 shows the same data segment as from Fig. 5.5, with the large-scale fit applied. The fit was first performed on the downsampled data, and then upsampled to match the full sample data. The bottom panel shows the small scale signal $\mathbf{s}$ when subtracting the large scale error, as given by Eq. 5.5. The resulting timestream is well-behaved and free from jumps, drifts and other obvious contaminations.

5.3 Mapmaking

The result from the large-scale filter is well-behaved timestreams in intensity and corresponding timestreams of the pointing direction given by longitude ϕ and latitude θ. In addition, we have a timestream of the polarization angle ψ of the detector with respect to some reference frame. We are now left with the task of projecting the data onto a map. At this point it has to be noted that the moon and Tau A maps presented earlier in the Chapter 4 (Fig. 4.8) are not considered as “real” sky maps, since each pixel was only hit once and the mapmaking process simply amounted to plotting azimuth versus elevation while the color code was given by the signal. This is a very rudimentary mapmaking process and only works if each pixel was only hit once, which stands in stark contrast to the thousands of hits that we have to deal with in the final mapmaking stage. In the following we will look at real mapmaking solutions that facilitate a full mapping of the time-ordered data.

The starting point for mapmaking is a cleaned data vector $\mathbf{d}$ of length n_t. Given the polarization sensitivity of the instrument, each measured sample d_t can be expressed as

$$d_t = I_t + Q_t \cos 2\psi_t + U_t \sin 2\psi_t + n_t, \tag{5.17}$$

where I, Q and U are the Stokes parameters and n_t is a noise term. The full data vector $\mathbf{d}$ can be compactly expressed in the form

$$\mathbf{d} = \mathbf{P}\mathbf{s} + \mathbf{n}. \tag{5.18}$$

Here, $\mathbf{P}$ is the so-called *pointing matrix*, which contains the information of the position for each individual sample d_t on the sky. It has dimension $(n_t \times 3n_{pix})$, where n_{pix} is the number of pixels on the sky, which is directly related to the resolution we are working with. We get the factor 3 from the fact that we are working with three Stokes parameters. Each row of $\mathbf{P}$ is therefore of length

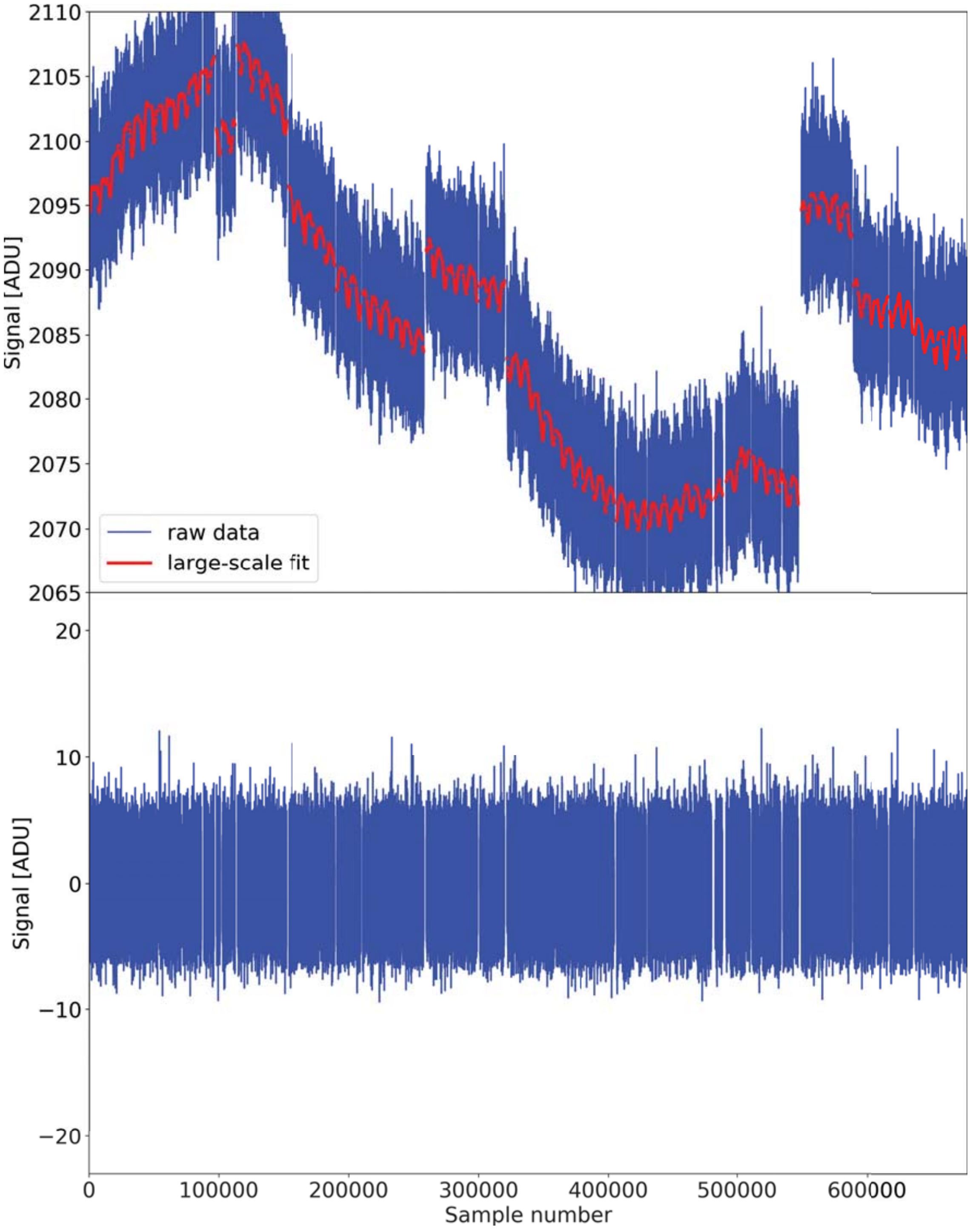


Figure 5.6: Top: Raw data (blue) with applied large-scale fit (red). Bottom: Raw data with large-scale fit subtracted.

$3n_{pix}$ and encodes the pixel number hit at any time t in the following way: the first n_{pix}-entries are reserved for the pointing of Stokes parameter I, while the remaining entries belong to Q and U. Each group of n_{pix} pointing elements contains only zeros, except for the element that corresponds to the pixel number hit at time t. For Stokes parameter I the non-zero value is simply 1, while for Q and U it is $\cos 2\psi_t$ and $\sin 2\psi_t$ respectively. The sky map that we are interested in is given by $\mathbf{s}$, and finding its solution is the goal of mapmaking. It has the dimension $3n_{pix}$, the factor 3 resulting again from the fact that we will solve for three sky maps, one for each Stokes parameter. The map vector $\mathbf{s}$ therefore has the structure

$$\mathbf{s} = \left(I_{p_1} \ldots I_{p_n}, \quad Q_{p_1} \ldots Q_{p_n}, \quad U_{p_1} \ldots U_{p_n}\right)^{\mathrm{T}}. \tag{5.19}$$

Lastly, $\mathbf{n}$ is a noise vector with zero mean $\langle \mathbf{n} \rangle = 0$ and noise covariance given by $\mathbf{N} = \langle \mathbf{n}\mathbf{n}^{\mathrm{T}} \rangle$. In the following we will first look at the most intuitive method of mapmaking, which puts strong and often unrealistic assumptions on the noise properties, but is widely used in CMB analysis to get a quick picture of the sky. This quick and handy mapping technique is called *binned mapmaking.*

5.3.1 Binned mapmaking

Under the assumption that the noise $\mathbf{n}$ is *Gaussian and uncorrelated*, we can compute the average of all samples that are mapped into one pixel. The least square estimate for the sky map $\mathbf{s}$ in Eq. 5.18 then reads as

$$\hat{\mathbf{s}} = \left(\mathbf{P}^{\mathrm{T}}\mathbf{P}\right)^{-1} \mathbf{P}^{\mathrm{T}}\mathbf{d}. \tag{5.20}$$

This equation has a very intuitive interpretation. The last term, $\mathbf{P}^{\mathrm{T}}\mathbf{d}$, maps all data samples of $\mathbf{d}$ into their respective pixels. The result has therefore the dimension of a map[2], but each pixel value is given by the sum of all samples mapped into the respective pixel. In order to get a more appropriate representation of the sky, we'll have to divide by the number of samples mapped into each pixel. This quantity, i.e. the number of observations at a certain pixel position, is given by the term in front, namely $\mathbf{P}^{\mathrm{T}}\mathbf{P}$. This term also bears the name *hitmap.* Hence, division by the hitmap results in the binned solution $\hat{\mathbf{s}}$. In practice, Eq. 5.20 is solved individually for each pixel and can be expressed in an analog per-pixel representation given by

$$\begin{pmatrix} I_p \\ Q_p \\ U_p \end{pmatrix} = \left[\sum_{t \in p} \begin{pmatrix} 1 & c_t & s_t \\ c_t & c_t^2 & c_t s_t \\ s_t & c_t s_t & s_t^2 \end{pmatrix}\right]^{-1} \left[\sum_{t \in p} \begin{pmatrix} d_t \\ d_t c_t \\ d_t s_t \end{pmatrix}\right], \tag{5.21}$$

where c_t and s_t denote $\cos 2\psi_t$ and $\sin 2\psi_t$ respectively [53, 54].

---

[2] Which may or may not contain polarization.

Preliminary maps

Now we are ready to apply the binned mapmaking method on the filtered data. The resulting maps for Stokes parameters I, Q, and U are presented in Figures 5.7a-5.7c. These are *preliminary maps* and are not final data products. They show however the current state of data filtering. While the I map looks perfectly clean, the polarization maps show large scale disturbances concentrated between galactic latitude -45 and -30. Because of the low amplitude of polarization maps, any systematic errors will show up a lot more clearly than in the I map. Nevertheless, the presented results look rather promising, but they seem to require a more stringent flagging approach. Presumably, the residual large scale structure is the result from the template fit that performed badly in parts of the data. Some of these poorly cleaned data sections seem to be missed by the subsequent flagging sequence and therefore slip into the mapmaker. These need to be detected and rejected. A future extension of the filtering approach requires the characterization of *correlated noise* in the timestreams, which will enable a more elaborate mapmaking technique. This method is called *optimal mapmaking* and will be described in the following.

5.3.2 Optimal mapmaking

The most general method of mapmaking is *optimal mapmaking*, also called *maximum likelihood mapmaking.* The assumption is that the noise is Gaussian, but it is more general than the binned mapmaking approach in the sense that it allows for the noise to be correlated. Any information about the correlation is then given in the noise covariance matrix $\mathbf{N}$. Let's rewind to the data model from Eq. 5.18,

$$\mathbf{d} = \mathbf{P}\mathbf{s} + \mathbf{n}, \tag{5.22}$$

and rewrite it as

$$\mathbf{n} = \mathbf{d} - \mathbf{P}\mathbf{s}. \tag{5.23}$$

Under the assumption that the noise is Gaussian, we can describe the probability of the noise as a multivariate Gaussian distribution,

