High fidelity maps of the CMB polarization from the first two campaigns of the Polarbear experiment and their statistical exploitation

par

Davide Poletti

présentée et soutenue publiquement le
18 octobre 2016

Thèse de doctorat de Sciences de l'Univers
dirigée par Radek Stompor

Devant un jury composé de:

Directeur de thèse:	Dr. Radek Stompor	–	APC Paris
Rapporteurs:	Prof. Davide Maino	–	Università degli Studi di Milano
	Dr. Juan Macias-Perez	–	LPSC Grenoble
Membres:	Dr. Nabila Aghanim	–	IAS Orsay
	Dr. Olivier Perdereau	–	LAL Orsay
	Prof. Stavros Katsanevas	–	Université Paris Diderot

Abstract

More than half a century passed since the first detection of the cosmic microwave background (CMB). From that moment on, the increasing accuracy of CMB observations, exploring both its spectral scaling and angular anisotropies, kept refining our understanding of the universe. The current frontier is the polarization of the CMB anisotropies with the B-modes—the curl-like polarization pattern—which are expected to convey invaluable information on both the primordial and the late universe.

In order to extract unique constraints on fundamental physics from the faint B-mode signal, the sensitivity of CMB experiments have been constantly increasing and the size of their datasets growing, approaching trillions of raw data samples. The reconstruction of the map is one of the main steps down any CMB data analysis pipeline, as it allows to compress by many orders of magnitude the data volume, while striving for preserving all the information.

The map-making is made difficult by the fact that typically the raw data are filtered in order to remove unwanted signals of instrumental or astrophysical origins. For instance, in the context of ground-based observations, filtering is frequently used to minimize the impact of low-frequency noise, atmospheric contributions and/or scan-synchronous signals on the resulting maps.

In this thesis, I study the CMB map-making problem in the presence of filtering. To accomplish this task I construct a general filtering operator, capable of unambiguously removing arbitrary and heterogeneous sets of filters, and amend the map-making procedure to incorporate and correct for it. I subsequently investigate the conditions under which this amended map-making procedure can render an unbiased estimate of the sky signal in realistic circumstances and discuss how this new formalism generalizes and incorporates the previous approaches.

This new formalism has been implemented in a massively parallel map-making code, capable of producing high fidelity renditions of the observed sky, whenever this is possible, and of characterising the degenerate modes—the sky patterns that cannot be reconstructed due to the filtering— as well as the map-domain correlations induced by the time-domain filtering.

These methodological developments have been motivated by needs of a specific CMB experiment, Polarbear, which I am a member of. Polarbear is one of the leading ground-based CMB telescopes, at the forefront of the B-mode quest. During my PhD training I have participated in the analysis of the first and second seasons of observations and, specifically, I have led the development of one of its two main data analysis pipelines. This has involved a reconstruction of high fidelity maps of the polarized sky from the Polarbear data and the analysis of their power spectra, capitalizing on the unique features of the map-making code.

In this thesis, I also describe the Polarbear data analysis methodology, the pipeline and its application to the actual Polarbear data. As part of this work, I investigate the role of the large-scale correlations present in the recovered sky maps, demonstrating that they can undermine the performance of popular numerically-efficient power spectrum techniques based on the pseudo-spectrum methods. I then propose a practical approach to compensate for this effect and apply it to the analysis of the Polarbear maps.

Keywords: cosmology – cosmic microwave background radiation – POLARBEAR – map-making

Résumé

Plus d'un demi-siècle s'est écoulé depuis la première détection du fond diffus cosmologique (CMB). A partir de ce moment, la précision croissante des observations du CMB, explorant à la fois son spectre d'émission et ses anisotropies angulaires, a continué à affiner notre compréhension de l'univers. La frontière actuelle est la polarisation des anisotropies du CMB, en particulier avec les modes B, c'est-à-dire la composante rotationelle de la polarisation, qui devraient fournir des informations inestimables à la fois sur l'univers primordial et sur l'univers tardif.

Afin d'extraire des contraintes uniques sur la physique fondamentale à partir du signal faible des modes B, la sensibilité des expériences CMB et le volume des données ont constamment augmenté, ce dernier approchant des milliers de milliards d'échantillons brutes. La reconstruction de la carte est l'une des principales étapes de tout pipeline d'analyse de données CMB, car elle permet de compresser le volume des données de plusieurs ordres de grandeur, tout en s'efforçant de préserver toute l'information physique.

La production des cartes est rendue difficile par le fait que, typiquement, les données brutes sont filtrées afin d'éliminer les contaminations d'origine instrumentale ou astrophysique. Par exemple, dans le contexte des observations au sol, le filtrage est fréquemment utilisé pour minimiser l'impact du bruit à basse fréquence, issu des contributions atmosphériques et / ou des signaux synchrones au scan.

Dans cette thèse j'étudie le problème de la création de cartes CMB en présence de filtrage. Pour accomplir cette tâche, je construis un opérateur de filtrage général, capable de traiter les données sans ambiguïté avec un ensemble de filtres arbitraires et hétérogènes. J'ai également modifié la procédure de création des cartes pour incorporer et corriger l'effet de cette opération. Par la suite j'étudie les conditions dans lesquelles cette procédure de cartographie modifiée peut rendre une estimation non biaisée du signal du ciel dans des circonstances réalistes et je discute comment ce nouveau formalisme généralise et incorpore les approches antérieures.

Ce nouveau formalisme a été implémenté dans un code massivement parallèle, capable de produire des rendus du ciel observé de haute fidélité, lorsque cela est possible, et de caractériser les modes dégénérés, c'est-à-dire les modes du ciel qui ne peuvent pas être reconstruits à cause du filtrage, ainsi que les corrélations induites dans la carte par le filtrage temporel.

Ces développements méthodologiques ont été motivés par les besoins d'une expérience spécifique, Polarbear, dont je suis membre. Polarbear est l'un des principaux télescopes CMB basés au sol, à l'avant-garde de la quête des modes B. Au cours de ma formation doctorale, j'ai participé à l'analyse de la première et deuxième saisons d'observation et j'ai dirigé le développement de l'un de ses deux principaux pipelines d'analyse des données. Cela a impliqué une reconstruction des cartes du ciel polarisé avec une haute fidélité, à partir des données de Polarbear et l'analyse de leurs spectres de puissance, en capitalisant sur les caractéristiques uniques du code de production de cartes.

Dans cette thèse, je décris également la méthodologie d'analyse de Polarbear, le pipeline et son application aux données Polarbear. Dans le cadre de ce travail, j'étudie le rôle des corrélations à grande échelle présentes dans les cartes du ciel estimé, démontrant qu'elles peuvent nuire aux performances des techniques populaires de construction des spectres en puissance, numériquement efficaces, et basées sur les méthodes de pseudo-spectre. Je propose ensuite une approche pratique pour compenser cet effet et l'appliquer à l'analyse des cartes de Polarbear.

Mots clés: cosmologie – fond diffus cosmologique – POLARBEAR – création des cartes

Acknowledgements

I am sincerely thankful to all those people that helped me during these years. I am very thankful to my advisor Radek Stompor for his guidance in this project as well as for the endless chats on the most diverse topics: they taught me a lot about both science and research. Likewise, the other members of the APC Polarbear team Giulio Fabbian, Josquin Errard, Julien Peloton and Maude Le Jeune had an important role in advising me, especially at the very beginning. I also thank our new member Dominic Beck for participating to all the rehearsals of my defence.

I would like to thank Nabila Aghanim, Olivier Perdereau and Stavros Katsanevas for accepting the invitation to be members of my committee and Davide Maino and Juan Macias-Perez for also thoroughly reading and evaluating this thesis.

I am very thankful to Adrian T. Lee and all the members of the Polarbear Collaboration: working on this experiment has been a great experience.

I am grateful to the whole APC, especially to my PhD colleagues and the Cosmology group, for making the laboratory a lively and stimulating environment.

Finally, I would like to thank Mathieu Langer, Jean-Baptiste Durrive and especially Eric Chassande-Mottin for their help in finding this PhD opportunity when I was a master student.

CONTENTS

Chapter 1

THE STANDARD COSMOLOGICAL MODEL

In this chapter I review the current understanding of the universe, its history, its content and dynamics. The pillars of the model are the *cosmological principle* and the *general relativity*. The former constrains the structure of space-time, the latter states how the content of the universe affects the dynamics of the space-time (and vice versa). We then illustrate the *Big Bang* model, which is at the basis of our cosmological model and assumes that the universe underwent a hot and dense phase in its early stages. This phase is predicted to leave a background radiation, filling the whole universe. The observation of this radiation, called *cosmic microwave background* (CMB), is one of the cornerstones of the Big Bang model. The purpose of this work is to contribute to the exploitation of the information conveyed by the CMB.

In order to explain observations, our cosmological model has to postulate two exotic components: a form of matter that interacts only gravitationally with the other components, the *cold dark matter* (CDM), and a form of energy analogous to the energy of the vacuum, the *cosmological constant* (Λ). Ironically, the standard cosmological model is named after the two components that we understand the least and is indeed called ΛCDM.

1.1 The cosmological principle and the geometry of the universe

The cosmological principle states that on sufficiently large scales the universe is spatially homogeneous and isotropic. It was introduced to mathematically describe the universe in the 1920s, when there was no observational evidence for it and when the existence of something beyond the Milky Way was still a subject of animated debate.

The equations of general relativity are too difficult to solve for a general distribution of matter. The cosmological principle simplifies enormously the problem. It implies that the four-dimensional space-time has a three dimensional maximally symmetric subspace. The most general metric of such a space is known as *Friedmann-Lemaître-Robertson-Walker* (FRLW) metric. Adopting (here as in the rest of this work) the natural units $c = \hbar = k_{\rm B} = 1$, the FRLW metric reads

$$\mathrm{d}s^2 = g^{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}t^2 - a(t)^2 \left[\frac{\mathrm{d}r^2}{1 - kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2) \right]$$
$$= \mathrm{d}t^2 - a(t)^2 \left[\mathrm{d}\chi^2 + S_k^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2) \right], \tag{1.1a}$$

where

$$S_k(\chi) = \begin{cases} \sin^{-1}(\chi) & k = 1, \\ \chi & k = 0, \\ \sinh^{-1}(\chi) & k = -1. \end{cases} \tag{1.2}$$

k is a parameter that can assume three values: $1, 0, -1$. Each of them identifies one of the three spatial topologies satisfying the symmetry principle imposed. $k = 0$ corresponds to the flat euclidean space. For $k = 1$ space is an hypersphere: a space with finite volume but no boundaries, easy understandable in analogy with a usual sphere in a 3-d space. $k = -1$ is an hyperbolic space. Though hard to visualize,

similarly to Euclidean space it is infinite but it has (constant) negative curvature.

This frame emphasizes the isotropy. Homogeneity is less apparent but ensured by a being a function of the time only and collecting the entire time dependence of the metric. The function $a(t)$ is called *scale factor*, it is the only dynamical degree of freedom of the metric. For $k = 0$ it has no absolute meaning because the space is scale invariant. Intuitively speaking, for $k = \pm 1$ a is the physical scale at which the non flat nature of space becomes apparent. a carries dimension while while the *comoving coordinates* r, χ, ϕ and θ are dimensionless. Other useful dimensionless quantities are the *conformal time* τ, defined by

$$\mathrm{d}\tau \equiv \frac{\mathrm{d}t}{a(t)}, \tag{1.3}$$

and the *redshift*

$$1 + z(t) \equiv \frac{a_0}{a(t)}. \tag{1.4}$$

The subscript $_0$ means "today". A photon emitted with wavelength λ_e is observed at a different frequency λ_0: because of the expansion of the universe, $\lambda \propto a$. The redshift was indeed first introduced as $1 + z(t_\mathrm{e}) = \frac{\lambda_0}{\lambda_\mathrm{e}}$.

About the conformal time, it enables to write the metric Eq. (1.1) as a^2 multiplying a Minkowski metric. Notice that $\mathrm{d}\tau$ is the comoving distance that light can travel in a time $\mathrm{d}t$ ($\mathrm{d}s^2 = 0$).

1.2   Friedmann equations

In general relativity the link between the metric and energy content of space-time is established by the Einstein field equations

$$G^{\mu\nu} = 8G\pi T^{\mu\nu}. \tag{1.5}$$

G is the gravitational constant. The Einstein tensor $G^{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda g_{\mu\nu}$, is built from the metric and its first and second derivatives (through $R_{\mu\nu}$, the Ricci tensor, and R, the scalar curvature). In the modern view (Gliner, 1966; Zel'dovich, 1968), the term containing the *cosmological constant* Λ is moved to the RHS and interpreted as the contribution from the vacuum, with energy momentum tensor $\Lambda g_{\mu\nu}/8\pi G$.

Because of Eq. (1.5), imposing the cosmological principle not only binds the metric to be Eq. (1.1) but also forces the energy-momentum tensor to have the same symmetry properties as the metric. Consequently, the content of the universe is bound to be a perfect fluid, i.e. to have a *energy momentum tensor* of the form

$$T^{\mu\nu} = (\rho + p)u^\mu u^\nu - pg^{\nu\mu}, \tag{1.6}$$

where ρ and p are the energy density and the pressure of the fluid and u^μ is the four-velocity of the fluid. The physical properties of the fluid define its *equation* of state $p(\rho)$, which in many cases can be approximated as linear $p \equiv w\rho$. A non-relativistic fluid has $w = 0$, a relativistic one has $w = 1/3$, while the cosmological constant is equivalent to a fluid with $w = -1$.