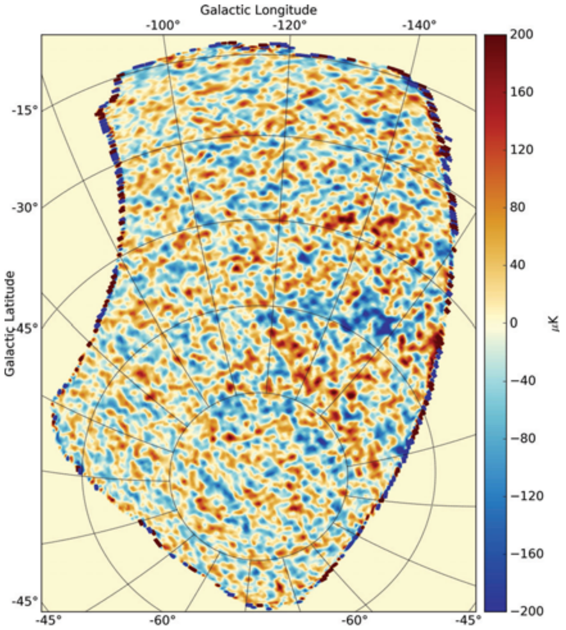
$$P(\mathbf{n}) \propto \exp\left[-\frac{1}{2}\mathbf{n}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{n}\right], \tag{5.24}$$

or alternatively using Eq. 5.23,

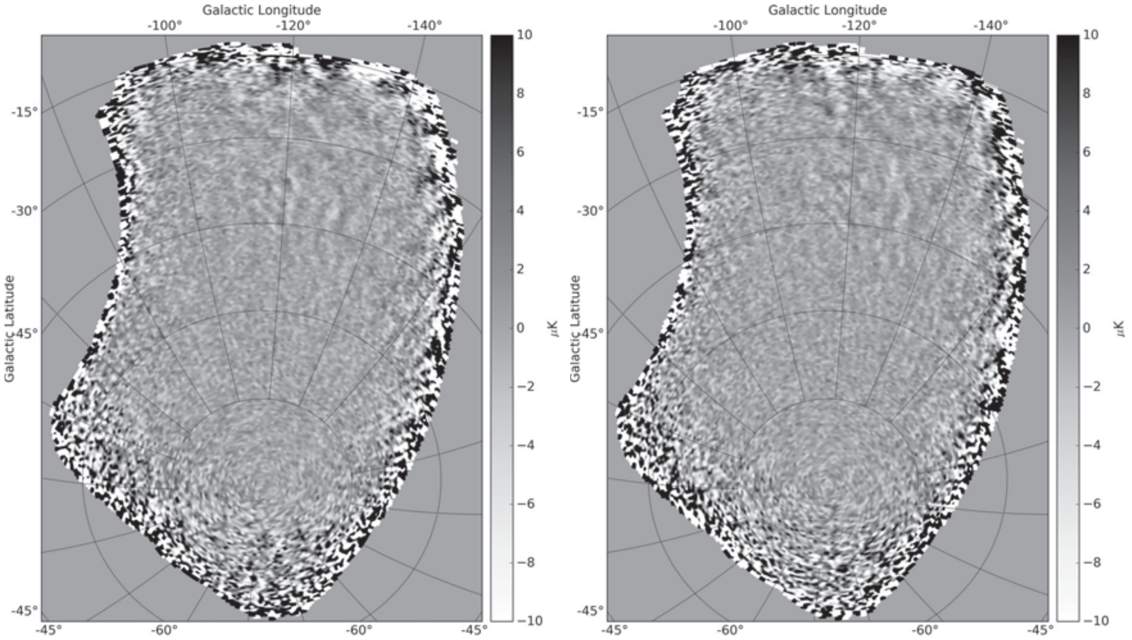
$$P(\mathbf{s}|\mathbf{d}) \propto P(\mathbf{d}|\mathbf{s}) \equiv \mathcal{L}(\mathbf{d}) \propto \exp\left[-\frac{1}{2}\left(\mathbf{d} - \mathbf{P}\mathbf{s}\right)^{\mathrm{T}}\mathbf{N}^{-1}\left(\mathbf{d} - \mathbf{P}\mathbf{s}\right)\right]^{\mathrm{T}}, \tag{5.25}$$

where we have used Bayes' theorem in connection with a uniform prior on $\mathbf{s}$, meaning we have no prior assumptions on how the sky looks like. We find the solution $\hat{\mathbf{s}}$ by maximizing the likelihood $\mathcal{L}$ or equivalently by minimizing the corresponding χ^2. The solution to this is given by

$$\hat{\mathbf{s}} = \left(\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{d}. \tag{5.26}$$



(a) Stokes parameter I



(b) Stokes parameter Q

(c) Stokes parameter U

Figure 5.7: Binned `SPIDER` maps @150 GHz

Noise weighting

In the following we will show how to facilitate correlated noise in the mapmaking process. The noise is assumed to follow a very common $1/f$ power distribution,

$$N(\nu) = \sigma_0^2 \left[1 + \left(\frac{\nu}{f_{\mathrm{knee}}} \right)^{\alpha} \right], \tag{5.27}$$

which is described by a high frequency noise floor σ_0, and a low frequency correlated component characterized by f_{knee} and a slope α. These parameters have to be fitted to the timestreams in frequency domain. Since no noise characterization has been performed at this stage of the analysis, we use filtered data and noise parameters from our former colleagues[3]. To avoid confusion, we will refer to these older version data as the 2017-data set. This set of data has been produced by filtering the same raw data with the same template fitting approach, but using a different implementation and solving technique. The resulting filter is more aggressive and leads to a significantly larger amount of rejected data.

In order to filter out the noise contribution given by Eq. 5.27, we rewrite Eq. 5.26 for reasons soon apparent, and also explicitly write out Fourier transformations $\mathcal{F}$ where required. This new expression reads

$$\left(\mathbf{P}^{\mathrm{T}} \mathcal{F}^{-1} \mathbf{N}^{-1} \mathcal{F} \mathbf{P} \right) \mathbf{s} = \left(\mathbf{P}^{\mathrm{T}} \mathcal{F}^{-1} \mathbf{N}^{-1} \mathcal{F} \right) \mathbf{d}. \tag{5.28}$$

Since our noise description is given in frequency space, it is necessary to transform timestreams such as $\mathbf{d}$ or $\mathbf{Ps}$ into Fourier domain, then divide out the noise contribution, and finally transform back into time domain. Hence, the expression $\mathcal{F}^{-1}\mathbf{N}^{-1}\mathcal{F}$ represents a filter operator in frequency domain, which filters out $1/f$ noise contributions given by Eq. 5.27 and the corresponding fitted parameters f_{knee}, σ_0 and α. The problem at hand is to solve for the sky map $\mathbf{s}$. When examining Eq. 5.28 more closely, one might guess the motive of rearranging the terms of the equation. The reason is that the equation above has exactly the form

$$\mathbf{Ax} = \mathbf{b}, \tag{5.29}$$

which we have encountered earlier in Sec. 5.2.1. The equation above can therefore again be solved iteratively using the *conjugate gradient method*.

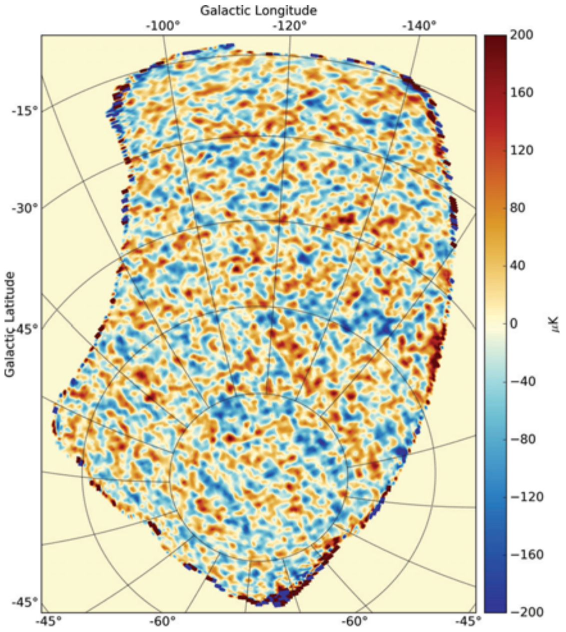
The result of solving Eq. 5.28 above is presented in Fig. 5.8. While the I-map looks virtually identical, differences can be spotted in the polarization maps. Since the data that underlie these maps is filtered more stringently, they look cleaner while slightly more noisy because of the loss of data.

5.4 Validation

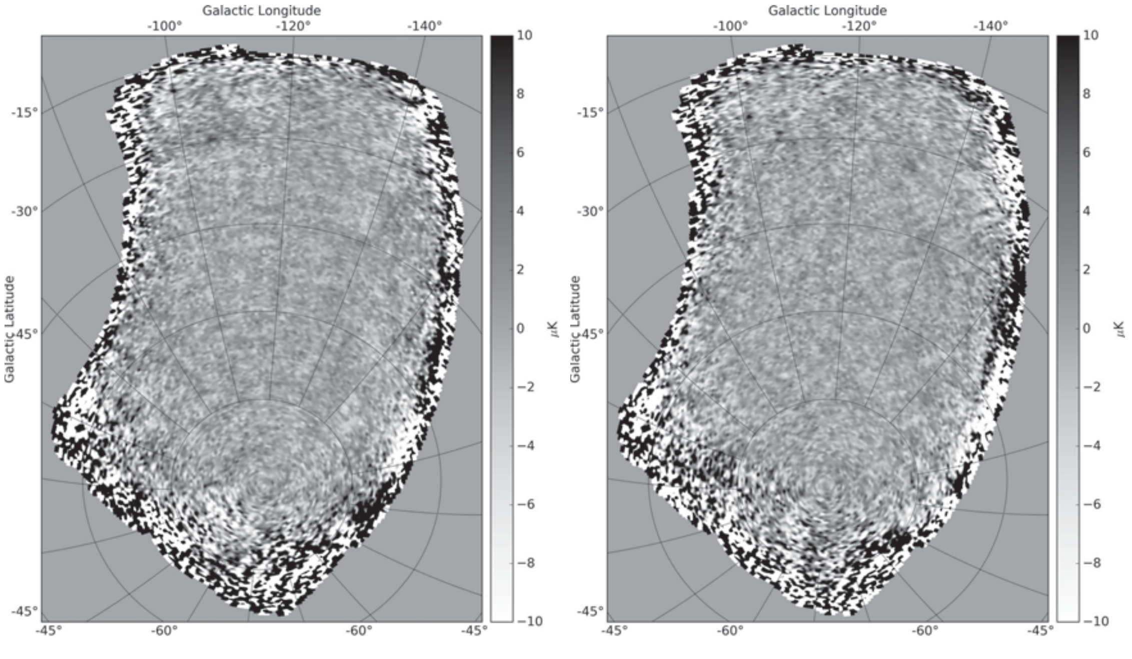
While the visual appearance of a CMB map gives a first impression of the quality of the underlying data, the validation of maps is ultimately done in power

---

[3] Data courtesy of T. M. Ruud and B. Racine.



(a) Stokes parameter I



(b) Stokes parameter Q (c) Stokes parameter U

Figure 5.8: Noise-weighted `SPIDER` maps @150 GHz using the 2017-data set

spectrum domain. This makes it necessary to first characterize the effect the filter has on the power spectrum of the true sky. The measure that quantifies this effect is called a *transfer function*. Since the implemented filter removes contributions below $\nu_{\max}$, in other words it acts as a high-pass filter, power spectra computed from these maps are consequently biased. To take this effect into account, one can propagate a known sky signal through the pipeline and measure the ratio between the output and the input power spectra. Ideally, this procedure starts with drawing an ensemble of CMB realizations from a theoretical power spectrum C_ℓ^{theory}. Then, one converts these maps into timestreams by "reobserving" them with the real pointing information of the `SPIDER` instrument. These are then propagated through the filtering and mapmaking stages and finally compared to the theory at power spectrum level. The transfer function is then given by

$$F_\ell = \frac{\langle C_\ell^{\text{sim}} \rangle}{C_\ell^{\text{theory}}}. \tag{5.30}$$

Since this ideal procedure is very computationally heavy, we obtained an approximation of this quantity by reobserving a single `Planck` CMB map. The transfer function was estimated by propagating the resulting timestreams through the entire pipeline and comparing the power spectra of the input and output map,

$$F_\ell = \frac{C_\ell^{\text{output}}}{C_\ell^{\text{input}}}. \tag{5.31}$$

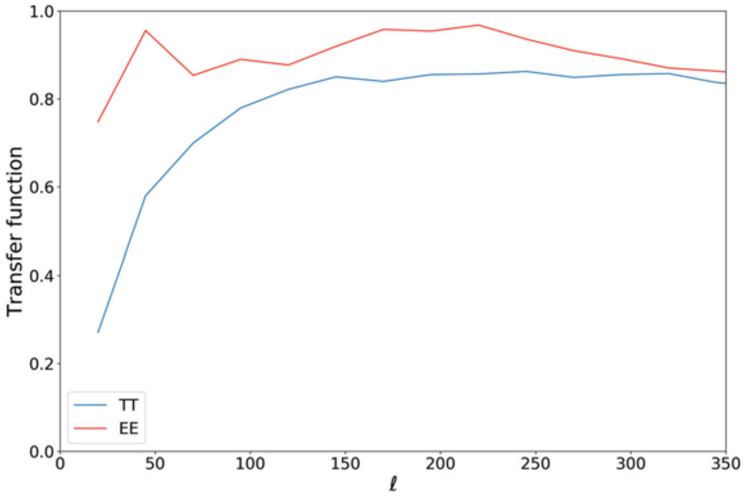
The approximate transfer functions for temperature and E-mode polarization resulting from this approach are plotted in Fig. 5.9a.

We produce noise-weighted maps for each half of the time-ordered data, which we then cross-correlate. The idea behind cross-spectra is that only the signal is correlated between the data splits, while the noise is not, and the resulting spectra should therefore be free of noise bias. Figures 5.9b and 5.9c show cross-spectra between the two `SPIDER` half-maps, corrected for the transfer function. These are then compared to a best-fit `Planck` ΛCDM spectrum. Furthermore, we compute the cross-correlation between the full `SPIDER` map and the reobserved `Planck` map, and also the cross-correlation between two reobserved `Planck` half-mission maps. This test does not show any big discrepancies between data and theory, and therefore serves as a first check of the validity of the maps.

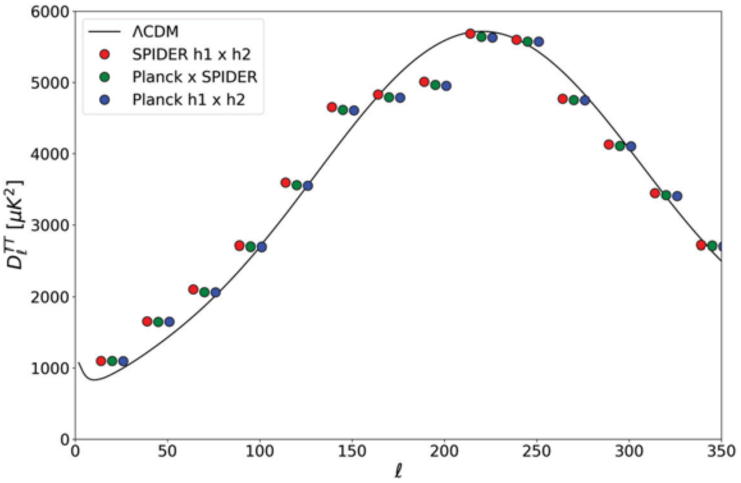
5.5 Summary and outlook

The main goal for the future is to implement characterization of correlated noise into the filter algorithm described in Sec. 5.2.1, which is required to perform optimal mapmaking. The current filter algorithm is an updated version of the one that produced the 2017-data set, in the sense that the latter is much more aggressive and therefore leads to the loss of large amounts of data. Furthermore, the previous implementation did not make use of the power of the CG method, making it at lot more computationally heavy. The new algorithm on the other hand rejects far less data, but unfortunately still leaves some residual large-scale structure in the polarization maps (Fig. 5.7), which requires a more thorough investigation of its cause. Finally, all of the `SPIDER` analysis so far has concentrated on the 150 GHz data. While applying the filter and mapmaking methods on the 90 GHz data set should be straightforward, this task is still outstanding at the time of writing.

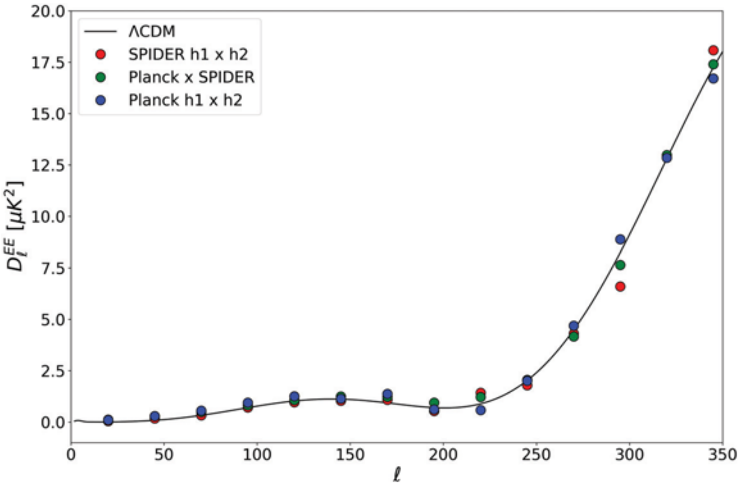
Taking a step back and looking at the big picture, we realize that we have almost reached the end of the generic CMB pipeline. The next task is to free the CMB map(s) from galactic foreground emission, a necessary procedure before cosmological parameter estimation. Since the `SPIDER` maps are not yet mature enough for this step, we now turn our attention to maps that are, namely those of `Planck`.



(a) Transfer function



(b) TT power spectrum



(c) EE power spectrum

Figure 5.9: Power spectra and transfer function for noise-weighted `SPIDER` maps, computed from the 2017-data set. Data points are plotted with an offset in order to improve their visibility.

Chapter 6

Planck: From Maps to Cosmology

`Planck` delivered some of the most valuable microwave sky maps to date, a result owing not only to the ideal observation environment in space, but also to the combined effort of a huge collaboration. In this Chapter we will use the `Planck` frequency maps as our basis and sketch the road ahead towards cosmological parameters. This happens in two steps: first, *component separation* has to be performed to reduce the frequency maps to one single CMB map. The maps we have at this point only show a composition of CMB *and* all the other contaminators presented earlier in Chapter 2. While Chapter 3 described how component separation works in *theory*, this Chapter will demonstrate how it's done in *practice*.

The second step is to use this map to infer cosmological parameters. One of these is the CMB dipole, which will be more thoroughly discussed in Sec. 6.3. For the sake of completeness, a more general approach for the estimation of the remaining parameters will be sketched in Sec. 6.5. But first, we will look at the instrument itself.

6.1 The instrument

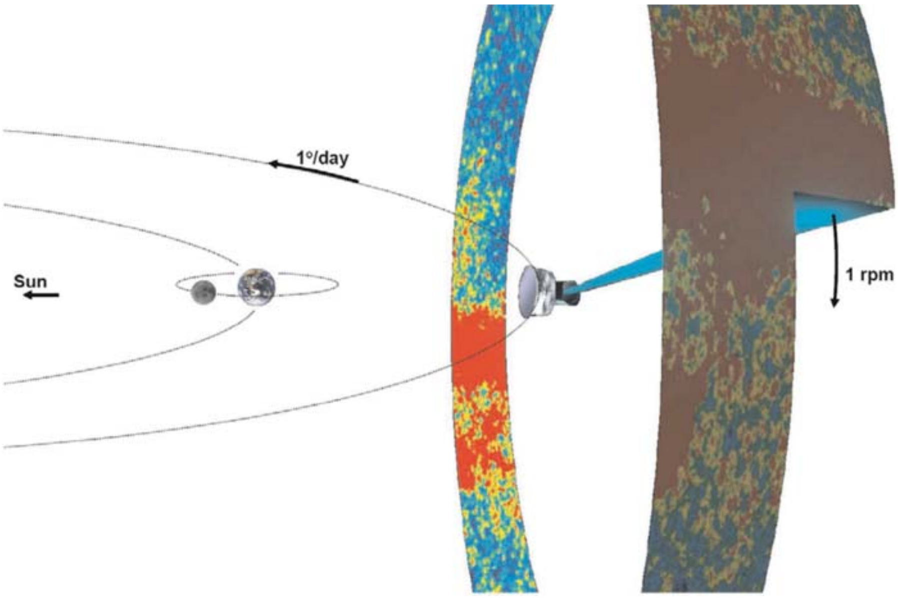
`Planck` was the fourth in a row of space-based CMB experiments, preceded by `RELIKT-1`, `COBE`, and `WMAP`. The idea of `Planck` was born in 1992, after the release of `COBE's` first scientific data. It took several years of development until it was finally launched in 2009 together with the Herschel space observatory. The `Planck` spacecraft is pictured in Fig. 6.1a, showing the telescope with a diameter of ~ 1.5 m. The primary reflector itself is surrounded by baffles that shield the optical path from any stray light coming from the Earth, the Moon, or the Sun. The collected light is guided through feedhorns and onto the focal plane unit. In order to examine the entire frequency range from synchrotron to dust emission, the FPU is shared by two instruments. The lower part of the spectrum is covered by the *Low Frequency Instrument* (LFI), while the upper part is mapped by the *High Frequency Instrument* (HFI). Mapping the sky with two independent systems also provides good control over systematic effects, since these will appear differently in the two instruments. The spacing of the frequency channels employed by `Planck` is plotted in Fig. 2.2 as shaded gray bands.

The LFI [56] covers the frequency range from 30 to 70 GHz in three bands[1]. The core of the instrument is an array of 22 radiometers fed by 11 feedhorns. Each horn supplies an orthomode transducer that divides the signal into its

---

[1] 30, 44 and 70 GHz

(a) The `Planck` instrument. Figure courtesy of [55].



(b) Illustration of scanning strategy. Figure courtesy of [55].

Figure 6.1: Top: The `Planck` instrument shortly before launch. Bottom: Placed at the second Lagrange point, `Planck` scans the sky while always avoiding the Sun.

two linear polarization states, which are then led to two independent radiometers. Hence, each horn feeds two detectors. The Stokes parameters can then be recovered by differencing the split signal. The detector technology used by the instrument consists of high-electron-mobility transistors (HEMTs) that are cooled to 20 K. In order to overcome the problem of variable gain, the detected signal is interleaved with a 4K-blackbody reference load. Since the sky signal and the signal from the reference load are subject to the same gain fluctuations, the true sky signal can be recovered [57].

The HFI [58] measured six frequencies[2] ranging from 100 to 857 GHz. As opposed to LFI, HFI employs bolometer technology that operates at 0.1 K. The four lowest channels are sensitive to linear polarization, while the highest two only measure intensity. Instead of using orthomode transducers, bolometers employ orthogonal sets of wires that only measure radiation whose electric field is parallel to these wires. The Stokes parameters can then be recovered by combining the data from detectors that are rotated by 45° with respect to each other.

`Planck` observed from the second Lagrangian point at a distance of ~ 1.5 million kilometers from Earth, which offers an unobstructed view of the sky. In order to avoid stray light, the optical path was shielded by the instrument itself and rotated on an axis pointing towards the Sun. By keeping the axis anchored at the Sun, while moving along the annual orbit, the instrument permanently avoids solar radiation and scans about 1 % of the sky per day [57]. The scanning strategy is illustrated in Fig. 6.1b. The resulting sky maps of Stokes parameter I were already presented earlier in Fig. 2.1 in Sec. 2. In the next section we will turn our attention to the corresponding polarization maps and use these to demonstrate the procedure of component separation.

6.2 Applied component separation

Figure 6.2 shows maps of the seven polarization-sensitive frequency channels of `Planck`, as seen through the Stokes parameters Q and U. The pre-processing and mapmaking of this data set has been performed by `Planck's` `NPIPE` pipeline. `NPIPE` is a processing pipeline for the raw timestreams of `Planck`, which is executed independently from the main `Planck` processing pipeline. The low and high frequencies are clearly more foreground dominated than the intermediate ones. Note that the foreground emission at 353 GHz is so strong that the colorbar had to be extended to avoid complete saturation of the figure.