Einstein equations then read

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \tag{1.7a}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \tag{1.7b}$$

called *Friedmann-Lemaître equations*. The covariant conservation of the energy momentum tensor $D_\mu T^{\mu\nu} = 0$ (embedded in Einstein equations, since $D_\mu G^{\mu\nu} = 0$) gives the following useful conservation expression

$$\dot{\rho} + 3(\rho + p)\frac{\dot{a}}{a} = 0. \tag{1.8}$$

If the components of the cosmic fluid interact just gravitationally, the previous equation holds also for each component i independently ($D_\mu T_i^{\mu\nu} = 0$) and, if w_i is constant, it can be written as

$$\rho_i = \rho_{i0}\left(\frac{a}{a_0}\right)^{-3(1+w_i)} = \rho_{i0}\left(1 + z\right)^{3(1+w_i)}. \tag{1.9}$$

As the universe expands components with low w become more important. Notice that in the Friedmann equations Eq. (1.7) the curvature acts like a component with equation of state $w_k = -1/3$ and density $\rho_k = 3k/8\pi G a^2$. Conversely, whether the universe expands and at which rate is determined by its content. For most of the history of the universe a single component with equation of state w_x dominates the energy content of the universe at a given time ($\sum_i \rho_i \simeq \rho_x$) and determines through Eq. (1.7) the time evolution of the scale factor

$$
a \propto \begin{cases} t^{\frac{2}{3(1+w_x)}} & w_x \neq -1 \\ e^{Ht} & w_x = -1. \end{cases} \tag{1.10}
$$

Thus, the smaller the value of w of the dominating component, the faster the universe expands and the expansion is exponential if $w_x = -1$. Eq. (1.7b) make a very important prediction about the past of an expanding universe: if one of the fluids have equation of state $w > -1/3$, $a(t) = 0$ at some point in the past. This singularity is called *Big Bang*.

1.3   Parametrising the dynamical and homogeneous universe

The *Hubble parameter* $H \equiv \dot{a}/a$ is the most important quantity characterizing the dynamical state of the universe. It defines two quantities (which coincide because we are adopting $c = 1$),

$$
t_H = r_H = \frac{1}{H}. \tag{1.11}
$$

The *Hubble time* t_H is the characteristic time of the expansion and the *Hubble radius* r_H is how far the light can travel in such a time, it quantifies the distance of causal connection at a given time.

Today's value of the Hubble parameters is usually expressed in terms of h

$$
H_0 = 100\,h\,\mathrm{kms^{-1}Mpc^{-1}}. \tag{1.12}
$$

The acceleration is quantified by the *deceleration parameter*

$$
q \equiv -\frac{a\ddot{a}}{\dot{a}^2} = -1 - \frac{\dot{H}}{H^2}, \tag{1.13}
$$

which is negative when the expansion is accelerating. The content of the universe is parametrised in terms of a characteristic density: the *critical density*

$$
\rho_c = \frac{3H_0^2}{8\pi G}. \tag{1.14}
$$

It is called critical because $\sum_i \rho_{i0} = \rho_c$ means that the universe is flat ($k = 0$). Moreover, if $\sum_i \rho_{i0} \leq \rho_c$ we get from Eq. (1.7a) that $\dot{a}$ can't be equal to zero and consequently the universe expands forever. The density of a component i today is often expressed as *density parameter*

$$
\Omega_i = \frac{\rho_i}{\rho_c}, \tag{1.15}
$$

Eq. (1.7a) says that $\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k = 1$. The first Friedmann equation can also be rewritten as

$$
H(z) = H_0 E(z) \tag{1.16}
$$

$$
E(z) \equiv \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_r(1+z)^4 + \Omega_\Lambda}. \tag{1.17}
$$

The age of the universe can be computed as

$$
t_0 = \int_0^{t_0} \mathrm{d}t = H_0^{-1} \int_0^\infty \frac{\mathrm{d}z}{(1+z)E(z)} \tag{1.18}
$$

and converges for any choice of the Ωs, if at least one of Ω_r and Ω_m is different from zero.
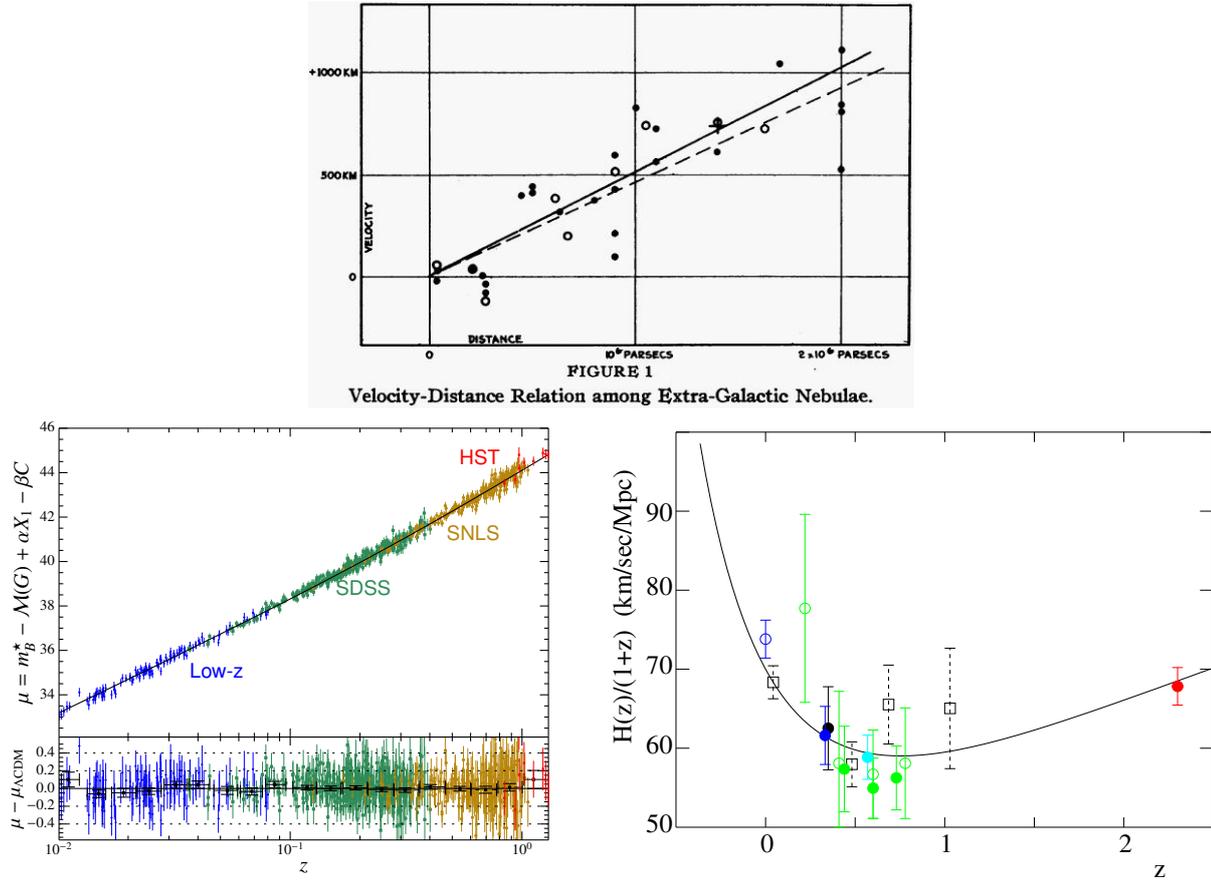
Figure 1.1: *Top panel*: Hubble diagram of the original work by Hubble (1929), using Cepheid variables he showed that more distant nebulae recede faster. It is considered the first evidence of the expansion of the universe. *Bottom left:* state of the art Hubble diagram using SDSS-II and SNLS supernova samples (Betoule et al., 2014). The similar measurements by Perlmutter et al. (1999) and Riess et al. (1998) showed that the expansion of the universe is accelerating, providing strong evidence for the existence of dark energy. *Bottom right:* Hubble parameter as a function of redshift derived from the BAO measurements (full circles) (Busca et al., 2012). By measuring the curve at high redshift the BOSS experiment (red) constrained for the first time the deceleration period during the matter-dominated era.

1.4   Expansion history of our universe

In this section I overview how the observations constrain the dynamical history of the universe, i.e. the evolution of the scale factor. Observing a given object or structure in the sky, we can have a sound measurement of the scale factor at the time of the emission from the redshift Eq. (1.4). Its distance is obtained by comparing an intrinsic property of the source with an observed one: comparing its intrinsic and apparent luminosity we get its *luminosity distance* or, equivalently, comparing its intrinsic size and angular size we get the *angular distance*. Objects of the former types are called *standard candles*, objects of the latter are called *standard rulers*. The Cepheid are variable stars with a strong relationship between the luminosity and the pulsation period, by measuring this latter they can be used as standard candles. Using them E.Hubble derived his famous diagram (Fig. 1.1), showing more distant nebulae receding faster. The measurement turned out to be affected by severe systematics errors but it was considered the first evidence of the expansion of the universe[1]. Using Cepheid stars, more recent estimations by Freedman et al. (2012) combined the measurements of Hubble and Spitzer Space Telescopes to get an estimation of the *Hubble constant*

$$h = 0.743 \pm 0.015 \text{ (stat)} \pm 0.021. \text{ (syst)} \tag{1.19}$$

---

[1]However, this interpretation was still surrounded by big scepticism, notably from Hubble himself

A value of H_0 can be derived also from CMB observations. Latest measurements by Planck Collaboration (2015f) find $h = 0.6751 \pm 0.0064$[2]. The tension between the results of these two probes is subject of debate.

The Supernovae Ia (SNIa) are another type of standard candle thanks to the similarity of their light curve during their explosion (Kowal, 1968). Due to their transient nature, they are more difficult to detect but they enable to probe deeper universe, even at $z \sim 1$. In 1998 using SNIa two independent groups (Riess et al., 1998; Perlmutter et al., 1999) have shown—with their great surprise—that the expansion of the universe is accelerating, unlike what Eq. (1.7b) predicts for a matter- or radiation-dominated universe. Successive measurements confirmed this discovery, Fig. 1.1 show the most up-to-date Hubble diagram obtained with these type of measurements (Betoule et al., 2014). Bearing in mind Eq. (1.7b), these results imply that the current universe is dominated by a mysterious component, called *dark energy*, with equation of state $w < -1/3$.

The distribution of baryonic matter in the universe has a special correlation scale: the *baryon acoustic oscillation (BAO) scale*. It is constant in comoving coordinates and therefore can be used as a standard ruler. Recently, the BOSS/SDSS experiment has been able to probe the distribution of baryonic matter up to $z = 3.5$. Measuring the BAO scale for $2.1 \leq z \leq 3.5$, they have been able for the first time to measure the deceleration period when the universe was dominated by matter (Fig. 1.1).

As mentioned earlier, this model of the universe predicts a singularity in space-time: $a(t) = 0$ for some finite t in the past. The CMB is the most direct evidence of this singularity. It was serendipitously detected for the first time by Penzias and Wilson (1965) and accurately characterised since then. We will illustrate the CMB, its physics and observation in a dedicated chapter (Chapter 3), here we limit ourselves to mention that it sets the age of the universe at (Planck Collaboration, 2015f)[3]:

$$t_0 = (13.807 \pm 0.026)\,\mathrm{Gyr} \tag{1.20}$$

The different probes mentioned study the universe at different ages, rely on different principles and are based on different instruments. Consequently, they are sensitive to different theoretical, astrophysical and instrumental systematics. In modern cosmology measurements are becoming so accurate that the systematic effects sometimes dominate over the statistical ones (e.g. in supernovae measurements), therefore probing the universe through a variety of complementary probes is increasingly important. The fact that all the current probes are in overall agreement is a success of the ΛCDM. Nevertheless, there are already some tensions between different probes. Besides the already mentioned determination of H_0, probably the most important in current observations is the tension between the amplitude of power spectrum of density perturbations inferred from the CMB and the one directly measured from the large-scale structures.

1.5 Cosmic inventory

The universe evolves differently depending on its content. In this section summarise the constituents of the universe and their properties.

1.5.1 Ordinary matter

Ordinary matter includes all the particles that belong to the standard model of particle physics, in particular it includes everything that can interact electromagnetically. It is organised in a variety of structures that can be probed through different means. The main source of visible light is the nuclear fusion inside stars, located inside the galaxies. Galaxies, in turn, are sometimes grouped in clusters of galaxies—gravitationally bound systems containing from hundreds to thousands of galaxies each.

Only when astronomers started to observe the sky outside of the visible frequencies they realised that only a small fraction of the ordinary matter is located in the stars. Galaxies contain also the *interstellar medium*, composed by hydrogen, helium and a small fraction of heavier atoms—the debris of previous generations of stars. However, most of the matter is located in the *intergalactic medium* (IGM), in the form of cold diffuse clouds that did not take part in the galaxy or star formation and, therefore, preserves the primordial composition (see next section). Within the galaxy clusters, the IGM falls in the same potential well that binds together the galaxies, its temperature rises to millions of Kelvin and therefore it becomes visible in the X-rays due to bremsstrahlung emission. In this context it is called *intracluster*

---

[2]The number quoted is the TT, TE,EE+lowP+lensing case
[3]Also in this case, the number quoted is the TT, TE, EE + lowP + lensing case

medium (ICM). Otherwise, the IGM does not emit significantly but can nevertheless be observed because of the absorption of a background light. The most remarkable studies that exploit this effect are probably the IGM mapping through the Lyman-α forest and the estimation of the primordial deuterium fraction.