The dominating polarized foregrounds are synchrotron and thermal dust emission, as illustrated earlier in Fig. 2.2b. In the following we will use the Bayesian component separation software `Commander`[3] [45], which is an implementation of the approach described in Sec. 3.1. For successful component separation, several

---

[2] 100, 143, 217, 353, 545 and 857 GHz

[3] `https://github.com/hke/Commander`

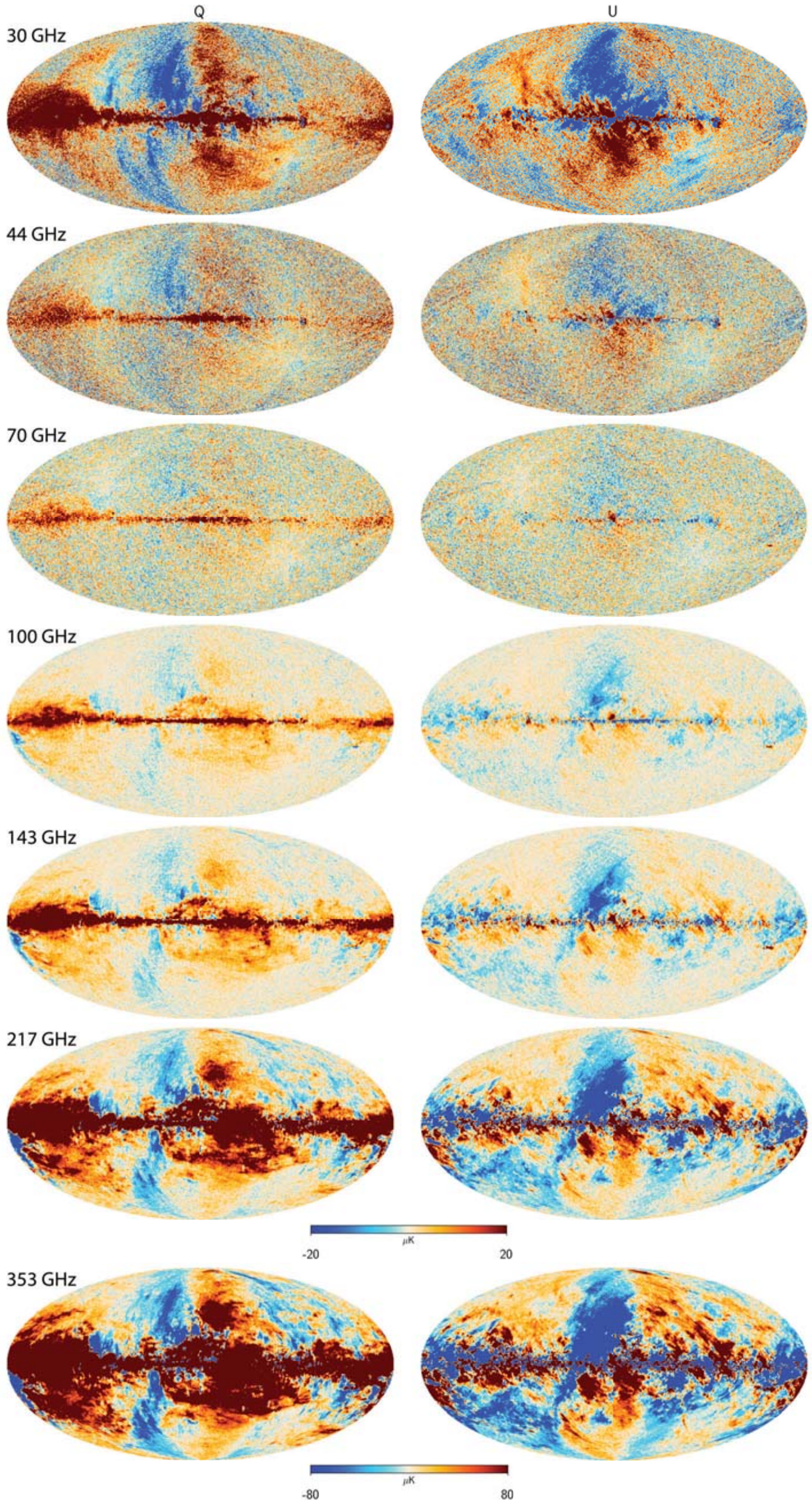


Figure 6.2: `Planck` `NPIPE` data maps for Stokes parameter Q (left) and U (right).

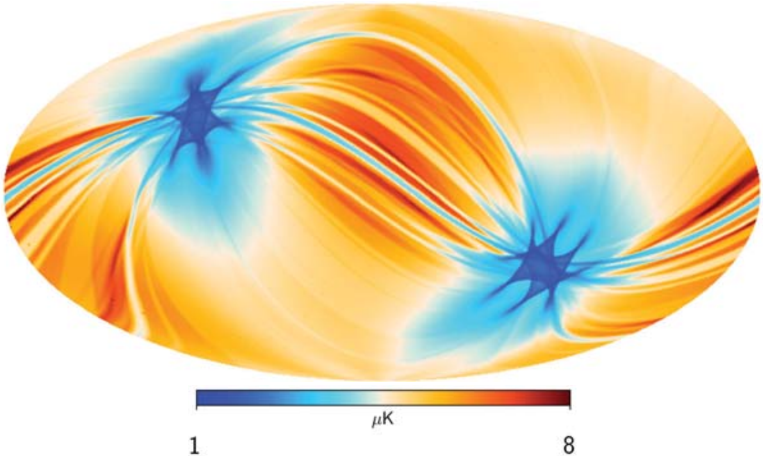


Figure 6.3: Noise RMS map for 70 GHz channel.

ingredients are needed. First, we need data maps for each frequency, as given in Fig. 6.2. Second, because of the scanning strategy of the instrument, some regions of the sky were observed more frequently than others, leading to a varying statistical uncertainty across the sky. This uncertainty is quantified in terms of RMS-maps, which contain the root mean square of the statistical noise per pixel for each frequency channel. The RMS map is therefore directly related to the integrated observation time for each pixel. One of these maps is pictured in Fig. 6.3. Since the spacecraft rotates while keeping the axis of rotation pointed towards the Sun (see Fig. 4.6), the northern and southern ecliptic poles were observed most frequently, and therefore feature a very low noise RMS. Third, the instrumental response is not limited to an infinitesimally small frequency range, but rather to a finite frequency interval called the *bandpass*. These are measured prior to launch and are illustrated in Fig. 6.4. Lastly, one has to consider the *instrumental beam*, which is the response of the instrument to a point source and is therefore related to the smallest resolvable scale. The data maps shown in Fig. 6.2 have had their instrumental beam deconvolved and were smoothed to a common resolution with a Gaussian beam with a FWHM of 40 arcmin.

Implemented internally in `Commander` are foreground models that parametrically describe the foreground SED. In polarization analysis, only synchrotron and thermal dust emission are relevant; their parametric models have been discussed in Eq. 2.2 and 2.1 respectively. Consequently, the free parameters per pixel are the foreground amplitudes (A_s, A_d), the spectral indices (β_s, β_d), the dust temperature (T_d), and the CMB amplitude (A_CMB). In the analysis presented here, amplitudes and spectral indices were fitted to each individual pixel, as illustrated in Fig. 3.1, while a fixed template was used for the dust temperature. The reason for not fitting the dust temperature along with the other parameters is that, for an adequate estimation, the higher `Planck` channels at 545 and 857 GHz are needed. The dust temperature template is therefore the

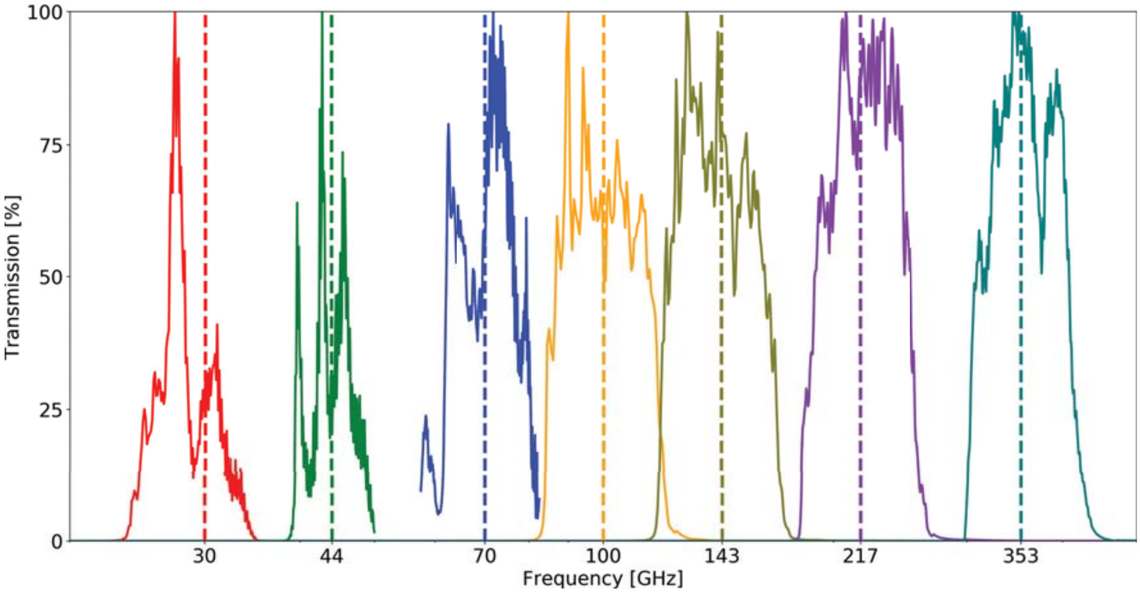


Figure 6.4: Bandpasses for `Planck` instrument.

result of a temperature-only analysis, which is then reused for polarization.

Apart from the parameters to be fitted per pixel, the software also allows for the treatment of global parameters, such as gains, bandpasses, monopoles, dipoles, and general templates. In this context, the gain is simply a multiplication factor that scales an input map up or down. The optimization of bandpasses can be achieved by rigidly shifting the individual transmission curves in Fig. 6.4 in frequency space. Finally, monopoles as well as templates for e.g. Zodiacal light can be neglected in polarization analyses. Even though the dipole directly only manifests itself in the temperature of the CMB, and not in its polarization, there are instrumental effects that make the dipole the largest contaminator of polarization maps. This problem will be more thoroughly discussed in Sec. 6.3.

In general, component separation is not performed in one attempt, but rather in several consecutive runs. It is often recommended not to free all parameters at once, but preferably start out by fitting only one parameter, namely the one with the biggest impact on the χ^2 (see Eq. 6.1). After convergence, one can initialize at the previous solution and continue with freeing another, until a joint fit is achieved. The next section presents the result of several dozens of such component separation iterations.

6.2.1 Results

Figures 6.5 through 6.7 show the results of component separation. The goodness of fit is quantified by the chi-squared

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

where $\mathbf{d}_\nu(p)$ are the data, $\mathbf{s}_\nu(p)$ the model, and $\sigma_\nu(p)$ the rms per pixel. The χ^2 is plotted in the second panel of Fig. 6.5. The better the model represents the data, the lower the χ^2 will be. Hence, the figure shows that the fit is very successful, and only regions along the galactic equator perform poorly. χ^2-maps as shown here are commonly used to create *masks*, which are templates that mark pixels excluded from further analysis. Any contribution to $\mathbf{d}_\nu(p)$ that is not explained by the model appears in the *residual*

$$\mathbf{r}_\nu(p) = \mathbf{d}_\nu(p) - \mathbf{s}_\nu(p), \tag{6.2}$$

which is shown for each channel in Fig. 6.7. Thus, the χ^2 tells us something about the performance of the overall fit, while the individual residuals give us additional information on where the contribution to χ^2 comes from. If the model is perfect, the residuals only contain noise. In our case, we can still see some structure along the galactic equator, but otherwise, the maps resemble noise, as they should. Again, the noise level is not spatially constant due to the scanning strategy.

The main result is given in Fig. 6.5, which shows the Q- and U-maps for the CMB at the top, and maps of thermal dust and synchrotron in the third and fourth panel respectively. The synchrotron amplitude map is evaluated at 30 GHz, therefore the diffuse morphology closely resembles that seen in the 30 GHz data map in Fig. 6.2, where synchrotron is the strongest. Conversely, the dust amplitude map is shown at a frequency of 353 GHz. Accordingly, the structure looks like that visible in the high frequency channels, which are dominated by thermal dust. For maps of the spectral indices and the dust temperature template, see Fig. 6.6.

6.3 The CMB Dipole

The CMB dipole was briefly introduced in Sec. 1.3, but has not been emphasized since. Even though it is neither an astrophysical foreground, nor a primary effect in polarization, it still deserves special treatment, which it will be given in this Section.

The movement of the `Planck` instrument with respect to the CMB rest frame leaves an imprint in the measured timestreams. This imprint has two contributions, caused by the *orbital* and the *solar* dipole. The former is the result of the instrument's motion around the Sun and has an annual variation, while the latter, much larger contribution is due to the solar system's orbit around the Milky Way's center. In subsequent data processing, this pattern is accounted for in one of two ways. The first approach is to remove both dipole contributions at timestream level, which was the method chosen for the main `Planck` processing pipeline. `NPIPE` on the other hand only removes the orbital dipole from the timestreams, while leaving the solar dipole untouched. Since the CMB dipole has exactly the same frequency behavior as the CMB itself, namely a

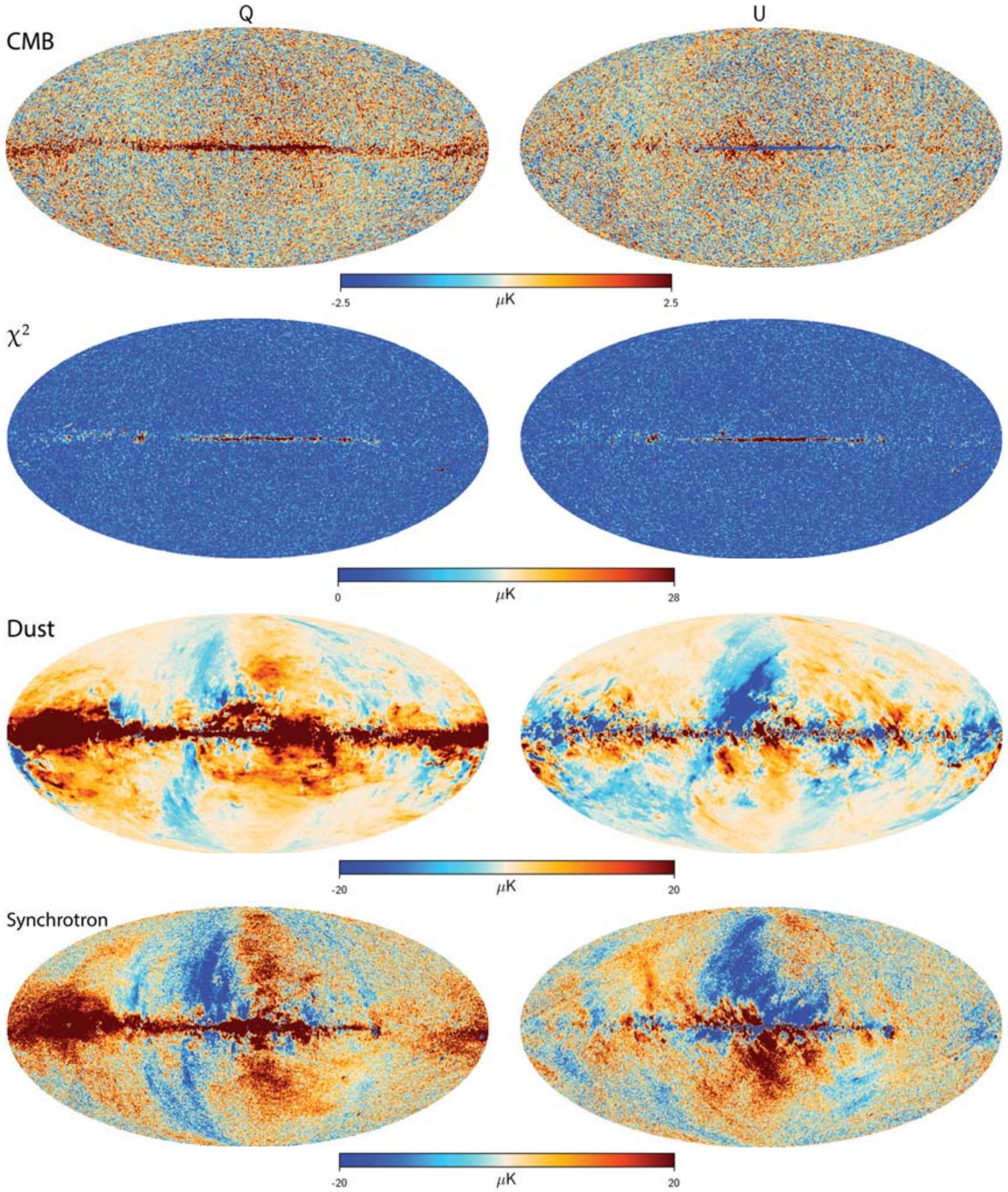


Figure 6.5: Results of component separation. Synchrotron and dust amplitude are evaluated at 30 and 353 GHz respectively.

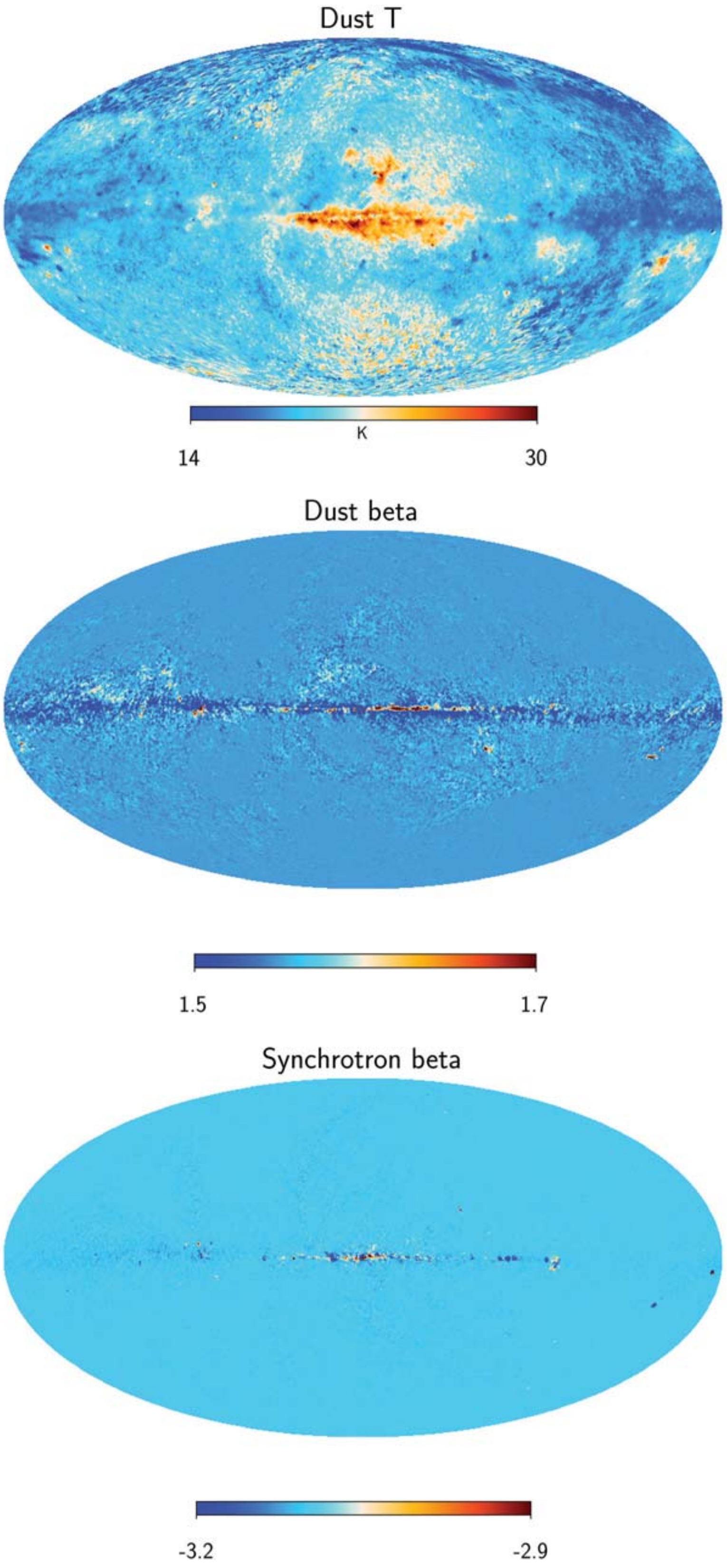


Figure 6.6: Dust temperature and spectral index maps.

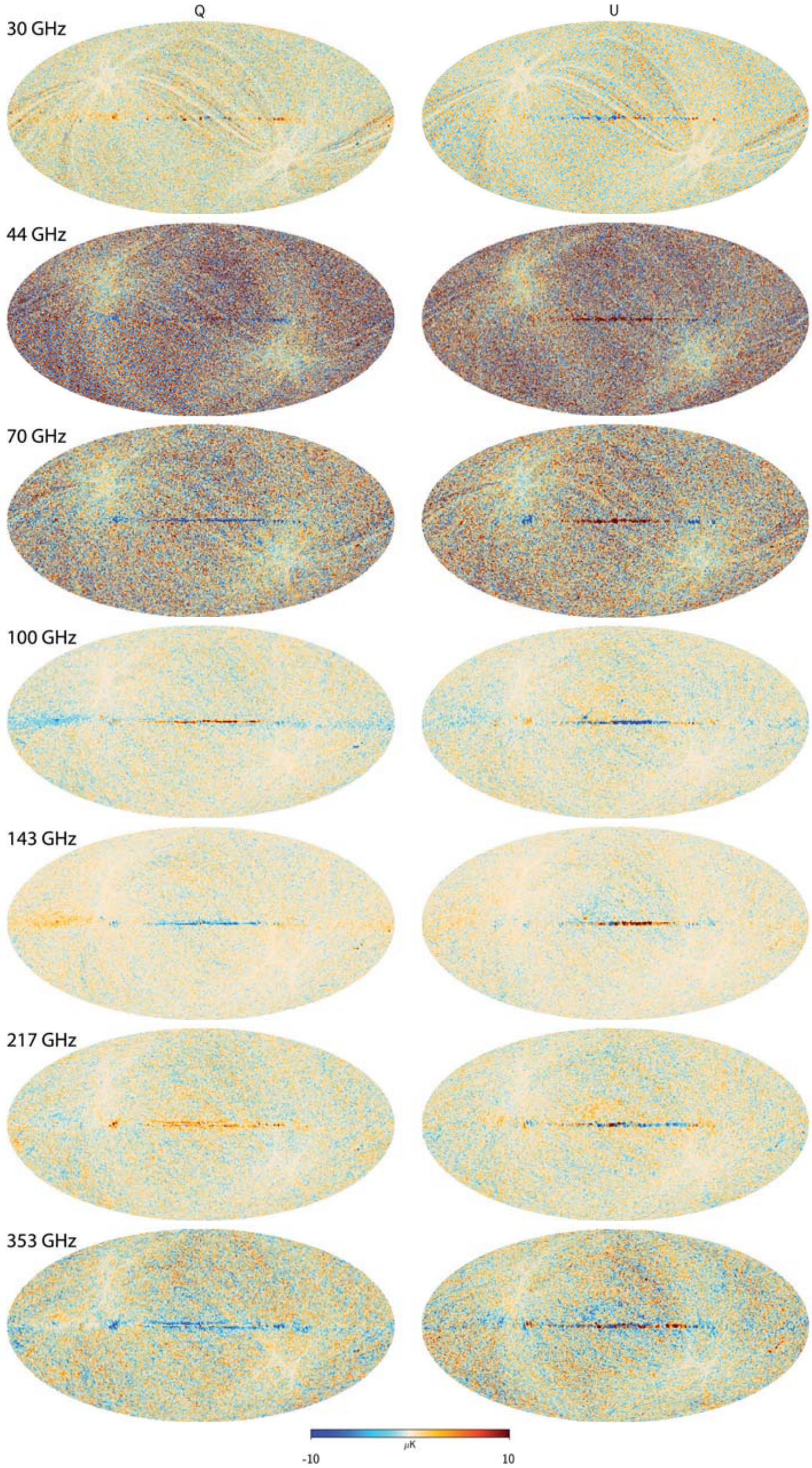


Figure 6.7: Residuals for Stokes parameter Q (left) and U (right).

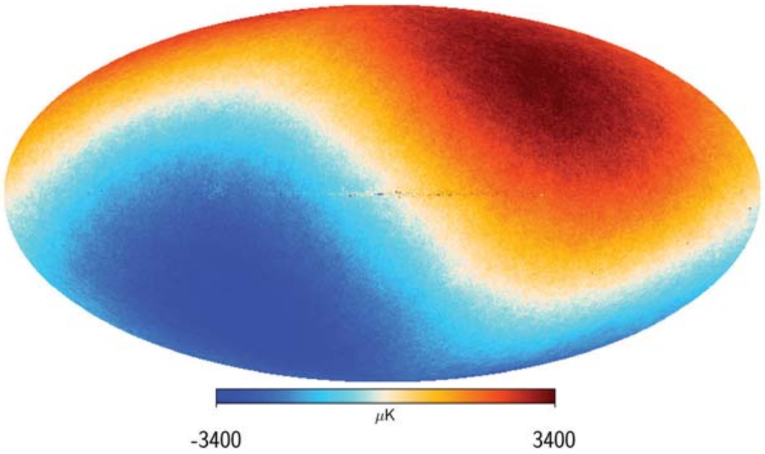


Figure 6.8: Component-separated `NPIPE` CMB map is dominated by solar dipole.

perfect blackbody, it will end up together with the CMB after component separation. This has the advantage that the dipole can be treated at map level after component separation. Figure 6.8 shows the component-separated CMB map in temperature after `NPIPE` processing, which is completely dominated by the solar dipole. While the dipole only manifests itself in Stokes parameter I, there are two reasons why knowing the dipole amplitude is crucial for polarization analysis as well.