1.5.2   Radiation

Exception made for the CMB, all the observable radiations were produced well inside the matter or dark energy dominated era and therefore had no impact on the energy budget of the universe. Moreover the radiation from astrophysical emissions has a negligible energy density compared to the CMB. The only moment in the history of the universe in which astrophysical emissions has a relevant cosmological effect is *reionisation*, during which the neutral hydrogen of the universe was ionised by the radiation of quasars and first generations of stars and galaxies. We will discuss further this epoch in Sec. 3.5.

1.5.3   Dark matter

Dark matter, as any dark component, is observed only through gravitational effects. Its existence is established by a variety of probes through the whole history of the universe (the rotation curves of galaxies, galaxies peculiar velocity inside clusters, ICM temperatures, the large-scale structures formation, the angular power spectrum of the CMB, etc.). Electromagnetic coupling aside, dark matter behaves exactly like ordinary matter: it has equation of state $w = 0$ and can collapse to form cosmic structures. Its abundance is roughly 5 times the one of ordinary matter and therefore drives the structure formation, with ordinary matter collapsing on dark matter overdensities.

From the side of the fundamental physics, the simplest models interpret dark matter as weakly interacting massive particles (WIMP). The particle physics community is active in the direct searches for dark matter, trying to detect an excess of interaction in their detectors unexplainable by the standard model. A yearly modulation of this signal would be consistent with the solar system moving in the dark matter bath of our galaxy.

1.5.4   Dark energy

Dark energy is the unknown component of the universe that produce the accelerated expansion. First postulated to explain supernovae measurements, it is nowadays confirmed by a variety of measurements (BAO, CMB, etc.).

The ΛCDM interprets dark energy as vacuum energy, which consequently has a constant density as the universe expands. This explanation is particularly compelling because the fields associated to the particles of the standard model are expected to have non-zero quantum energy density. However, the particle physics prediction and the observed dark energy density differ by many tens of orders of magnitudes. Reconciling the particle physics prediction with the cosmological observations is one of the most active interfaces between the two disciplines.

1.6   The early universe

In this section we give an overview of the processes that characterise the early phases of the universe in the Big Bang model.

1.6.1   The singularity and the Planck epoch

As explained in Sec. 1.2, all expanding isotropic and homogeneous cosmological models containing a perfect fluid with equation of state $w > -1/3$ have a singularity in the past. In the vicinity of the Big Bang, the universe was at extremely high energies and quantum effects involved cosmological scales. In this regime, the model of the universe described previously—based on a non-quantum description of gravity and the strict separation between the concepts of gravity field and matter—is no longer valid. Therefore, the Big Bang model does not state that a singularity necessarily occurred but it does predict an initial epoch in which the standard model of particle physics and the general relativity are no longer adequate to describe the physical phenomena. In the absence of a quantum theory of gravity it is impossible to know when gravitational quantum effects had a cosmological impact. From qualitative arguments we estimate the characteristic quantities of this epoch, called *Planck epoch*: the Planck

time $t_{\mathrm{P}} = \sqrt{G} \sim 10^{-43}\,\mathrm{s}$, the Planck energy $E_{\mathrm{P}} = 1/\sqrt{G} \sim 10^{19}\,\mathrm{GeV}$ (or, expressed as temperature, $T_{\mathrm{P}} = 1/\sqrt{G} \sim 10^{32}\,\mathrm{K}$). Another useful quantity is the Planck mass $m_{\mathrm{P}} \equiv 1/\sqrt{8\pi G} = 2.435 \times 10^{18}\,\mathrm{GeV}$.

1.6.2 Equilibrium in the radiation dominated epoch

The universe in its very early stages was substantially different from Sec. 1.2: the content of the universe was in a very hot and dense state and therefore all the particles of (and beyond) the standard model are thought to have been interacting.

The particle species were able to establish equilibrium if their rate of interaction was faster than the expansion rate

$$\frac{\Gamma}{H} \gg 1 \tag{1.21}$$

This condition was satisfied by all the particles of the Standard Model for $T > 100\,\mathrm{GeV}$. In this regime they are all relativistic, and their energy density is $\rho \sim T^4$ and thus $H \sim \sqrt{\rho}/m_{\mathrm{P}} \sim T^2/m_{\mathrm{P}}$. From a dimensional analysis we can evaluate the cross section of two particles to be $\sigma \sim \alpha^2/T^2$, where α is the structure constant associated to the gauge boson through which the two particles interact. Since the particles are relativistic, $v \sim 1$ and $n \sim T^3$, thus we can estimate the rate of interaction to be $\Gamma \equiv n\sigma v \sim \alpha^2 T$. Assuming $\alpha \sim 10^{-1}$ (reasonable value for an electro-weak interaction) we can evaluate the LHS of Eq. (1.21)

$$\frac{\Gamma}{H} \sim \frac{\alpha^2 m_{\mathrm{P}}}{T} \sim \frac{10^{16}\,\mathrm{GeV}}{T}. \tag{1.22}$$

For $10^{16}\,\mathrm{GeV} \geq T \geq 100\,\mathrm{GeV}$ all the particles are in equilibrium. At higher temperatures (earlier times) the expansion is too fast compared to the interaction rate. At lower temperature (later times) some of the particles are no longer relativistic.

The number density of some massive species is a factor $(mT)^{3/2}\exp(m/T)$ smaller than the number the density of massless particles, which is proportional to T^3. Therefore, when $T \lesssim m$ the massive species become non-relativistic and, if they are still in equilibrium with the relativistic components, their number density exponentially decreases.

Consequently, as long as the massive particles stay in equilibrium with the relativistic thermal bath, the relativistic components dominate the energy content of the universe

$$\rho \simeq \rho_{\mathrm{r}} = \frac{\pi^2}{30} g_*(T) T^4. \tag{1.23}$$

g_* is the *effective number of relativistic degrees of freedom*. It is obtained by summing the degrees of freedom from all the relativistic species i

$$g_* \equiv \sum_{\mathrm{Bosons}} g_i + \frac{7}{8} \sum_{\mathrm{Fermions}} g_i. \tag{1.24}$$

Assuming the particles of the standard model we can predict the g_* at different temperatures see Fig. 1.2. Every time the temperature decreases below the mass of a particle-antiparticle pair, g_* falls because those particles are no longer relativistic. Another sharp drop happens when the temperature reaches the confinement-deconfinement temperature, between 10^{12} and $10^{13}\,\mathrm{K}$: the strong force binds the quarks into hadrons and, exception made for the π mesons, they are all non relativistic.

At the same time, the particle-antiparticle annihilation or the binding of quarks in hadrons releases heat. Since the comoving entropy density is conserved , $a^3 s \propto a^3 g_*(T) T^3 = \mathrm{const}$, we get the relation between temperature and scale factor during the radiation dominated epoch

$$T(t) = T_{\mathrm{P}} \frac{a(t_{\mathrm{P}})}{a(t)} \left[\frac{g_*(T_{\mathrm{P}})}{g_*(T)}\right]^{1/3}. \tag{1.25}$$

The error in using the formula that neglects the varying number of relativistic species, $T \propto a^{-1}$, is very small: the term containing g_* varies by a factor of few while T wildly spans 30 orders of magnitude. The correction is however relevant when estimating, for example, the temperature of the neutrino background, as I expand on in the next section.
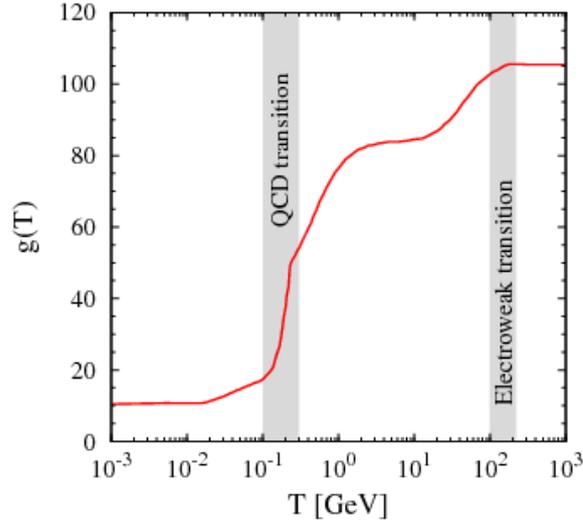
Figure 1.2: Number of relativistic degrees of freedom as a function of temperature according to the standard model. The sharp drop happens at the confinement-deconfinement temperature, when quarks start to bind and form hadrons.

1.6.3 Beyond equilibrium

Equilibrium condition Eq. (1.21) can not be indefinitely maintained. Even though the expansion rate slows down ($H \propto T^2 \propto a^{-2}$), the interaction rate decreases faster, both because $n \propto a^{-3} \propto T^3$ and because at, for different reasons, the cross-section of the relevant interactions drops. In this section we consider few examples.

Nucleosyntesis

As mentioned in Sec. 1.6.2 protons and neutrons (and their antiparticles) start forming as their quarks get confined into hadrons around $T \sim 130\,\text{MeV}$. However, since they are non relativistic, their abundance exponentially decreases with the temperature. We now assume that in the universe there was some asymmetry in the mater and antimatter amounts (see Sec. 1.7.1). Consequently, baryon and antibaryons keep annihilating until no antibaryon is left and the number of baryons freezes.

The high neutron-proton conversion rate keeps the ratio of their abundances equal to $n/p = e^{-Q/T}$, where $Q \equiv m_\text{n} - m_\text{p} = 1.293\,\text{MeV}$. However, the interaction rate $\Gamma_{n\leftrightarrow p} \sim G_\text{F}^2 T^5$ falls faster than the expansion rate $H \sim T^2$ and at $T_\text{fr} \simeq 1\,\text{MeV}$ the weak interaction ceases to keep equilibrium between protons and neutrons. Their relative abundance does not vary with temperature any longer, $n/p = e^{-Q/T_\text{fr}} \simeq 1/6$, and neutrons are said to experience *freeze-out*. Freeze-out is the state when the number density of the particles of the same type is fixed because the interactions with the rest of the cosmic plasma are not frequent enough to keep equilibrium.

When temperature falls below $T \simeq 0.1\,\text{MeV}$ protons and neutrons can combine without being immediately photo-dissociated by a high energy photon, the production of light nuclei can start (notice that by this moment the n/p has dropped to $\sim 1/7$ due to neutron β-decay). This process is called *nucleosynthesis*: a series of 2-body interactions that produce D, ^{3}He, ^{4}He, ^{7}Li. The rates of these interactions depend on the baryon density, usually expressed in terms of n_γ (known from the CMB temperature) $\eta \equiv n_\text{b}/n_\gamma$ which in turn is related to the baryonic fraction $h^2\Omega_\text{b} \simeq \eta \cdot 10^{10}/274$.

The cross sections of the nuclear reactions involved have all been measured and enable to predict the abundances of the light element for different values of η or, equivalently, $h^2\Omega_\text{b}$. Conversely, by measuring these abundances we can measure η and check the consistency between the abundances of the different elements, setting constraints not only on cosmology but also on the Standard Model. The challenging part of these measurements is either finding sites that preserved the primordial abundances or properly accounting for the products of the stellar nucleosynthesis (the strong constraining power of deuterium comes indeed from the fact that stars destroy all the deuterium they produce). The values of η inferred by the measurements of the light element abundances agree between themselves and with the value estimated from the CMB constraints, with the only exception of the lithium (see Fig. 1.3). The
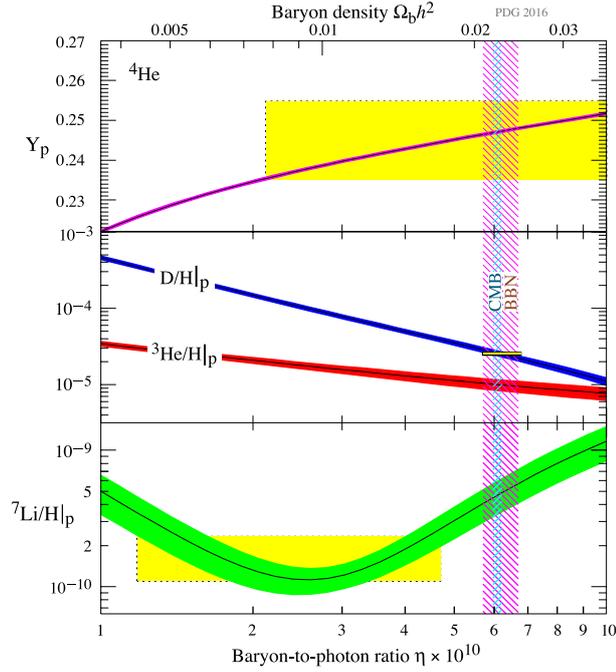
Figure 1.3: The abundances of ^{4}He, D, ^{3}He and ^{7}Li as predicted by standard model of nucleosynthesis, the corresponding bands represent the 95% CL range. Yellow boxes show the observed abundances, while the vertical bans indicate the 95% CL concordance interval from light element measurements (red) and the constraints from the CMB (blue). (Olive et al., 2014)

Lithium abundance inferred from the CMB measurement of $\Omega_b h^2$ is inconsistent at more than 5σ level with the observed Lithium abundance (Cyburt et al., 2008). While striving for better constraints on the astrophysics and nuclear physics involved, the lithium problem is also stimulating explanations involving new physics.

For further readings on nucleosynthesis we refer to Schramm and Turner (1998), Steigman (2007) and Iocco et al. (2009).