First, the dipole is the primary calibration source for many CMB experiments. Since both the distance from the instrument to the Sun and the revolution period around the Sun are known with very high precision, one can use the orbital dipole for *absolute calibration*. The solar dipole on the other hand enables *relative calibration*, as it is brighter than most other signals and available across the full sky. Any uncertainty in the dipole parameters therefore biases the normalization of the computed power spectra and, correspondingly, the derived cosmological parameters.

The second reason is an instrumental effect called *temperature-to-polarization leakage*, caused by gain mismatches between pairs of detectors. This has the consequence that a small part of the temperature signal is interpreted as polarization. Since the dipole amplitude is extremely large, even very little leakage leads to strong contamination of polarization maps. Any cosmological parameter that is derived from polarization power spectra, such as the optical depth of reionization τ, would consequently be biased.

The estimation of the dipole parameters from maps is complicated by the process of *masking*. Since the component-separated CMB map is subject to residual contamination in the galactic plane, which is even stronger than the dipole amplitude itself (see. Fig 6.8), one has to mask out these regions in order to avoid biasing the estimate. This seemingly harmless process however afflicts the dipole estimate with an error, whose origin can be better understood in the spherical harmonic domain. The subtle problem is that masking a CMB map goes hand in hand with masking the base functions of the corresponding spherical harmonic representation, which thereby lose their orthogonality. Since this changes the spherical harmonic coefficients of the CMB map, the resulting effect is leakage from higher-order multipoles into the dipole component[4]. We are therefore confronted with the dilemma of choosing a mask that is big enough to suppress unwanted emission, while being small enough to keep the mask-induced error at a minimum.

I played a central role in `NPIPE's` dipole estimation. One of my main projects was the implementation of an estimation technique that allows for a reduction of the mask-induced error. The Paper that emerged from this work, presented in Part III, describes an estimation method based on Wiener filtering. By exploiting basic properties of the CMB, it is possible to partially reconstruct the sky inside the mask, thereby decreasing the related error. Monte Carlo simulations have demonstrated that this method drastically reduces the dipole uncertainty. The resulting dipole parameter estimates are much more accurate than those achieved by standard estimation approaches. For a full review of the theory and the implementation of this method, I refer the reader to Paper I.

The ultimate goal is to apply the estimation technique described in Paper I on real data. I was responsible for this task in the `NPIPE` analysis, and the dipole parameter estimation performed on the component-separated CMB map is presented in Sec. 8 in Paper III. The result might be considered the most accurate dipole estimation to date.

6.4 Validation

Before estimating cosmological parameters from component-separated CMB maps, it is common practice to first perform a *validation* to guarantee the control over systematic effects. Even though χ^2 and residual maps offer a way to visually identify systematic effects in the data maps, since it's there where all unaccounted emission manifests itself, there is a more common way of doing so. The common way of identifying nonidealities is by means of *null tests* that involve data splits and end-to-end simulations. The overall aim is to check the consistency of an ensemble of simulated power spectra with that of the component-separated CMB map, a process I have been involved with in both `Planck` and `NPIPE` analysis (see Paper II and III). This procedure will be outlined in the following.

In order to characterize the instrumental noise and systematic effects of the `Planck` instrument, there are three measures of interest. The first is odd-even half-difference (OEHD) maps, which entails splitting the time-ordered data

---

[4] Which is multipole $\ell = 1$.

into two sets. The first set consists of even-numbered rings of `Planck's` scanning strategy, while the second is composed of odd-numbered rings. These two data sets are then propagated independently through all steps of the processing pipeline, including calibration, mapmaking, and component separation. From the clean and component-separated CMB maps resulting from the odd and even data splits, $\mathbf{m}_{\mathrm{odd}}$ and $\mathbf{m}_{\mathrm{even}}$, one computes the half-difference map given by

$$\mathbf{m}_{\mathrm{OEHD}} = \frac{\mathbf{m}_{\mathrm{odd}} - \mathbf{m}_{\mathrm{even}}}{2}. \tag{6.3}$$

The reason behind computing the difference of the split maps is that the sky signal present in both maps cancels. Furthermore, because of the intertwined data split based on picking out even and odd-numbered data rings, most of the systematic effects are likely to cancel as well, as these will be present in both splits. Consequently, OEHD maps are a good measure of the *instrumental noise* contained in the CMB maps.

The second measure of interest is half-mission half-difference (HMHD) maps. The procedure is completely analogous to that of OEHD maps, only differing in the choice of data split. As the name suggests, the entire data volume is split into two chronological halves, each half consisting of years of uninterrupted time-ordered data. For that reason, systematic effects that vary on long time scales are not expected to cancel. Hence, HMHD maps are a measure of the joint effect of *instrumental noise and systematic effects*.

Lastly, in order to verify that all systematic effects are under control, one needs a point of comparison. This is provided by simulations, that are used as a benchmark against which we compare the data. These are full end-to-end simulations, taking into account all known instrumental effects and sky components. The starting point is simulations for the foreground emission using the *Planck sky model* [59] to which simulated CMB skies are added. Using the real pointing information of the instrument, the simulated sky is converted into timestreams for each bolometer and radiometer. Following that, instrumental effects and noise are modeled and superimposed at timestream level. These are then propagated through the full pipeline, following the identical data splits and processing steps as the real time-ordered data, such as mapmaking and component separation. Since these simulations are extremely computationally heavy, the number of realizations is limited. In the case of `Planck` and `NPIPE`, 300 end-to-end simulations were performed.

In the null test, data and model are compared at the power spectrum level. The procedure above delivers clean component-separated CMB maps for Stokes parameters I, Q, and U, for the full mission as well as for OEHD and HMHD splits. In addition, we have the corresponding simulations. The task now is to compute the CMB power spectrum for each map, for instance by using `PolSpice` [60], where we also apply a mask to suppress residual galactic emission. The comparison is then made by computing the percentage deviation

$$\eta_\ell \equiv \frac{D_\ell^{\mathrm{data}} - \langle D_\ell^{\mathrm{sim}} \rangle}{\langle D_\ell^{\mathrm{sim}} \rangle} \cdot 100, \tag{6.4}$$

where $D_\ell = C_\ell \cdot \ell(\ell+1)$. Hence, this quantity tells us by what percentage the observed power spectrum deviates from the mean of the simulated ones, and can be computed for temperature and polarization power spectra (D^{TT}, D^{EE}, D^{BB}) as well as for full-mission, OEHD, and HMHD data splits. Any strong deviation would then indicate systematic effects or foregrounds that have not been accounted for, and calls for further processing of the data. Fortunately, an adequate characterization of instrumental effects and noise has been achieved for `Planck` and `NPIPE`. Apart from the low-ℓ regime, where cosmic variance is the dominant player, the percentage deviation η_ℓ remains at the ± 5 %-level. For a full review of the null tests performed for the four `Planck` pipelines and `NPIPE`, I refer the interested reader to Fig. 11 and 69 in Paper II and III respectively.

Once the component-separated CMB maps have been validated, one can proceed to cosmological parameter estimation.

6.5 The final step towards cosmological parameters

The heading of this section might be somewhat of an understatement, as the process of cosmological parameter estimation envelops several substeps and can as such be considered an entire field of research. Nonetheless, this final part of the thesis will briefly outline the basic strategy of how cosmological parameters are estimated from maps, while skipping a real-world demonstration and any technicalities.

As already mentioned in the introductory section, it is the CMB power spectrum through which the data are compared to the theory. Trying to explain the essence of cosmological parameter estimation compressed into one single sentence might sound something like this: The task at hand is to find the cosmological parameters which produce CMB power spectra that best fit the data. There are three main problems involved. First, we need to reduce the information contained in the data maps to power spectra. Second, we need to compute the theoretical power spectra, given a set of cosmological parameters. Third, we need a measure to compare the two and find the best fit.

A standard software for the computation of both temperature and polarization power spectra, including their cross-terms, is `PolSpice` [60]. While `PolSpice` derives power spectra from data maps, `CAMB`[5] [61] is a code that computes the anisotropies of the CMB and its polarization based on the ΛCDM model, and therefore supplies us with theoretical power spectra that result from a set

---

[5] **C**ode for **A**nisotropies in the **M**icrowave **B**ackground; `https://camb.info/`

of parameters. This set of parameters consists of A_s and n_s, the amplitude and spectral index of primordial density perturbations; τ, the optical depth of reionization; H_0, the present-day Hubble parameter; and finally $\omega_b h^2$ and $\omega_c h^2$, the present-day baryon and dark matter densities, where $h = H_0/100$. To find the set of parameters that best fits the data is then a problem of Bayesian inference. Let's denote $\vec{D}$ as the data, i.e. a set of measured power spectra, which we want to describe using a cosmological parameter vector $\vec{\theta}$. The goal is to estimate the posterior distribution of the model parameters given the data, which according to Bayes' theorem is given by

$$P(\vec{\theta}|\vec{D}) \propto \mathcal{L}(\vec{\theta}) P(\vec{\theta}), \tag{6.5}$$

where $\mathcal{L}(\vec{\theta}) \equiv P(\vec{D}|\vec{\theta})$ is the likelihood and $P(\vec{\theta})$ a set of priors. The prior represents the best knowledge we have on the parameters, and is in practice often chosen to be uniform. The posterior probability is then maximized by maximizing the likelihood function, which can be shown to be a simple χ^2-minimization problem. The difficulty of this task is not the size of the data set, but rather the dimensionality of the parameter space and therefore requires an efficient engine to explore it. `CosmoMC`[6] [62] is a widely used Markov Chain Monte Carlo algorithm that is designed to do just that. By stepping through the parameter space, while relying on `CAMB` for producing the theoretical spectra, it is possible to map out the posterior distribution given in Eq. 6.5 and find the set of parameters that gives the maximum.

6.6 Summary and outlook

This Chapter described the final steps of the generic analysis pipeline, starting out at sky maps and ending up at cosmological parameters. The key steps to reach this goal are *component separation* (Sec. 6.2), *map validation* (Sec. 6.4), and finally *parameter estimation* (Sec. 6.5). Component separation was illustrated by means of polarized sky maps in Stokes parameters Q and U; the process in temperature proceeds analogously, but involves a few additional subtleties. Since there are only two polarized foregrounds, namely thermal dust and synchrotron emission, the `Planck` frequencies alone are sufficient to achieve a satisfying separation. In the temperature domain on the other hand, the large variety of unpolarized foregrounds makes it necessary to find a way to break degeneracies that arise between free-free and synchrotron emission, which both are described by power laws. In practice, one therefore adds additional sky maps of alternative CMB missions to not only increase the number of channels but also to extend the covered frequency range. In `Planck` analysis, this is achieved by including maps of `WMAP` and the low frequency map of the Haslam survey, which is a great tracer of synchrotron emission. Only then can one fit for the full number of free parameters. Since `WMAP` also provides polarized sky maps,

---

[6] **Cosmo**logical **M**onte**C**arlo; `https://cosmologist.info/cosmomc/`

using these in polarization analysis leads to a more stable solution.

The validation of component-separated CMB maps (Sec. 6.4) is done by means of null tests, which makes use of various data splits and detailed modeling of instrumental behavior in order to characterize instrumental effects and noise. Both data and simulation are propagated through all stages of the analysis pipeline and ultimately compared at power spectrum level. One of my tasks in `Planck` and `NPIPE` analysis was to demonstrate the power spectrum consistency using simulations. For the resulting plots showing the percentage deviation between data and simulation, see Fig. 11 and 69 in Paper II and III respectively.