Neutrino decoupling

Neutrinos interact only through the weak interaction. As the temperature decreases, the W and Z bosons—responsible for the interaction of the neutrinos—get mass. In this regime the cross section scales as $\sigma G_F^2 T^2$, where G_F is Fermi's constant. The LHS of Eq. (1.21) evaluates to $\Gamma/H \sim G_F^2 m_P T^3 \sim (T/1\,\mathrm{MeV})^3$. Therefore, as the temperature approaches $\sim 1\,\mathrm{MeV}$ neutrinos decouple from the other particles and their (independent) evolution now follows Eq. (1.8). On the contrary, the photon fluid is still interacting with $e^\pm$, its temperature follows Eq. (1.25) and will be affected when $e^\pm$ become non relativistic. Consequently,

$$\frac{T_\nu}{T_\gamma^{\text{after } e^\pm}} = \frac{T_\gamma^{\text{before } e^\pm}}{T_\gamma^{\text{after } e^\pm}} = \left(\frac{g_*^{\text{after } e^\pm}}{g_*^{\text{before } e^\pm}}\right)^{1/3} = \left(\frac{4}{11}\right)^{1/3}. \tag{1.26}$$

Taking into account the different relation between T and n for fermions and bosons, we can express Eq. (1.26) in terms of number density $n_\nu = (3/11)n_\gamma$. If a neutrino family has mass $m_\nu > T_\nu \sim T_\gamma^{(\text{CMB})} \sim 2 \times 10^{-4}\,\mathrm{eV}$, it is not relativistic today and its energy density would be $\rho_\nu = m_\nu n_\nu$. Therefore, the neutrino background can represent a relevant energy content in today's universe:

$$\Omega_\nu = \frac{\sum_\nu m_\nu}{h^2\,94\,\mathrm{eV}} \tag{1.27}$$

where the sum is only over the neutrino and anti-neutrino of the different flavours. Therefore setting constraints on Ω_ν can set tight constraints on the total mass of the neutrinos. This is particularly interesting because atmospheric and solar neutrino oscillations imply a non-zero squared mass differences
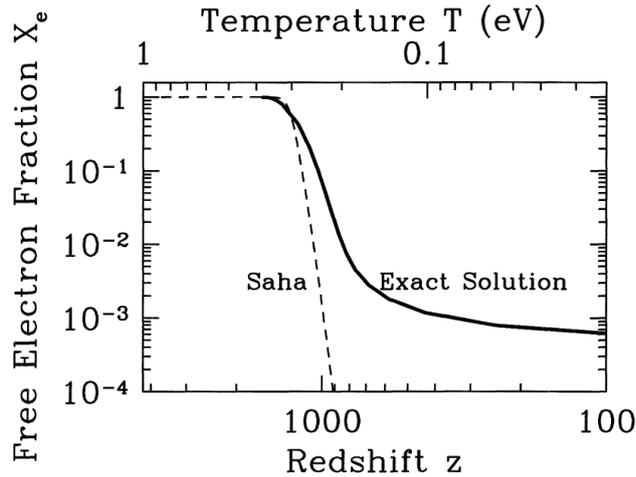
Figure 1.4: The decrease of the free electron fraction as the universe cools. Assuming that at each temperature the reaction Eq. (1.28) reaches the equilibrium proportions, we get the exponential fall described by the Saha equation. However, the temperature drops too fast and the equilibrium is not reached: the exact solution of the free electron fraction differ considerably from the Saha approximation. (Dodelson, 2003)

between the three neutrino flavours (Fukuda et al., 2000), suggesting a lower limit on the neutrino mass of 0.05 eV.

Recombination

When the temperature is few tens of eV the universe is composed of photons, electrons, decoupled neutrinos and ionized light nuclei (mainly hydrogen and some helium). Free electron tend to bind to protons because neutral Hydrogen is energetically favoured: $m_p + m_e - m_H \equiv \epsilon_0 = 13.6\,\text{eV}$. However, energetic photons scatter off the electron and, if they are energetic enough, they ionise again the Hydrogen.

The reaction we are discussing is

$$e^- + p \leftrightarrow H + \gamma \tag{1.28}$$

where the abundance of the left side of the equation is disfavoured by a factor $e^{-\epsilon_0/T}$. As for the deuterium production, the high density of photons above the ionisation energy threshold delays the formation of neutral hydrogen until $T \sim 1/4\text{eV}$. At this temperature the exponential suppression of the ionised state causes a steep decrease in the free electron fraction, this scenario is described by the famous Saha equation. However, at this stage the density of the ingredients in Eq. (1.28) is low and the reaction rate cannot reach the equilibrium value because Eq. (1.21) is not satisfied. As a result, recombination—the period of transition from ionised to neutral hydrogen—requires more time than what Saha's equation would predict and, moreover, some freeze-out fraction of free electrons is left (see Fig. 1.4).

1.7 Problems of the standard model

1.7.1 Baryogenesis

The Big Bang model owes a considerable part of its success to the capability of explaining the abundances of the elements in the universe, giving them a cosmological origin. The standard primordial nucleosynthesis is based on an important assumption: a primordial baryon-antibaryon asymmetry. When the temperature drops below the mass of the baryons they annihilate with antibaryons. Nevertheless, when all antibaryons have annihilated, some baryons are left and can not be converted into lighter particles because all the interactions of the Standard Model preserve the baryon number. For the same reason, this asymmetry can not have been produced when the universe is in the energy regime that is well described by the standard model. Trying to avoid the assumption of a primordial asymmetry, many models have

been proposed to produce dynamically this asymmetry, typically at the GUT energy scale. Without adding further detail, they all satisfy the three Sakharov (1967) conditions: baryon number violation, C and CP violation and interactions out of thermal equilibrium.

1.7.2 Horizon problem

Since the singularity photos have travelled only for a finite time. The comoving distance they can have covered is then

$$\chi_{\mathrm{ph}}(\bar{t}) \equiv \int_0^{\bar{t}} \frac{\mathrm{d}t}{a(t)} = \int_0^{\bar{a}} \frac{\mathrm{d}a}{a\dot{a}}, \tag{1.29}$$

which is called *comoving particle horizon*. From Eq. (1.7a) we know that $\dot{a} \propto \sqrt{\rho a^2} \propto a^{-(1+3w)/2}$ where w is the equation of state of the dominating component of the universe. Therefore the integral Eq. (1.29) converges if $\rho a^2 \xrightarrow{a \to 0} +\infty$, i.e. $w > -1/3$ in the very early universe. This is the case in the Big bang model since radiation ($w = 1/3$) dominates eventually in the past and, consequently, there is a finite comoving distance of causal connection.

The comoving particle horizon at the last scattering of the CMB photons was about 150 Mpc, which corresponds to $\sim 1\,\mathrm{deg}$ in angular size. A large number of CMB regions are expected to be causally disconnected. How is it possible that the CMB is so homogeneous across the sky?

1.7.3 Flatness problem

There is no a priori reason for k to be equal to zero. Nevertheless Ω_k today is very small. This is even more surprising if we consider how it evolves with the expansion of the universe

$$\Omega_k(a) = \frac{\Omega_k}{\Omega_k + \Omega_\Lambda(a/a_0)^2 + \Omega_{\mathrm{m}}(a/a_0)^{-1} + \Omega_{\mathrm{r}}(a/a_0)^{-2}}. \tag{1.30}$$

As a goes to zero, the denominator goes to infinity: in order to match a non-zero value of Ω_k today, its initial value have be very finely tuned. For example, if we suppose that the initial value of Ω_k was set during the Planck epoch, in order to reproduce $|\Omega_k| \sim 1$ its initial value has to be $|\Omega_k(a_{\mathrm{P}})| \lesssim 10^{-60}$.

1.7.4 Magnetic monopoles

As mentioned earlier the very early universe is supposed to be described by a GUT. As the universe cools many of these theories undergo a phase transition and produce topological defects. The most popular is the magnetic monopole, a zero-dimensional defect. Depending on the theory higher dimensional analogues can be produced: strings (one dimensional), domain walls (two dimensional) and textures (three dimensional). All these theories are exotic but nevertheless share this copious production of defects while they are not observed in the universe.

1.7.5 The origins of structures

The framework of the Big Bang model assumes the homogeneity of the universe. However, the universe can be considered as such only on large scales ($> 100\,\mathrm{Mpc}$). The structures that we observe at smaller scales are the result of the gravitational collapse seeded by some initial perturbation. These initial conditions are an assumption of the standard cosmological model. Their natural explanation would be a desirable property of its possible extensions.

1.8 Summary

The standard cosmological model, ΛCDM, has been found extremely successful over the last few decades, explaining satisfactorily most if not all existing observations. The CMB radiation and the abundance of the light elements provide some of the most robust pieces of evidence in its favour. As is, the model is however incomplete. Indeed, it has to postulate the existence of dark matter and energy, the nature of which is still unknown. In addition, it has to assume the presence of initial perturbations. In this latter case a well-motivated theory, called inflation, provides an elegant solution to the initial perturbations problem and, moreover, it successfully address many other theoretical issues of the cosmological model.

Chapter 2

INFLATION

As explained in the previous chapter, the Big Bang model has several problems. Though they are not true inconsistencies, they require very accurate tuning of the initial conditions in order to reproduce the observed universe. Inflation is the most popular solution to these problems: by postulating a period of accelerated expansion in the early universe, it dynamically reproduces these unlikely initial conditions. Moreover, inflation also naturally explains the inhomogeneities at the origin of CMB anisotropies and cosmic structures.

Since it was first postulated (Starobinskiĭ, 1979; Guth, 1981), the inflationary paradigm has been extensively studied and has become the most widely accepted extension of the standard cosmological model.

In the next section we illustrate the qualitative features that (basically) all inflationary models share and explain why the solve the problems mentioned above. We then consider in more detail the simplest inflationary scenario, the inflation produced by a single scalar field. Our discussion focuses on the salient aspects, without being complete nor fully pedagogical. Nowadays any cosmology textbook has at least a chapter on inflation. Among the wide literature available, Baumann and Peiris (2008) can be a concise and self contained first reading, more detailed treatments can be found in Linde (1986), Kolb and Turner (1994) or in the more recent book Liddle and Lyth (2000) and review article Bassett et al. (2005). While writing this chapter, I have also made use of the following lecture notes Baumann (2011) and Langlois (2004).

2.1 The inflation mechanism

In the flatness and horizon problems the comoving Hubble horizon $(aH)^{-1}$ plays a crucial role.

By combining Eq. (1.7a) and Eq. (1.9), we get that, in a universe dominated by a component with equation of state w, the comoving Hubble radius evolves as

$$\frac{1}{aH} \propto a^{\frac{1}{2}(1+3w)}. \tag{2.1}$$

In particular, it grows during matter and radiation dominated epochs. We now show that both the flatness and horizon problems are related to the steady increase of the comoving Hubble parameter.

For the flatness, this is apparent if we use Eq. (1.7a) for rewriting Eq. (1.30) as

$$|\Omega_k(a)| = \frac{1}{(aH)^2} \tag{2.2}$$

For the horizon problem, we first notice that Eq. (1.29) can be rewritten as

$$\chi_{\mathrm{ph}}(t_f) = \tau_f - \tau_i = \int_0^{a_f} \frac{\mathrm{d}\ln a}{aH}. \tag{2.3}$$

and plugging in Eq. (2.1) we notice that $\chi_{\mathrm{ph}}(t) \sim 1/a(t)H(t)$: the particle horizon today is approximately equal to the Hubble radius. As a consequence, unless we introduce by hand homogeneous initial conditions, regions separated by more than a Hubble radius are expected to have completely different

properties because they have never been in causal contact. This contrasts with the observations, which show homogeneous properties far beyond the Hubble radius. The CMB is one of the most striking examples: its properties are homogeneous across the sky despite we observe nearly 10^4 causally disconnected regions.

These two problems can be solved by supposing that in the very early universe, between a time t_i and t_e, there was actually an epoch in which the comoving Hubble radius decreased, this period is called *inflation*. As Eq. (2.1) says, in order to have inflation we need a universe dominated by a component with

$$w < -1/3 \quad \text{during inflation.} \tag{2.4}$$

The condition for a decreasing comoving Hubble radius coincides with the one for an accelerated expansion because $1/aH = 1/\dot{a}$.

Without inflation, the fact that the primordial Ω_k value has to be 10^{-60} times smaller than the current value requires the fine tuning of the initial conditions. Inflation solves the problem because it is capable of dynamically producing arbitrarily small values of Ω_k (see Eq. (2.2)).

Inflation also naturally solve the horizon problem. The presence of an inflationary period in the integral of Eq. (2.3) can make the comoving particle horizon grow arbitrarily big, by taking the comoving Hubble radius at the beginning of inflation arbitrarily large. In the inflationary scenario, comoving regions that in the standard Big Bang model would be causally disconnected could have been in causal contact before inflation shrunk the comoving Hubble radius.

Finally, inflation also solves the monopole problem and naturally explains the origin of the perturbations to homogeneity. The condition expressed Eq. (2.4) says that during inflation the universe undergoes a phase of accelerated expansion during which its size increase by an enormous factor (see Eq. (2.15)). If topological defects were produced before inflation, by the end of inflation they would be extremely diluted, explaining why we do not see any in our observations. Since during inflation microscopic lengths are blown up to cosmological scales, quantum fluctuation before inflation can be the seed of the inhomogeneities at the origin of cosmic structures.

Summarizing, the key of inflation is the accelerated expansion. In the inflationary context this condition is expressed as a constraint on H rather then a. Noticing that $\ddot{a} = a(\dot{H} + H^2)$, the accelerated expansion is equivalent to $-\dot{H}/H^2 < 1$. During inflation it is sometimes convenient to measure the time in terms of *e-folds*

$$\mathrm{d}N \equiv H\mathrm{d}t = \mathrm{d}\ln a. \tag{2.5}$$

We therefore define the *first* and *second slow-roll parameter*—the reason of the name will be apparent later—as follows

$$\varepsilon \equiv -\frac{\dot{H}}{H^2} \tag{2.6}$$

$$\eta \equiv \frac{\mathrm{d}\ln\varepsilon}{\mathrm{d}N} = \frac{\dot{\varepsilon}}{H\varepsilon}. \tag{2.7}$$

They both have to be small: $\varepsilon, |\eta| < 1$. The condition on ε produce the accelerated expansion, the other condition ensures that the acceleration lasts for a sufficiently long period of time.

As a last remark for the rest of this chapter, in the inflationary context the conformal time is typically measured choosing the end of inflation as reference. With this choice $|\tau|$ is approximately equal to the comoving Hubble radius both before and after inflation: in Eq. (2.3) the dominant contribution to the integral comes from early times before inflation and from late times after inflation.

2.2   Implementation with a scalar field

The easiest way produce inflation is through a scalar field ϕ, referred to as *inflaton*, minimally coupled to gravity. This last property means that the interaction between the inflaton and the metric is given just by the kinetic term $\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$ in the action. We will consider the self-interaction term of the inflaton to be given by a potential $V(\phi)$.

The energy-momentum tensor of the inflaton is

$$T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - \mathrm{g}^{\mu\nu}\left(\frac{1}{2}\partial^\sigma\phi\partial_\sigma\phi\mathrm{V}(\phi)\right). \tag{2.8}$$

In the case of a homogeneous field one extract the following energy density and pressure

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{2.9a}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{2.9b}$$

For a universe dominated by a scalar field, the Friedmann equations and the continuity equation now read

$$H^2 = \frac{1}{3m_{\text{Pl}}^2}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right) \tag{2.10a}$$

$$\frac{\ddot{a}}{a} = -\frac{1}{6m_{\text{Pl}}}\left(2\dot{\phi}^2 - V(\phi)\right) \tag{2.10b}$$

$$\ddot{\phi} + 3H\dot{\phi} + V_{,\phi} = 0, \tag{2.10c}$$

where $V_{,\phi} \equiv dV/d\phi$.

If during the dynamics of the scalar field the kinetic energy is sufficiently smaller than the potential energy, the slow-roll conditions are satisfied. This can be easily seen in the extreme case: the so called *slow-roll approximation* (SRA)

$$\frac{1}{2}\dot{\phi}^2 \ll V(\phi) \tag{2.11a}$$

$$|\ddot{\phi}| \ll |V_\phi(\phi)|. \tag{2.11b}$$

The second condition—derived as the time derivative of the first—ensures that the first one is satisfied for sufficiently long period of time. In this approximation Eqs. (2.10) read

$$H^2 \simeq \frac{V(\phi)}{3m_{\text{Pl}}^2} \sim \text{const} \tag{2.12a}$$

$$\dot{H} = -\frac{\dot{\phi}^2}{2m_{\text{Pl}}^2} \simeq -\frac{1}{6}\frac{V_{,\phi}^2}{V} \tag{2.12b}$$

$$3\dot{\phi}H + V_{,\phi} \simeq 0 \tag{2.12c}$$

A nearly constant value of the Hubble parameter means that the expansion is quasi-De Sitter: $a(t) \sim e^{Ht}$. In the SRA the slow-roll parameters take the form

$$\varepsilon \simeq \frac{m_{\text{Pl}}^2}{2}\left(\frac{V_{,\phi}}{V}\right)^2 \ll 1 \tag{2.13a}$$

$$\eta \simeq m_{\text{Pl}}^2\left[\frac{V_{,\phi\phi}}{V} - \frac{1}{2}\left(\frac{V_{,\phi}}{V}\right)\right] \ll 1 \tag{2.13b}$$

These expressions highlight two things. First, inflation is obtained by requiring that the potential have a gentle slope[1]. Second, measuring the slow-roll parameters sets constraints on the potential of the scalar field.

Notice that inflation can be obtained even if the SRA is not satisfied but we wouldn't obtain the simple expressions in Eq. (2.12) and the relationship between potential and slow-roll parameters wouldn't be as simple. This case would require the *potential slow-roll parameters* (Liddle and Lyth, 1992) and infinite hierarchy of slow-roll parameters (Liddle et al., 1994).

2.3 End of inflation and reheating

At some point, inflation has to end and the standard Big Bang scenario begins. By definition, inflation ends when the scalar field acquire relevant kinetic energy and, consequently, the first slow-roll parameter

---

[1]Actually, one also have to constrain the initial velocity $\dot{\phi}$ since it doesn't necessarily satisfy Eq. (2.11b). This is a minor problem because usually Eq. (2.11b) is an attractor solution and wide range of initial values of the velocities quickly converges to it, see e.g. Liddle et al. (1994).

becomes of order unity.

During inflation the universe expanded by a number of e-folds equal to

$$N = \frac{\ln a_{\text{end}}}{\ln a_{\text{start}}} = \int_{t_{\text{start}}}^{t_{\text{end}}} H \text{d}t = \int_{\phi_{\text{start}}}^{\phi_{\text{end}}} \frac{H}{\dot{\phi}} \text{d}\phi \overset{SRA}{\simeq} \frac{1}{m_{\text{Pl}}^2} \int_{\phi_{\text{start}}}^{\phi_{\text{end}}} \frac{V}{V_\phi} \text{d}\phi \tag{2.14}$$

In order to solve the flatness problem, the comoving Hubble radius at the beginning of inflation has to be at least the equal to the comoving Hubble radius today. Since the end value is fixed, this requirement sets a lower bound on the duration of inflation: assuming $H = \text{const}$ during inflation and the temperature at its end $T_{\text{end}} \sim 10^{15}$ GeV, one can set the following famous requirement on the number of e-folds

$$N = \frac{\ln a_{\text{end}}}{\ln a_{\text{start}}} \sim 64. \tag{2.15}$$

The transition between the end of inflation and the early universe of the standard Big Bang scenario is an important part of every inflationary model. Inflation typically leaves the universe in a cold and non-thermal state while the very early universe of the Big Bang model is hot and thermalised. This transition is therefore called *reheating*. Traditionally, in this process the scalar field was treated as a collection of statistically independent particles decaying into the particles of the standard model. In the 1990s this picture has considerably evolved, together with the awareness of the impact of reheating on baryogenesis and nucleosynthesis. More information on reheating can be found in the comprehensive review by Bassett et al. (2005).

2.4   The origin of inhomogeneities

Inflation also has the extraordinary property of naturally explaining the inhomogeneities in the universe that are observed in the CMB anisotropies and seed the structure formation. In this section we will summarize how the quantum fluctuation of the inflation can give rise to inhomogeneities.

2.4.1   Statistical characterisation of the perturbation

We now make a small digression to introduce the power spectrum, in our discussion it will be the most important statistical characterisation of a field X.

The correlation function is defined

$$\xi(\boldsymbol{x}, \boldsymbol{x}') \equiv \langle X(\boldsymbol{x}) X(\boldsymbol{x}') \rangle = \xi(|\boldsymbol{x} - \boldsymbol{x}'|) \tag{2.16}$$

where $\langle \cdots \rangle$ denotes the expected value and the last equality holds when the X is statistically homogeneous and isotropic, which we will assume henceforth.

X can be represented in Fourier domain

$$X_{\boldsymbol{k}} \equiv \int \text{d}^3\boldsymbol{x} \, X(\boldsymbol{x}) \, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{2.17}$$

$$X(\boldsymbol{x}) = \frac{1}{(2\pi)^3} \int \text{d}^3\boldsymbol{k} \, X_{\boldsymbol{k}} \, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{2.18}$$

The power spectrum is defined as

$$\langle X_{\boldsymbol{k}} X_{\boldsymbol{k}'} \rangle \equiv (2\pi)^3 \delta^3(\boldsymbol{k} + \boldsymbol{k}') \mathcal{P}_X(k) \tag{2.19}$$

$$\Delta_X^2(k) \equiv \frac{k^3}{2\pi^2} \mathcal{P}_X(k). \tag{2.20}$$

$\mathcal{P}_X(k)$ so-defined has the property of being the Fourier transform of the correlation function.

The spectral index (or tilt) of the power spectrum is defined as

$$n \equiv \frac{\text{d}\ln \Delta_X^2}{\text{d}\ln k} \tag{2.21}$$

and $n = 0$ means flat spectrum. n has to be defined at some reference scale since it is independent of k only for power-law power spectra. The running of the spectral index is typically parametrised with the second derivative of the power-spectrum. However, no current cosmological observation detected departure from power law primordial density spectra. Beware that for scalar perturbation, the actual parameter traditionally adopted is $n_s \equiv n + 1$.

2.4.2 Perturbed homogeneous universe

We review here few facts about cosmological perturbation theory that are necessary to link inflation to the inhomogeneities of the observable universe. For a complete treatment see e.g the review article by Malik and Wands (2008).

First of all, we have to account for first-order perturbations of the FLRW metric. The most general form is

$$\mathrm{d}s^2 = -(1 + 2\Psi)\mathrm{d}t^2 + 2aB_i\mathrm{d}x^i\mathrm{d}t + a^2[(1 - 2\Phi)\delta_{ij} + E_{ij}]\mathrm{d}x^i\mathrm{d}x^j. \tag{2.22}$$

where the perturbations are further divided in scalar, tensor and vector perturbation. $B_i \equiv \partial_i\mathrm{B} - \mathrm{S}_i$ and $E_{ij} \equiv 2\partial_{ij}\mathrm{E} - \partial_j\mathrm{F}_i - \partial_j\mathrm{F}_i + \mathrm{h}_{ij}$, where Φ, Ψ, E and B are scalar quantity, F_i and S_i are divergence-less vector fields and h_{ij} is a symmetric tensor field satisfying $\partial^i\mathrm{h}_{ij} = \mathrm{h}_i^i = 0$.

Similarly to the metric, the stress-energy tensor in Eq. (1.6) now has to allow for perturbations. Its density and pressure can have time and space dependent perturbations

$$\delta\rho(t, x^i) \equiv \rho(t, x^i) - \bar{\rho}(t) \tag{2.23}$$

$$\delta p(t, x^i) \equiv p(t, x^i) - \bar{p}(t) \tag{2.24}$$

and it also can have a first order anisotropic stress $\Sigma^{\mu\nu}$ satisfying $\Sigma^{00} = \Sigma^{i0} = \Sigma^{ii} = 0$.

These perturbations are not necessarily physical, they could have been produced simply by applying a first order gauge transformation to the FLRW metric. For example, in constant-time slices of a FLRW universe the content of the universe is homogeneous but if we squeeze the spatial coordinates at some point in space an unphysical underdensity appears.

Ψ, Φ, E, B, F_i, S_i, $\delta\rho$, δp are indeed gauge dependent. On the contrary, the tensor component of Σ^{ij} and h^{ij} are gauge independent. The ambiguities related to the gauge dependency of the perturbations are avoided by expressing the results in terms of gauge independent quantities (Bardeen, 1980; Ellis and Bruni, 1989). The one relevant for our discussion will be

$$\mathcal{R} \equiv \Phi - \frac{H}{\bar{\rho} + \bar{p}}\delta q. \tag{2.25}$$

δq is the scalar component of the 3-momentum density. In the comoving frame, it is by definition zero and $\mathcal{R}$ reduces to the curvature Φ. $\mathcal{R}$ is indeed called *comoving curvature perturbation.*

2.4.3 Perturbations during inflation

Plugging the perturbed metric and stress-energy tensor into the Einstein equations Eq. (1.5) we get the differential equations governing the time evolution of the perturbations. Einstein equation, in general, are coupled second-order partial differential equations but, thanks to the symmetries of the background, to first order they simplify to a set of uncoupled ordinary differential equations in which the scalar, vector and tensor components evolve independently.