Finally, a major project was dedicated to the dipole, whose importance was explained in Sec. 6.3. At the current stage of `Planck` analysis, the uncertainty of the dipole is no longer limited by instrumental effects and noise, but rather by an error introduced when masking unwanted residual emission. This problem is thoroughly discussed in Paper I and describes a method to mitigate this effect by using a Wiener filter algorithm. For an application of this estimation method on real `NPIPE` data, see. Sec. 8 in Paper III. The resulting dipole parameter estimates may be considered as the most accurate ones to date.

Chapter 7

Summary and Outlook

The unified goal of today's CMB experiments is to detect the precious B-mode signal, which according to theory, is a relic from the epoch of inflation. To achieve this task, two major challenges have to be overcome. The first is to build detector technology sensitive enough to pick it up in the first place; the second is the characterization of polarized foregrounds, which are orders of magnitude larger than the signal we are looking for. This thesis described the generic analysis workflow of such experiments, illustrated by the example of three very different missions, `GreenPol`, `SPIDER`, and `Planck`. While the `Planck` instrument has reached the end of its lifetime, its data live on and are presently being reprocessed. Furthermore, both `GreenPol` and `SPIDER` are awaiting future deployments.

The CMB experiments introduced in this thesis are by far not the only ones that search the sky for B-modes. Presently, huge efforts are put into ground-based telescopes that concentrate on smaller patches on the sky. These can be distinguished by the size of their aperture. Small-aperture telescopes like `BICEP2` and `Keck` at the South Pole observe with a larger beam and target the large-scale B-mode polarization signal. Larger telescopes on the other hand have the ability to resolve smaller scales and therefore have a wider ranger of scientific goals. The main players among these are the *South Pole Telescope* (SPT) in Antarctica, and the *Atacama Cosmology Telescope* (ACT) and *Simons array* both located in Chile. Ground-based CMB experiments are commonly classified in stages, depending on their sensitivity, which is related to the number of detectors used. Figure 7.1 shows the breathtaking evolution of raw experimental sensitivity over last two decades. Both SPT and ACT are currently observing with their third generation receivers, namely SPT-3G and Advanced ACTPol. In order to further increase the sensitivity, major collaborations are planning to join forces by building a globe-spanning network of already existing and future telescopes. This will be the ultimate ground-based experiment, called CMB-S4.

The future doesn't only look promising for ground-based missions; the list of space-based missions is also awaiting a new member. A new project has only recently received funding to develop a CMB observatory called `LiteBIRD`[1]. The goal for this satellite is to perform a full-sky survey of the polarized CMB on the hunt for B-modes and thereby constrain the scalar-to-tensor ration down to $r = 0.001$ [63]. Just like `Planck`, it will observe from the second Lagrangian point at several frequency bands, which will allow for a good foreground char-

---

[1] Lite (Light) satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation

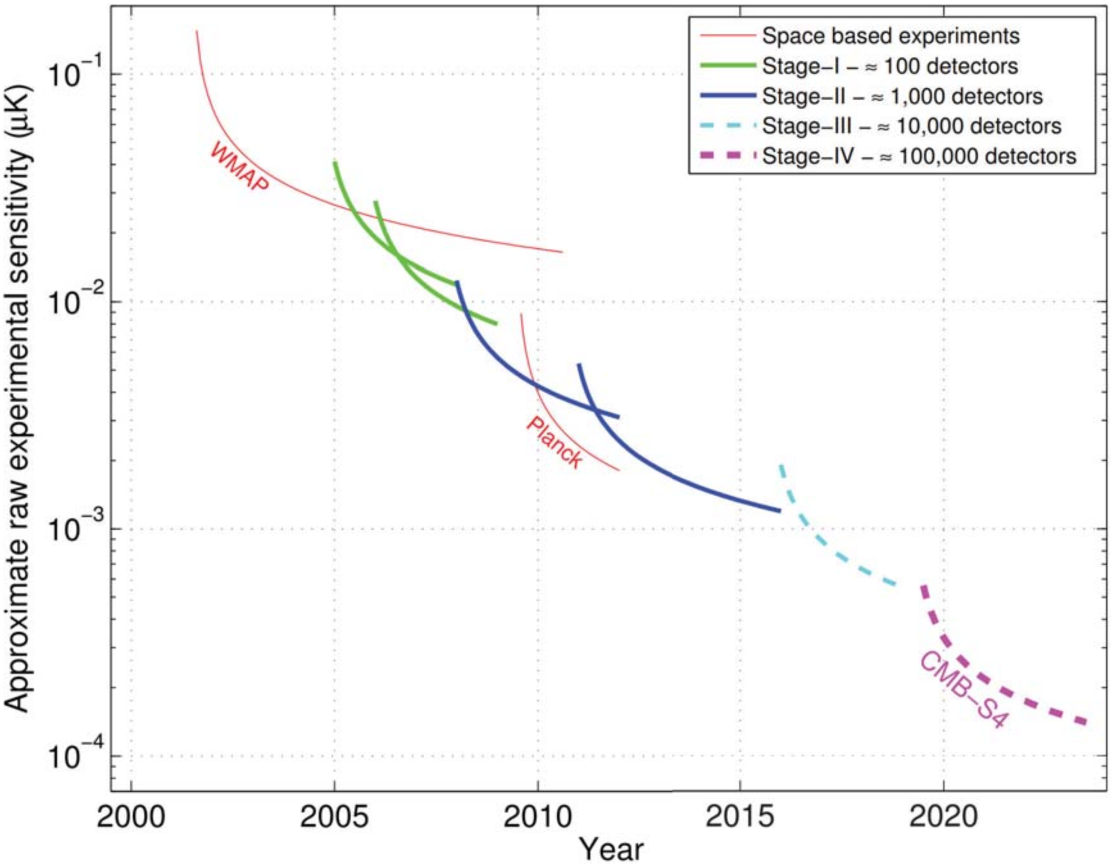


Figure 7.1: Evolution of raw sensitivity of CMB experiments. Figure courtesy of [64].

acterization.

The handling of the upcoming flood of data will require more sophisticated analysis pipelines and computer algorithms than currently available. The future therefore holds many exciting challenges and I'm happy to be part of the community that dares to take them on.

Bibliography

[1] E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929.

[2] H. Bondi and T. Gold. The Steady-State Theory of the Expanding Universe. *MNRAS*, 108:252, 1948.

[3] R. A. Alpher, H. Bethe, and G. Gamow. The origin of chemical elements. *Phys. Rev.*, 73:803–804, Apr 1948.

[4] Ralph A. Alpher and Robert C. Herman. On the relative abundance of the elements. *Phys. Rev.*, 74:1737–1742, Dec 1948.

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

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

[7] Wikipedia contributors. Lists of cosmic microwave background experiments — Wikipedia, the free encyclopedia, 2019. [Online; accessed 11-May-2019].

[8] N. W. Boggess, J. C. Mather, R. Weiss, C. L. Bennett, E. S. Cheng, E. Dwek, S. Gulkis, M. G. Hauser, M. A. Janssen, T. Kelsall, S. S. Meyer, S. H. Moseley, T. L. Murdock, R. A. Shafer, R. F. Silverberg, G. F. Smoot, D. T. Wilkinson, and E. L. Wright. The COBE mission - Its design and performance two years after launch. *Ap.J.*, 397:420–429, October 1992.

[9] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright. The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set. *Ap.J.*, 473:576, December 1996.

[10] C. L. Bennett, A. J. Banday, K. M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright. Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results. *Ap.J.*, 464:L1, Jun 1996.

[11] C. L. Bennett, M. Bay, M. Halpern, G. Hinshaw, C. Jackson, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, D. T. Wilkinson, E. Wollack, and E. L. Wright. The Microwave Anisotropy Probe Mission. *Ap.J.*, 583:1–23, January 2003.

[12] Planck Collaboration. Planck 2018 results. I. Overview and the cosmological legacy of Planck. *arXiv e-prints*, page arXiv:1807.06205, Jul 2018.

[13] D. J. Fixsen. The Temperature of the Cosmic Microwave Background. *Ap.J.*, 707(2):916–920, Dec 2009.

[14] C. H. Lineweaver. The CMB dipole: the most recent measurement and some history. In *Microwave Background Anisotropies*, volume 16, pages 69–75, Jan 1997.

[15] Planck Collaboration. Planck 2018 results. IV. Diffuse component separation. *arXiv e-prints*, page arXiv:1807.06208, Jul 2018.

[16] Wayne Hu and Martin White. A CMB polarization primer. *New Astronomy*, 2(4):323–344, Oct 1997.

[17] J. W. Henning, J. T. Sayre, C. L. Reichardt, P. A. R. Ade, A. J. Anderson, J. E. Austermann, J. A. Beall, A. N. Bender, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. C. Chiang, H. M. Cho, R. Citron, C. Corbett Moran, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, W. Everett, J. Gallicchio, E. M. George, A. Gilbert, N. W. Halverson, N. Harrington, G. C. Hilton, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, N. Huang, J. Hubmayr, K. D. Irwin, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, D. Li, A. Lowitz, A. Manzotti, J. J. McMahon, S. S. Meyer, L. Mocanu, J. Montgomery, A. Nadolski, T. Natoli, J. P. Nibarger, V. Novosad, S. Padin, C. Pryke, J. E. Ruhl, B. R. Saliwanchik, K. K. Schaffer, C. Sievers, G. Smecher, A. A. Stark, K. T. Story, C. Tucker, K. Vanderlinde, T. Veach, J. D. Vieira, G. Wang, N. Whitehorn, W. L. K. Wu, and V. Yefremenko. Measurements of the Temperature and E-mode Polarization of the CMB from 500 Square Degrees of SPTpol Data. *Ap.J.*, 852(2):97, Jan 2018.

[18] Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, Jan 1981.

[19] The BICEP/Keck Collaboration, :, P. A. R. Ade, Z. Ahmed, R. W. Aikin, K. D. Alexander, D. Barkats, S. J. Benton, C. A. Bischoff, J. J. Bock, H. Boenish, R. Bowens-Rubin, J. A. Brevik, I. Buder, E. Bullock, V. Buza, J. Connors, J. Cornelison, B. P. Crill, M. Crumrine, M. Dierickx, L. Duband, C. Dvorkin, J. P. Filippini, S. Fliescherj J. Grayson, G. Hall, M. Halpern, S. Harrison, S. R. Hildebrandt, G. C. Hilton, H. Hui, K. D. Irwin, J. Kang, K. S. Karkare, E. Karpel, J. P. Kaufman, B. G. Keating, S. Kefeli, S. A. Kernasovskiy, J. M. Kovac, C. L. Kuo, K. Lau, N. A. Larsen, E. M. Leitch, M. Lueker, K. G. Megerian, L. Moncelsi, T. Namikawa, B. Netterfield, H. T. Nguyen, R. O'Brient, IV Ogburn, R. W., S. Palladino, C. Pryke, B. Racine, S. Richter, R. Schwarz, A. Schillaci, C. D. Sheehy, A. Soliman, T. St. Germaine, Z. K. Staniszewski, B. Steinbach, R. V. Sudiwala, G. P. Teply, K. L. Thompson, J. E. Tolan, C. Tucker, A. D. Turner, C. Umilta, A. G. Vieregg, A. Wandui, A. C. Weber, D. V. Wiebe, J. Willmert, C. L. Wong, W. L. K. Wu, H. Yang, K. W. Yoon, and C. Zhang. Measurements of Degree-Scale B-mode Polarization with the BICEP/Keck

Experiments at South Pole. *arXiv e-prints*, page arXiv:1807.02199, Jul 2018.

[20] Matias Zaldarriaga and Uroš Seljak. All-sky analysis of polarization in the microwave background. *Phys. Rev. D*, 55(4):1830–1840, Feb 1997.

[21] BICEP2 Collaboration, P. A. R. Ade, R. W. Aikin, D. Barkats, S. J. Benton, C. A. Bischoff, J. J. Bock, J. A. Brevik, I. Buder, E. Bullock, C. D. Dowell, L. Duband, J. P. Filippini, S. Fliescher, S. R. Golwala, M. Halpern, M. Hasselfield, S. R. Hildebrandt, G. C. Hilton, V. V. Hristov, K. D. Irwin, K. S. Karkare, J. P. Kaufman, B. G. Keating, S. A. Kernasovskiy, J. M. Kovac, C. L. Kuo, E. M. Leitch, M. Lueker, P. Mason, C. B. Netterfield, H. T. Nguyen, R. O'Brient, R. W. Ogburn, A. Orlando, C. Pryke, C. D. Reintsema, S. Richter, R. Schwarz, C. D. Sheehy, Z. K. Staniszewski, R. V. Sudiwala, G. P. Teply, J. E. Tolan, A. D. Turner, A. G. Vieregg, C. L. Wong, and K. W. Yoon. Detection of B-Mode Polarization at Degree Angular Scales by BICEP2. *Phys. Rev. Lett.*, 112(24):241101, Jun 2014.