The Einstein equation for the $\boldsymbol{k}$ Fourier mode of the scalar perturbation and the amplitude h of one of the two polarizations of the tensor perturbations reduce to the *Mukhanov equation*

$$v_k'' + \left(k^2 + \frac{z''}{z}\right)v_k = 0 \tag{2.26}$$

where the prime denotes the derivative with respect to the conformal time,

$$v_k = \begin{cases} ah_k m_{\mathrm{Pl}}/\sqrt{2} & \text{tensor} \\ z\mathcal{R}_k m_{\mathrm{Pl}} & \text{scalar} \end{cases} \tag{2.27}$$

and

$$z = \begin{cases} a & \text{tensor} \\ a|\dot{\phi}|/(H m_{\mathrm{Pl}}) = a\sqrt{2\varepsilon} & \text{scalar}. \end{cases} \tag{2.28}$$

In general the Mukhanov equation have to be solved numerically because of the dependence of z from the evolution of the background. However, we can get interesting insights by considering the SRA, for which we get $\varepsilon \to 0$ and a quasi-De Sitter evolution for the background. In this case also for scalar

perturbations $z''/z \to a''/a$ and $a''/a \to 2/\tau^2$. In this approximation $(2k)^{1/2}e^{-k\tau}(1 - i/k\tau)$ and its complex conjugate are a basis for the solutions of Eq. (2.26). The standard quantization of v_k [2] (similar to the canonical quantization of the harmonic oscillator) allows to compute the following expected value on the vacuum state

$$\frac{1}{a^2}\langle v_{\boldsymbol{k}}v_{\boldsymbol{k'}}\rangle = (2\pi)^3\delta^3(\boldsymbol{k}+\boldsymbol{k'})\frac{H^2}{2k^3}(1 + k^2\tau^2) \tag{2.29}$$

$$\xrightarrow{|k\tau|<<1} (2\pi)^3\delta^3(\boldsymbol{k}+\boldsymbol{k'})\frac{H^2}{2k^3} \tag{2.30}$$

At some point the modes k leave the comoving horizon because during inflation it shrinks, the second equation shows that when it happens the amplitude of the power spectrum freezes. If we allowed slight variations of H, its value in Eq. (2.30) would be the one at horizon crossing, $H_\star$. Using Eq. (2.27) we estimate the spectrum of scalar and tensor perturbations

$$\Delta_s^2(k) \equiv \Delta_{\mathcal{R}}^2(k) = \frac{1}{8\pi^2}\frac{H_\star^2}{m_{\mathrm{Pl}}^2}\frac{1}{\varepsilon_\star} \tag{2.31}$$

$$\Delta_t^2(k) \equiv 2\Delta_h^2(k) = \frac{2}{\pi^2}\frac{H_\star^2}{m_{\mathrm{Pl}}^2} \tag{2.32}$$

The amplitude of the tensor perturbations is often expressed in terms of the *tensor to scalar ratio*

$$r \equiv \frac{\Delta_t^2}{\Delta_s^2} = 16\varepsilon_\star. \tag{2.33}$$

These power spectra stay frozen until the k modes re-enter the comoving Hubble horizon in the standard Big Bang model. They therefore set the initial condition for the perturbations of the observable universe. $H_\star^2 m_{\mathrm{Pl}}^2$ is basically the energy at which inflation happened. $\varepsilon_\star$ quantifies how close the universe was to De-Sitter. Measuring scalar perturbations constrains a combination of these two quantities while measuring the tensor perturbations would directly probe the energy scale of inflation. Notice that H and ε are assumed to vary very little during inflation. Therefore, the spectra of the perturbations are predicted to be scale invariant, their small tilt would be the consequence of their slight variation during inflation – imprinted on the power spectrum because different scales cross the horizon at different times.
These results actually exploited only the approximation of the quasi-De Sitter expansion. The SRA allows to go further and predict more relations between observables and inflationary quantities, most notably $n_s - 1 = 2\eta - 4\varepsilon$, the consistency relation

$$r = -8n_t \tag{2.34}$$

and

$$V^{\frac{1}{4}} \simeq \left(\frac{r}{0.01}\right)^{\frac{1}{4}} 10^{16}\mathrm{GeV} \tag{2.35}$$

2.5 Constraints on inflation

The most important (and currently the only) probe for inflation is the CMB, discussed in the next chapter. In addition to solving the problems of the standard model mentioned earlier, inflation makes predictions that found confirmation in the CMB observations. The simplest inflationary models predict Gaussian primordial perturbations and current CMB observations set remarkable constraints on the primordial non-Gaussianity (Planck Collaboration et al., 2015c). Simplest inflationary models also predict a nearly scale-invariant spectrum for scalar perturbations, with a slight tilt due to the different time at which different scales cross the horizon. This departure from the scale invariance was recently detected by Planck Collaboration (2015g), $n_s = 0.9677 \pm 0.0060$ (remind that the scale-invariant spectrum corresponds to $n_s = 1$).

In general, CMB observations characterise inflation mostly by constraining the power spectrum of primordial scalar and tensor perturbations. Their "lowest order" properties are parametrised by n_s and

---

[2]The normalisation also assumes the choice of the Minkowski vacuum in the far past as vacuum state.
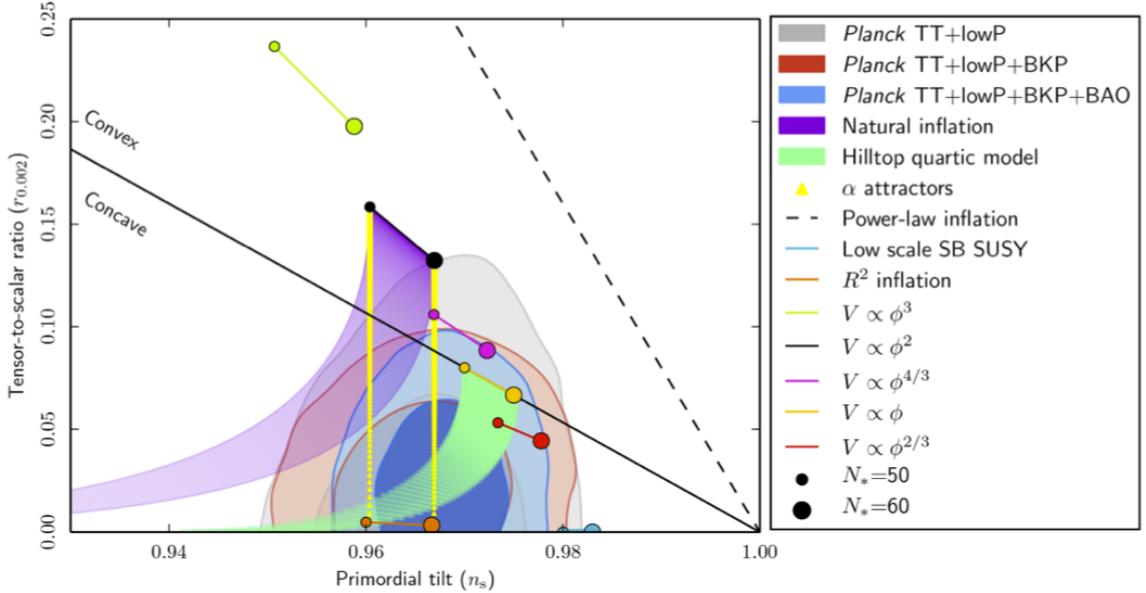
Figure 2.1: State-of-the-art constraints on the $n_s - r$ space. Deep and light contour regions are the 68% and 95% CL regions from *Planck* alone and in combination with its cross-correlation with BICEP2/*Keck Array* and/or BAO data. The other lines represent the prediction of different inflationary model. Most notably ϕ^2-inflation and natural inflation are now disfavoured compared to other models that predict lower values of r. (Planck Collaboration et al., 2015b)

r, of which every inflationary model give a prediction. Therefore, the constraints on inflation set by an experiment are typically expressed as a contour plot in the n_s-r plane. As Fig. 2.1 shows, currently we have only upper bounds on r. In the next chapter we explain that the polarization B-modes of the CMB are the most promising probe of primordial gravitational waves. The measurement of r is indeed one of the central motivations of the hunt for primordial B modes.

2.6 Summary

Inflation is an attractive theory for the very first instants after the Big Bang. It proposes an elegant solution to the numerous problems of the cosmological model, such as flatness, horizon, and abundance of monopoles. It also explains the origin of the initial perturbations, making predictions about their nature and amplitude as a function of scale.

These predictions are testable. In particular, the existence of the primordial gravity waves—expected in many inflationary models—is considered one of the possibly most tantalizing signatures of inflation. At present, CMB anisotropies, and in particular B-mode CMB polarization, are the best way to detect such signature. While confirming inflation, this will also provide unique insights on the physics at the energies 12 orders of magnitude higher than those currently investigated by LHC.

Chapter 3

THE COSMIC MICROWAVE BACKGROUND

One of the consequences of the Big Bang model, as summarised in Chapter 1, is the existence of a background radiation, the CMB. It was first predicted by Alpher and Herman (1948) and serendipitously detected nearly twenty years later by Penzias and Wilson (1965). Its observation has been a crucial evidence in favour of the Big Bang model.

The CMB is the oldest observable light and thus it is a unique window on the primordial universe. Great efforts were already made to exploit its scientific potential. Remarkable turning points were the measurement by the COBE satellite of its black body spectrum (Mather et al., 1990), which left no doubt on the primordial origin, and the detection of the anisotropies (Smoot et al., 1992). The information they convey is a goldmine for cosmology and their increasingly accurate characterisation led to precise constraints on a number of cosmological parameters.

In this chapter we explain why the CMB is an invaluable probe for cosmology and how it can be used to constrain the primordial and late universe. We also describe the observational status and the current forefront: the B-mode polarization.

3.1 Primordial photon background

In this section we explain further the link the between the spectrum of the CMB and the physics of the early universe.

To start with, a photon gas in equilibrium with a thermal bath at temperature T has the *Bose-Einstein* distribution function[1]

$$f(\nu; T, \mu) = \frac{1}{e^{(x-\mu)} - 1}. \tag{3.1}$$

where, for consistency with the literature, in this section $x \equiv 2\pi\nu/T$, ν is the frequency of the photon and μ is the chemical potential.

If the gas is free to gain or lose photons, its chemical potential is $\mu = 0$. This case is called *black body*. If $\mu > 0$ the gas has a deficit of photons, while it has an excess of photons if $\mu < 0$.

In the primordial plasma photons interact mainly with electrons. The most relevant processes are Compton scattering, double Compton scattering and bremsstrahlung. They are very efficient in the very early universe but progressively become negligible as the universe expands. *Compton scattering* consists in $e + \gamma \to e + \gamma$, energy can be efficiently transferred from electron to photons (and vice versa) but the total number of particles in each of the species is left unchanged. On the contrary, the *double Compton* scattering and *bremsstrahlung* produce (or destroy) a photon in each interaction.

Consider a particle with quadrimomentum P^μ in the FLRW universe, we can show that the time component of the geodesic equation can be written as $p^2 \equiv -P^i P^j g_{ij} \propto 1/a^2$. For a photon this implies that $p = \nu \propto 1/a$. The fact that Eq. (3.1) can be written in terms of $x \equiv \nu/T$ means that the Bose-Einstein distribution is preserved by the expansion of the universe but the temperature parameter decreases as $T \propto 1/a$.

---

[1] remind that we are using $k_{\mathrm{B}} = \hbar = 1$
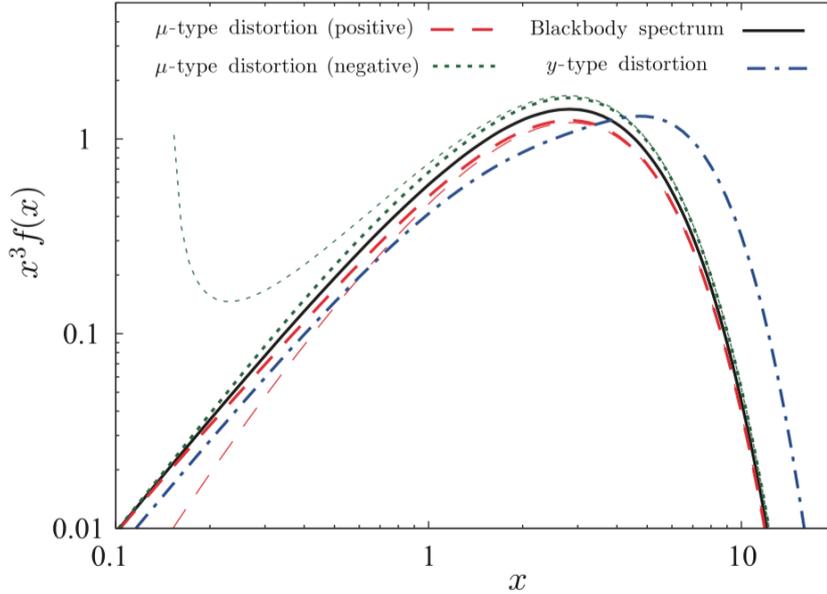
Figure 3.1: The black-body spectrum (black solid line) and examples of μ-type and y-type distortions. The red (dashed) and green (dotted) lines represent μ-type distortions with a negative and positive chemical potential, respectively. The blue (dot-dashed) line is for the y-type distortion. For comparison, the black-body spectrum with the same temperature is plotted as the black (solid) line. The thin lines show the spectra in the Bose–Einstein distributions with constant chemical potentials. (Tashiro, 2014)

3.1.1   The black-body spectrum

Before $z \simeq 2 \times 10^6$ (i.e. temperature higher than $T \simeq 500\,\text{eV}$) all the three interactions mentioned above were very efficient. Photon momenta are easily redistributed and, consequently, the radiation is thermalised and follows the distribution in Eq. (3.1). Moreover, double Compton and bremsstrahlung are able to increase and decrease the number of photons. Their contribution to the Boltzmann equation of photons is proportional to $1 - f(e^x - 1)$, thus enforcing $\mu = 0$ in the photon distribution: radiation has a black body spectrum. Fig. 3.1 shows the black body spectrum together with some (exaggerated) examples of the following spectral distortions

3.1.2 μ-type distortions

After $z \simeq 2 \times 10^6$ double Compton and bremsstrahlung become inefficient: their interaction rate no longer satisfies Eq. (1.21). Compton scattering with electrons still thermalise the photons but no mechanism can keep $\mu = 0$, the so-called μ-type of distortions can be produced. In particular, if energy is injected into the plasma the temperature of the electrons rises. The temperature of the photons rises too but, since the number of photons cannot change, there is a deficit of photons and thus μ increases. Conversely, if for some reasons photons are injected the temperature of the radiation keeps tracking the temperature of the electrons but the excess in the number of photons produces $\mu < 0$.

Notice that the $e^\pm$ annihilation happens well before $z \simeq 10^6$, therefore the consequent massive photon injection does not produce spectral distortion. The electrons present at this epoch are the leftover of the matter-antimatter asymmetry.

3.1.3 y-type distortions

At $z \simeq 2 \times 10^4$ also the Compton scattering starts to become inefficient and energy injections in the photons cannot be thermalised. In the non-relativistic limit for the electrons, the evolution of the distribution function is described by the Kompaneets equation (Kompaneets, 1957). For a black body radiation in an optically thin cloud of electrons, it simplifies to (Zeldovich and Sunyaev, 1969)

$$\frac{\delta f}{f} = y \frac{x e^x}{(e^x - 1)^2} \left(x \coth(x/2) - 4 \right). \tag{3.2}$$

where

$$y \equiv \int dt n_e \sigma_{\mathrm{T}} \frac{T_e - T}{m_e} \tag{3.3}$$

is called *Compton parameter*. Therefore, if $T_e > T$ some photons are scattered to higher energies creating the so-called *y-type distortion*. A *y*-type distortion can also be produced if two black bodies with a temperature difference ΔT are averaged: $y = \frac{1}{2} \langle \Delta T/T \rangle$.

3.1.4 Sources of spectral distortions

A wide variety of phenomena lead to spectral distortions: recombination, reionization, dissipation of primordial density fluctuations, decaying or annihilating particles, primordial magnetic fields, etc. They involve both standard and exotic physics, making the spectral distortions of the CMB a particularly valuable probe for both testing our understanding of the universe and constraining out standard models of cosmology and particle physics. For a concise discussion on the spectral distortions predicted by the ΛCDM see Chluba (2016), while a recent review on spectral distortions can be found in Tashiro (2014).

3.1.5 Constraints on CMB spectrum

The Far Infrared Absolute Spectrophotometer (FIRAS) (Mather et al., 1990) aboard the COBE satellite (Boggess et al., 1992) accurately measured for the first (and last) time the CMB spectrum on a wide range of wavelengths, from 2 to 21 cm. The temperature of the CMB radiation was measured to be $T = (2.728 \pm 0.004)$ K at 95% confidence level. No deviation from the black body spectrum was found, setting the following 95% confidence level upper limits $|y| < 1.5 \times 10^{-5}$ and $|\mu| < 9.0 \times 10^{-5}$. These results were confirmed by the COBRA rocket experiment (Gush et al., 1990) soon after. The CMB power spectrum was further probed from the ground at low frequencies by the TRIS (Gervasi et al., 2008) and the ARCADE (Seiffert et al., 2011) experiments, setting upper bounds on the chemical potential respectively equal to $|\mu| < 6.0 \times 10^{-5}$ at 95% confidence level and $|\mu| < 6.0 \times 10^{-4}$ at 2σ level. A new satellite mission, PIXIE (Kogut et al., 2011), is being considered, aiming at improving by a factor $\sim 10^3$ the constraints set by the COBE satellite.

3.2 Describing the CMB anisotropies

Before discussing the origin and the implications of the CMB anisotropies, we introduce the physical quantities and the mathematical tools necessary for their description.

3.2.1 Stokes parameters

The cosmic microwave background is polarised and is characterised with the help of the *Stokes parameters*.

Consider a monochromatic plane wave with frequency ω propagating along the z axis, it can be fully described in terms of the time evolution of the electric vector on the x-y plane:

$$E_x(t) = A \cos(\omega + \phi_x) \tag{3.4}$$
$$E_y(t) = B \cos(\omega + \phi_y). \tag{3.5}$$

The x and y components of the field are independently oscillating at the same frequency with a relative phase $\phi_x - \phi_y$. We can also express the electric field as a single complex 2-vector $\boldsymbol{E}(t) = \boldsymbol{E}_0 e^{i\omega t}$, where the information about the phase offset is moved in the argument of the components of the complex vector $\boldsymbol{E}_0$.

This simple description, however, is not adequate for the observation of the cosmic microwave background: the photons observed at a given frequency have indeed random phases and, even if the photons were coherent, in the microwave-band observations we typically measure power, quantified as the average power absorbed on a period of time $\Delta t \gg 1/\omega$. In this section we will denote this averaging operation as $\langle \cdots \rangle$. Therefore, the formalism adopted for describing the CMB measurements is the one of the *Stokes*

parameters, which are defined as

$$I \equiv \langle E_x^2 + E_y^2 \rangle$$
$$Q \equiv \langle E_x^2 - E_y^2 \rangle$$
$$U \equiv \langle 2\mathrm{Re}(E_x E_y^*) \rangle$$
$$V \equiv \langle -2\mathrm{Im}(E_x^* E_y) \rangle. \qquad (3.6)$$

I is the total intensity, while the other parameters describe the polarization of the radiation. Q and U characterise the linear polarization state while V the circular one.

One can show that the following relation holds

$$Q^2 + U^2 + V^2 \leq I^2. \qquad (3.7)$$

For the monochromatic wave considered above the equality holds, but doesn't for incoherent radiation. We define the *polarization intensity* as

$$P^2 \equiv Q^2 + U^2 + V^2. \qquad (3.8)$$

The relation in Eq. (3.7) express the fact that the polarized energy cannot exceed the total energy.

It is important to notice how the Stokes parameters change under parity change and rotation of the reference coordinate system. I is a scalar while V is a pseudo-scalar (it changes sign under parity transformation). More important to us, Q and U change if the x-y plane is rotated by an angle θ. One can easily show that in the new system x'-y'

$$(Q' \pm iU') = e^{\mp 2\theta i}(Q \pm iU) \qquad (3.9)$$

or, equivalently, one can gather Q and U in a 2-vector and express the same relation as

$$\begin{pmatrix} Q \\ U \end{pmatrix}' = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ -\sin(2\theta) & \cos(2\theta) \end{bmatrix} \begin{pmatrix} Q \\ U \end{pmatrix}. \qquad (3.10)$$

This relations express the spin-2 nature of the linear polarization field that can thus be represented as a head-less vector with magnitude P, if $V = 0$, and orientation

$$\alpha = \frac{1}{2}\arctan\left(\frac{U}{Q}\right) \qquad (3.11)$$

3.2.2 Radiation on a sphere

Astronomical observations measure the Stokes parameters illustrated in the previous section in many directions $\hat{\boldsymbol{n}}$ (i.e. many (θ, ϕ) angles in polar coordinates), potentially covering the whole celestial sphere. This section illustrates how to describe and characterise a radiation on a sphere. In Sec. 3.2.1 we explained that the radiation field can be decomposed into a scalar component, I, and a spin ± 2 component $Q \mp iU$. From now on we will ignore the pseudo-scalar component V because it is not produced by Compton scattering, the dominant process when the CMB radiation was generated. In addition, in the CMB context, the intensity is typically expressed as temperature fluctuation T with respect to the average temperature of the CMB (see Sec. 3.3.1).

Spin-weighted spherical harmonics

We can expand these components in the corresponding spin weighted basis

$$T(\hat{\boldsymbol{n}}) = \sum_{lm} a_{lm}\, Y_{lm}(\hat{\boldsymbol{n}}) \qquad (3.12)$$

$$(Q \mp iU)(\hat{\boldsymbol{n}}) = \sum_{lm}\, {}_{\pm 2}a_{lm}\; {}_{\pm 2}Y_{lm}(\hat{\boldsymbol{n}}), \qquad (3.13)$$

where Q and U are defined with respect to the $(\hat{\boldsymbol{e}}_\theta, \hat{\boldsymbol{e}}_\phi)$. The $Y_{lm}(\hat{\boldsymbol{n}})$ functions are the usual scalar *spherical harmonics* while ${}_sY_{lm}(\hat{\boldsymbol{n}})$ functions are the *spin-weighted spherical harmonics* with spin s

([Goldberg](), 1967). They are related to the scalar spherical harmonics by spin-rising and spin-lowering operators

$$\eth = -\sin^s\theta\left[\frac{\partial}{\partial\theta} - i\csc\theta\frac{\partial}{\partial\phi}\right]\sin^{-s}\theta \tag{3.14}$$

$$\bar{\eth} = -\sin^{-s}\theta\left[\frac{\partial}{\partial\theta} - i\csc\theta\frac{\partial}{\partial\phi}\right]\sin^s\theta \tag{3.15}$$

where s is the spin of the function on which the operator is being applied. If f is a spin-s function, then $\eth f$ has spin $s+1$. Conversely, if f has spin-s function, then $\bar{\eth} f$ has spin $s-1$.
In particular, when applied to spherical harmonics,

$$\eth\,_sY_{lm} = [(l-s)(l+s+1)]^{1/2}\,_{s+1}Y_{lm} \tag{3.16a}$$

$$\bar{\eth}\,_sY_{lm} = -[(l+s)(l-s+1)]^{1/2}\,_{s+1}Y_{lm} \tag{3.16b}$$

therefore the spherical harmonics with spin ± 2 can be derived from the spin-0 spherical harmonics as

$$_2Y_{lm} = [(l-2)!/(l+2)!]^{1/2}\eth\eth Y_{lm} \tag{3.17}$$

$$_{-2}Y_{lm} = [(l-2)!/(l+2)!]^{1/2}\bar{\eth}\bar{\eth} Y_{lm}. \tag{3.18}$$

Combining the these two relations we obtain that $\eth\eth\bar{\eth}\bar{\eth} = \bar{\eth}\bar{\eth}\eth\eth = \nabla^2(\nabla^2+2)$.

E-B decomposition

A vector field (spin 1) on a two dimensional manifold can be decomposed into a gradient and a curl component by the Helmholtz theorem. Similarly the spin ± 2 polarization field can be naturally expressed in terms of its gradient and curl components, respectively called E and B fields in analogy with the electromagnetism. They are both spin-0 fields, their harmonic coefficients can be derived from those of the harmonic coefficients of the spin-2 field

$$_Ea_{\ell m} = -\tfrac{1}{2}(_2a_{\ell m} + _{-2}a_{\ell m}) \tag{3.19}$$

$$_Ba_{\ell m} = \tfrac{i}{2}(_2a_{\ell m} - _{-2}a_{\ell m}). \tag{3.20}$$

Conversely, Q and U can be written in terms of E and B

$$Q = -\sum_{\ell m}(_Ea_{\ell m}\,_1X_{\ell m} + i\,_Ba_{\ell m}\,_2X_{\ell m}) \tag{3.21}$$

$$U = -\sum_{\ell m}(_Ba_{\ell m}\,_1X_{\ell m} + i\,_Ea_{\ell m}\,_2X_{\ell m}) \tag{3.22}$$

where $_1X_{\ell m} \equiv (_2Y_{\ell m} + _{-2}Y_{\ell m})/2$ and $_2X_{\ell m} \equiv (_2Y_{\ell m} - _{-2}Y_{\ell m})/2$.
In the vector notation of Eq. (3.10), Eq. (3.22) reads

$$\boldsymbol{P} = -\sum_{\ell m}(_Ea_{\ell m}\,_E\boldsymbol{Y}_{\ell m} + _Ba_{\ell m}\,_B\boldsymbol{Y}_{\ell m}) \tag{3.23}$$

where the basis for the electric and magnetic component of the spin-2 spherical harmonics is defined as

$$_E\boldsymbol{Y}_{\ell m} \equiv \begin{pmatrix} _1X_{\ell m} \\ -i\,_2X_{\ell m} \end{pmatrix} = \boldsymbol{D}_E Y_{\ell m} \tag{3.24}$$

$$_B\boldsymbol{Y}_{\ell m} \equiv \begin{pmatrix} i\,_2X_{\ell m} \\ _1X_{\ell m} \end{pmatrix} = \boldsymbol{D}_B Y_{\ell m} \tag{3.25}$$

The second order differential operators in the second equality are defined as

$$\boldsymbol{D}_E \equiv \frac{1}{2}\begin{pmatrix} \eth\eth + \bar{\eth}\bar{\eth} \\ -i(\eth\eth - \bar{\eth}\bar{\eth}) \end{pmatrix} \tag{3.26}$$

$$\boldsymbol{D}_B \equiv \frac{1}{2}\begin{pmatrix} i(\eth\eth - \bar{\eth}\bar{\eth}) \\ \eth\eth + \bar{\eth}\bar{\eth} \end{pmatrix}. \tag{3.27}$$

The following properties of these operators

$$\boldsymbol{D}_E^\dagger \cdot \boldsymbol{D}_B = \boldsymbol{D}_B^\dagger \cdot \boldsymbol{D}_E = 0 \tag{3.28}$$

$$\boldsymbol{D}_E^\dagger \cdot \boldsymbol{D}_E = \boldsymbol{D}_B^\dagger \cdot \boldsymbol{D}_B = (\nabla^2 + 2)\nabla^2 \tag{3.29}$$

make apparent that any B mode is orthogonal to any E mode.