[22] P. A. R. Ade, Z. Ahmed, R. W. Aikin, K. D. Alexander, D. Barkats, S. J. Benton, C. A. Bischoff, J. J. Bock, R. Bowens-Rubin, J. A. Brevik, I. Buder, E. Bullock, V. Buza, J. Connors, B. P. Crill, L. Duband, C. Dvorkin, J. P. Filippini, S. Fliescher, J. Grayson, M. Halpern, S. Harrison, G. C. Hilton, H. Hui, K. D. Irwin, K. S. Karkare, E. Karpel, J. P. Kaufman, B. G. Keating, S. Kefeli, S. A. Kernasovskiy, J. M. Kovac, C. L. Kuo, E. M. Leitch, M. Lueker, K. G. Megerian, C. B. Netterfield, H. T. Nguyen, R. O'Brient, R. W. Ogburn, A. Orlando, C. Pryke, S. Richter, R. Schwarz, C. D. Sheehy, Z. K. Staniszewski, B. Steinbach, R. V. Sudiwala, G. P. Teply, K. L. Thompson, J. E. Tolan, C. Tucker, A. D. Turner, A. G. Vieregg, A. C. Weber, D. V. Wiebe, J. Willmert, C. L. Wong, W. L. K. Wu, and K. W. Yoon. Improved constraints on cosmology and foregrounds from bicep2 and keck array cosmic microwave background data with inclusion of 95 ghz band. *Phys. Rev. Lett.*, 116:031302, Jan 2016.

[23] Planck Collaboration. Planck 2015 results. X. Diffuse component separation: Foreground maps. *Astron. Astrophys.*, 594:A10, Sep 2016.

[24] N. Odegard, A. Kogut, D. T. Chuss, and N. J. Miller. Assessment of Models of Galactic Thermal Dust Emission Using COBE/FIRAS and COBE/DIRBE Observations. *Ap.J.*, 828(1):16, Sep 2016.

[25] Planck Collaboration. Planck 2013 results. XI. All-sky model of thermal dust emission. *Astron. Astrophys.*, 571:A11, Nov 2014.

[26] W. A. Stein and B. T. Soifer. Dust in galaxies. **??***jnlARA&A*, 21:177–207, 1983.

[27] Planck Collaboration. Planck intermediate results. XIX. An overview of the polarized thermal emission from Galactic dust. *Astron. Astrophys.*, 576:A104, Apr 2015.

[28] V. L. Ginzburg and S. I. Syrovatskii. Cosmic Magnetobremsstrahlung (synchrotron Radiation). **??***jnlARA&A*, 3:297, 1965.

[29] M. S Longair. *High energy astrophysics*. Cambridge University Press, Cambridge, 3rd ed. edition, 2011.

[30] Bradley W Carroll. *An introduction to modern astrophysics*. Pearson, Harlow, 2nd ed. pearson new int. ed. edition, 2014.

[31] C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. A 408 MHz all-sky continuum survey. II - The atlas of contour maps. **??***jnlA&AS*, 47:1, January 1982.

[32] M. Remazeilles, C. Dickinson, A. J. Banday, M. A. Bigot-Sazy, and T. Ghosh. An improved source-subtracted and destriped 408-MHz all-sky map. *MNRAS*, 451(4):4311–4327, Aug 2015.

[33] E. M. Berkhuijsen, C. G. T. Haslam, and C. J. Salter. Are the galactic loops supernova remnants? *Astron. Astrophys.*, 14:252–262, September 1971.

[34] George F. Smoot. Galactic Free-free and H-alpha Emission. *arXiv e-prints*, pages astro–ph/9801121, Jan 1998.

[35] A. J. Banday, C. Dickinson, R. D. Davies, R. J. Davis, and K. M. Górski. Reappraising foreground contamination in the COBE-DMR data. *MNRAS*, 345(3):897–911, Nov 2003.

[36] E. M. Leitch, A. C. S. Readhead, T. J. Pearson, and S. T. Myers. An Anomalous Component of Galactic Emission. *Ap.J.*, 486(1):L23–L26, Sep 1997.

[37] B. T. Draine and A. Lazarian. Diffuse Galactic Emission from Spinning Dust Grains. *Ap.J. Lett.*, 494:L19–L22, February 1998.

[38] Clive Dickinson, Y. Ali-Haïmoud, A. Barr, E. S. Battistelli, A. Bell, L. Bernstein, S. Casassus, K. Cleary, B. T. Draine, R. Génova-Santos, S. E. Harper, B. Hensley, J. Hill-Valler, Thiem Hoang, F. P. Israel, L. Jew, A. Lazarian, J. P. Leahy, J. Leech, C. H. López-Caraballo, I. McDonald, E. J. Murphy, T. Onaka, R. Paladini, M. W. Peel, Y. Perrott, F. Poidevin, A. C. S. Readhead, J. A. Rubiño-Martín, A. C. Taylor, C. T. Tibbs, M. Todorović, and Matias Vidal. The State-of-Play of Anomalous Microwave Emission (AME) research. **??***jnlNew A Rev.*, 80:1–28, Feb 2018.

[39] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barand ela, J. Aumont, C. Baccigalupi, and A. J. Banday. Planck 2013 results. XIV. Zodiacal emission. *Astron. Astrophys.*, 571:A14, Nov 2014.

[40] R. Fernández-Cobos, P. Vielva, R. B. Barreiro, and E. Martínez-González. Multiresolution internal template cleaning: an application to the Wilkinson Microwave Anisotropy Probe 7-yr polarization data. *MNRAS*, 420:2162–2169, March 2012.

[41] J. Delabrouille, J.-F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, and F. Guilloux. A full sky, low foreground, high resolution CMB map from WMAP. *Astron. Astrophys.*, 493:835–857, January 2009.

[42] J. Delabrouille, J.-F. Cardoso, and G. Patanchon. Multidetector multicomponent spectral matching and applications for cosmic microwave background data analysis. *MNRAS*, 346:1089–1102, December 2003.

[43] H. K. Eriksen, I. J. O'Dwyer, J. B. Jewell, B. D. Wand elt, D. L. Larson, K. M. Górski, S. Levin, A. J. Banday, and P. B. Lilje. Power Spectrum Estimation from High-Resolution Maps by Gibbs Sampling. *Ap.J. Suppl.*, 155(2):227–241, Dec 2004.

[44] H. K. Eriksen, C. Dickinson, C. R. Lawrence, C. Baccigalupi, A. J. Banday, K. M. Górski, F. K. Hansen, P. B. Lilje, E. Pierpaoli, M. D. Seiffert, K. M. Smith, and K. Vanderlinde. Cosmic Microwave Background Component Separation by Parameter Estimation. *Ap.J.*, 641(2):665–682, Apr 2006.

[45] H. K. Eriksen, J. B. Jewell, C. Dickinson, A. J. Band ay, K. M. Górski, and C. R. Lawrence. Joint Bayesian Component Separation and CMB Power Spectrum Estimation. *Ap.J.*, 676(1):10–32, Mar 2008.

[46] Benjamin D. Wandelt, David L. Larson, and Arun Lakshminarayanan. Global, exact cosmic microwave background data analysis using Gibbs sampling. *Phys. Rev. D*, 70(8):083511, Oct 2004.

[47] Jonathan Y. Suen, Michael T. Fang, and Philip M. Lubin. Global Distribution of Water Vapor and Cloud Cover—Sites for High-Performance THz Applications. *IEEE Transactions on Terahertz Science and Technology*, 4(1):86–100, Jan 2014.

[48] A. A. Fraisse, P. A. R. Ade, M. Amiri, S. J. Benton, J. J. Bock, J. R. Bond, J. A. Bonetti, S. Bryan, B. Burger, and H. C. Chiang. SPIDER: probing the early Universe with a suborbital polarimeter. *Journal of Cosmology and Astro-Particle Physics*, 2013(4):047, Apr 2013.

[49] R. Gualtieri, J. P. Filippini, P. A. R. Ade, M. Amiri, S. J. Benton, A. S. Bergman, R. Bihary, J. J. Bock, J. R. Bond, and S. A. Bryan. SPIDER: CMB Polarimetry from the Edge of Space. *Journal of Low Temperature Physics*, 193(5-6):1112–1121, Dec 2018.

[50] J. E. Gudmundsson, P. A. R. Ade, M. Amiri, S. J. Benton, J. J. Bock, J. R. Bond, S. A. Bryan, H. C. Chiang, C. R. Contaldi, and B. P. Crill. The thermal design, characterization, and performance of the SPIDER long-duration balloon cryostat. *Cryogenics*, 72:65–76, Dec 2015.

[51] D. Sutton, J. A. Zuntz, P. G. Ferreira, M. L. Brown, H. K. Eriksen, B. R. Johnson, A. Kusaka, S. K. Næss, and I. K. Wehus. Fast and precise map-making for massively multi-detector CMB experiments. *MNRAS*, 407(3):1387–1402, Sep 2010.

[52] Jonathan R Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, Pittsburgh, PA, USA, 1994.

[53] A. S. Rahlin. *The First Flight of the SPIDER Balloon-Borne Telescope*. PhD thesis, Princeton University, 2016.

[54] T. M. Ruud. *Mapping the Polarized CMB Sky with QUIET and Spider*. PhD thesis, University of Oslo, 2017.

[55] J. A. Tauber, N. Mandolesi, J. L. Puget, T. Banos, M. Bersanelli, F. R. Bouchet, R. C. Butler, J. Charra, G. Crone, and J. Dodsworth. Planck pre-launch status: The Planck mission. *Astron. Astrophys.*, 520:A1, Sep 2010.

[56] M. Sandri, F. Villa, M. Bersanelli, C. Burigana, R. C. Butler, O. D'Arcangelo, L. Figini, A. Gregorio, C. R. Lawrence, and D. Maino. Planck pre-launch status: Low Frequency Instrument optics. *Astron. Astrophys.*, 520:A7, Sep 2010.

[57] The Planck Collaboration. The Scientific Programme of Planck. *arXiv e-prints*, pages astro–ph/0604069, Apr 2006.

[58] J. M. Lamarre, J. L. Puget, P. A. R. Ade, F. Bouchet, G. Guyot, A. E. Lange, F. Pajot, A. Arondel, K. Benabed, and J. L. Beney. Planck pre-launch status: The HFI instrument, from specification to actual performance. *Astron. Astrophys.*, 520:A9, Sep 2010.

[59] J. Delabrouille, M. Betoule, J. B. Melin, M. A. Miville-Deschênes, J. Gonzalez-Nuevo, M. Le Jeune, G. Castex, G. de Zotti, S. Basak, and M. Ashdown. The pre-launch Planck Sky Model: a model of sky emission at submillimetre to centimetre wavelengths. *Astron. Astrophys.*, 553:A96, May 2013.

[60] G. Chon, A. Challinor, S. Prunet, E. Hivon, and I. Szapudi. Fast estimation of polarization power spectra using correlation functions. *MNRAS*, 350:914–926, May 2004.

[61] Antony Lewis, Anthony Challinor, and Anthony Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *Ap.J.*, 538(2):473–476, Aug 2000.

[62] Antony Lewis and Sarah Bridle. Cosmological parameters from CMB and other data: A Monte Carlo approach. *Phys. Rev. D*, 66(10):103511, Nov 2002.

[63] T. Matsumura, Y. Akiba, J. Borrill, Y. Chinone, M. Dobbs, H. Fuke, A. Ghribi, M. Hasegawa, K. Hattori, and M. Hattori. Mission Design of LiteBIRD. *Journal of Low Temperature Physics*, 176(5-6):733–740, Sep 2014.

[64] Kevork N. Abazajian et al. CMB-S4 Science Book, First Edition. 2016.

Part III

Papers

Paper I

A Monte Carlo comparison between template-based and Wiener-filter CMB dipole estimators

Paper II

Planck 2018 results. IV. Diffuse component separation

II

Paper III

NPIPE – Joint Planck LFI and HFI data processing