Small scale limit

In the small scale limit the sky can be considered to be flat and the harmonic decomposition becomes the Fourier transform. In this case the spin rising and lowering operators take the form

$$\eth = -(\partial_x + i\partial_y) \tag{3.30}$$

$$\bar{\eth} = -(\partial_x - i\partial_y) \tag{3.31}$$

and Eq. (3.22) takes the form

$$\boldsymbol{P}(\boldsymbol{x}) \equiv \begin{pmatrix} Q \\ U \end{pmatrix}(\boldsymbol{x}) = \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \left[E(\boldsymbol{k}) \begin{pmatrix} \cos(2\phi) \\ \sin(2\phi) \end{pmatrix} + B(\boldsymbol{k}) \begin{pmatrix} -\sin(2\phi) \\ \cos(2\phi) \end{pmatrix} \right] e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{3.32}$$

The wave vector $\boldsymbol{k} \equiv k(\cos\phi, \sin\phi)$ is the direction along which the magnitude of the vector $\boldsymbol{P}$ varies. $\boldsymbol{P}$ and $\boldsymbol{k}$ are not necessarily aligned. We can make two interesting observations from this expression. First, The E/B decomposition becomes local in Fourier space, as in Eq. (3.20). Second, consider a single Fourier mode $\boldsymbol{P}(\boldsymbol{k})$, if we perform a rotation such that in the new frame $\phi = 0$, we obtain that $E = Q$ and $B = U$. In other words for the modes of type E the vector $\boldsymbol{P}$ is either aligned or orthogonal to the direction in which its magnitude increases, while in the modes of type B the vector $\boldsymbol{P}$ forms an angle of 45 deg. This last statement is true also if we don't restrict ourselves to a single Fourier mode $\boldsymbol{P}(\boldsymbol{k})$ and even beyond the flat-sky approximation: it express indeed the heuristic properties of a gradient-like or curl-like spin-2 field.

Statistical characterisation of a field on the sphere

A field X on the sphere can be characterised in terms of the two point correlation function.

$$C(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}') \equiv \langle X^*(\hat{\boldsymbol{x}})X(\hat{\boldsymbol{x}}') \rangle \tag{3.33}$$

where the angle brackets denote the average over the ensemble of possible realisation of the field X. If we assume statistical isotropy $C(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}')$ becomes a function of only $\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{x}}'$ and it can be expanded on the Legendre polynomial

$$C(\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{x}}') = \frac{1}{4\pi} \sum_{\ell=0}^{\infty} (2\ell + 1) C_\ell P_\ell(\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{x}}'), \tag{3.34}$$

where C_ℓ is called *angular power spectrum*, though often it simply referred to as power spectrum, for brevity. With this assumption, the RHS of Eq. (3.33) is equal to the one of Eq. (3.34), they can be multiplied by $Y_{\ell m}(\hat{\boldsymbol{x}})Y_{\ell'm'}^*(\hat{\boldsymbol{x}}')$ and integrated over $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{x}}'$, the equality reduces to[2]

$$\langle a_{\ell m}^* a_{\ell'm'} \rangle = \delta_{\ell\ell'}\delta_{mm'}C_\ell. \tag{3.35}$$

This implies that, given the spherical harmonic decomposition of X, the C_ℓ can be estimated with

$$\hat{C}_\ell = \sum_{m=-\ell}^{\ell} \frac{a_{\ell m}^* a_{\ell m}}{2\ell + 1} \tag{3.36}$$

In current observations the CMB anisotropies are compatible with Gaussian fluctuations, as predicted by inflation. Assuming this has two important implications. The first one is that the power spectrum

---

[2]Remember that the Legendre polynomials can be written in terms of the spherical harmonics:

$$P_\ell(\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{x}}') = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y_{\ell m}(\hat{\boldsymbol{x}}')Y_{\ell m}^*(\hat{\boldsymbol{x}})$$

(or, equivalently, the two point correlation function) fully describes the statistical properties of the field. Second, even assuming no uncertainty in the harmonic coefficients, we get the following variance on the estimator in Eq. (3.36)

$$\text{Var}(\hat{C}_\ell) = \frac{2}{2\ell + 1} C_\ell^2. \tag{3.37}$$

called *cosmic variance*. In the power spectrum estimation in Eq. (3.36) we try to evaluate the variance of the distribution of the $a_{\ell m}$s from $a_{\ell m}$s themselves. Since for each ℓ we have only $2\ell + 1$ different values of m, the measurement is affected by the sampling variance in Eq. (3.37). Approximately speaking, if only a fraction of the sky f_{sky} is available, the number of modes available is accordingly reduced and the cosmic variance is boosted by a factor $1/f_{\text{sky}}$.

All the considerations made in this sections can be generalised to the case involving two (potentially different) fields X and Y. If the two fields are isotropic, similarly to Eq. (3.35), the harmonic coefficients are related to the power spectrum through

$$\langle {}_Y a_{\ell m}^* \, {}_X a_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell^{XY}. \tag{3.38}$$

and it can be estimated with

$$\hat{C}_\ell^{XY} = \sum_{m=-\ell}^{\ell} \frac{{}_X a_{\ell m}^* \, {}_Y a_{\ell m}}{2\ell + 1} \tag{3.39}$$

As explained in the previous sections, the CMB radiation can be expressed in two scalar and one pseudo scalar field: T, E and B. Therefore, in order to fully characterise its 2-point correlation, we have to estimate the six spectra coming from the combinations TT, EE, BB, TE, TB and EB.

3.3 Inhomogeneities in the primordial plasma

In this section we outline the dynamics of the primordial plasma, focusing in particular on the photon fluid and its properties after recombination. The formalism is the linear theory of cosmological perturbations, applied to the baryon, dark matter and photon fluid (Peebles and Yu, 1970; Bond and Efstathiou, 1984; Bond and Efstathiou, 1987), the aim is to understand the link between the observed CMB angular power spectrum and the physics of the primordial universe. This field has been very active between the 80s and the 90s. The discussion presented here is based on Hu and Sugiyama (1995), Hu et al. (1996), Seljak and Zaldarriaga (1996) and Dodelson (2003).

3.3.1 Perturbations of the photon fluid

The inhomogeneities of a fluid and their evolution are studied in terms of the time dependence of its distribution function, which is governed by the *Boltzmann equation*

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{x}, \boldsymbol{p}, t) = C[f], \tag{3.40}$$

where C is the possible interaction term with other fluids. What we are mostly interested in is the time evolution of the departure δf from the background distribution $\bar{f}$ (the bar denotes the solution if the universe was homogeneous).

For dark matter, δf is usually expressed in terms density fluctuation $\delta(\boldsymbol{x}, t) \equiv (\rho(\boldsymbol{x}, t) - \bar{\rho}(t))/\bar{\rho}(t)$ and the velocity $\boldsymbol{v}$ defined as the average of $\boldsymbol{p}/E$ over the momenta of the background distribution. The density and velocity of baryons, δ_{b} and v_{b}, are defined in the same way.

For photons, δf is expressed in terms of *fractional temperature fluctuations* $\Theta(\boldsymbol{x}, \hat{\boldsymbol{p}}, t)$ around the background black body homogeneous distribution

$$f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}, t) \equiv \frac{1}{\exp\{\frac{p}{T(t)[1+\Theta]}\} - 1} \tag{3.41}$$

$$= \bar{f}(t) - p \frac{\partial \bar{\mathrm{f}}}{\partial \mathrm{p}}(t) \Theta(\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}, t) + \mathcal{O}(\Theta^2). \tag{3.42}$$

As in Sec. 2.4, the study of linear perturbations is more natural in Fourier space. Therefore, henceforth we adopt the convention (standard in the literature) of denoting simply as Θ, δ, v, etc. the generic $\boldsymbol{k}$

Fourier mode $\tilde{\Theta}(\boldsymbol{k}, \hat{\boldsymbol{p}}, t)$, $\tilde{\delta}(\boldsymbol{k}, t)$, $\tilde{v}(\boldsymbol{k}, t)$, etc. Assuming that perturbations are axisymmetric around $\boldsymbol{k}$, the angular dependence of Θ is usually expressed as a function on $\mu \equiv \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{p}}$ and further decomposed on Legendre polynomials

$$\Theta = \Theta(\boldsymbol{k}, \mu, t) = \sum_{\ell=0}^{\infty} (-i)^{\ell} (2\ell + 1)\Theta_{\ell}(\boldsymbol{k}, t) P_{\ell}(\mu). \tag{3.43}$$

If Θ is isotropic, only the monopole Θ_0 is different from zero.

Finally, the evolution of the perturbation is usually studied in conformal time, hence the overdot will denote derivative with respect to the conformal time, for consistency with the literature and in contrast with previous chapters.

3.3.2   The dynamics of the plasma

Since in the primordial plasma photons and baryons are strongly interacting, they behave as a single fluid. The dynamics of this fluid is affected by the inhomogeneous gravitational potential, to which dark matter contributes. Gravity attracts the photon-baryon fluid in the potential well, photon pressure increases, eventually inverting the infall: the results is an oscillatory dynamics that is at the origin of the CMB anisotropies and we shall outline in this paragraph.

The Boltzmann equations for baryons are the following

$$\dot{\delta}_{\mathrm{b}} + ikv_{\mathrm{b}} = -3\dot{\Phi} \tag{3.44}$$

$$\dot{v}_{\mathrm{b}} + \frac{\dot{a}}{a}v_{\mathrm{b}} = -ik\Psi + \frac{\dot{\kappa}}{R}[v_{\mathrm{b}} + 3i\Theta_1]. \tag{3.45}$$

Eq. (3.44) is just the evolution of matter perturbation in a perturbed metric. On the contrary, Eq. (3.45) also contains in its last them the interaction with photons through Coulomb scattering, which tends to enforce isotropy in the rest frame of the electrons, $v_{\mathrm{b}} = -3i\Theta_1$. How efficiently the interaction can establish this condition depends on the frequency of Compton scattering and the inertia of the baryons. The former is represented by the *differential optical depth* $\dot{\kappa} = n_e \sigma_{\mathrm{T}} a \chi_e$, where χ_e is the ionization fraction, while the baryon inertia is expressed by the inverse proportionality on $R \equiv 3\bar{\rho}_b/4\bar{\rho}_\gamma$, the baryon to photon density normalized at $3/4$ at photon/baryon equality.

Neglecting polarization, the Boltzmann equation of photons reads

$$\dot{\Theta} + ik\mu\Theta = -\dot{\Phi} - ik\mu\Psi + \dot{\kappa}[\Theta_0 - \Theta + \mu v_{\mathrm{b}}]. \tag{3.46}$$

In this case the interaction with electrons, proportional to the differential optical depth, tends to isotropise the radiation in the rest frame of the electrons.

In the rest of this section, we discuss the dynamics of the fluids well before recombination, when the differential optical depth is very large. In this *tight coupling regime* $\Theta_1 \simeq -3v_{\mathrm{b}}$ and higher multiples are negligible because Compton scattering isotropises the photon distribution. Using this condition, we integrate Eq. (3.46) both before and after multiplying it by μ. We obtain two coupled first order differential equations in Θ_0 and Θ_1 that can be combined in the following famous second order differential equation

$$\ddot{\Theta}_0 + \frac{\dot{a}}{a}\frac{R}{1+R}\dot{\Theta}_0 + k^2 c_s^2 \Theta_0 = -\frac{k^2}{3}\Psi + \frac{\dot{a}}{a}\frac{R}{1+R}\dot{\Phi} - \ddot{\Phi} \tag{3.47}$$

which describes the driven oscillatory dynamics of the monopole of the temperature fluctuations. $c_s \equiv 1/\sqrt{3(1+R)}$ is the speed of sound in the photon-baryon fluid, which slowly decrease with the expansion of the universe since $R \propto a$. The maximum distance sound could travel by time τ is called *sound horizon* and is given by $s \equiv \int_0^\tau c_s \mathrm{d}\tau'$. In the rest of this section, we outline the oscillatory dynamics of the fluid in a schematic fashion.

Oscillatory dynamics

To start with, if there were no baryons and the potentials were constant the solution of Eq. (3.47) would be $\Theta + \Psi = \frac{1}{3}\Psi \cos(ks)$, where the amplitude of the harmonic function and the absence of the $\sin(ks)$ term comes from imposing deep in the radiation dominated era that the initial conditions match the adiabatic perturbations produced by inflation.

When at time τ_* recombination takes place, photons decouple from baryons and start free streaming. Since they have to climb (descend) the potential Ψ, they lose (gain) energy and photons appear colder (hotter) then they were at recombination. This is the so-called *Sachs-Wolfe effect*. $\Theta + \Psi$ is the effective temperature visible to an observer outside of the potential well that caused the oscillation.

In this first crude approximation, the effective temperature undergoes an oscillatory dynamics with a phase, ks, that grows with the sound horizon s until recombination.

At recombination, super-horizon modes ($ks_* \ll 1$) did not evolve significantly and they still reflect the primordial perturbations, $\Theta + \Psi = \frac{1}{3}\Psi$. Assuming $n_s \simeq 1$, as predicted by inflation, we have $\mathcal{P} \propto k^{-3}$ which, as we will see in the next section, produces an angular power spectrum $C_\ell \propto \ell^{-2}$. Super-horizon modes produce $\ell(\ell+1)C_\ell \sim$ const, which is called *Sachs-Wolfe plateau*.

Modes such that $ks_* \geq 1$ had time to experience the oscillatory dynamics. If $ks_* = \pi$ the mode has just reached its first compression and produce a peak in the power spectrum. In general, peaks are produced wherever at recombination the oscillations was at an extreme, i.e. for k multiple of π/s_*.

The "velocity" of the effective temperature also oscillates with the same period but $\pi/2$ out of phase: it has extrema when the effective temperature equals to zero.

Adding baryons

We now add baryons but suppose that $R = $ const. The equation of the oscillatory dynamics becomes $\Theta + \Psi = \frac{1}{3}\Psi(1+3R)\cos(ks) - R\Psi$.

The mass of the baryons shifts the zero point of the oscillation enhancing the compression of the fluid. The consequence is that the absolute amplitude of the compressional peaks is $2R\Psi$ higher than the rarefaction peaks. The mass of the baryons also decreases the speed of sound thus reducing the size of the sound horizon at a given time. Finally, R actually slowly evolve with time, since it is proportional to the scale factor. It can be shown that the consequence is a small decay with time proportional to $(1+R)^{-1/4}$.

Diffusion dumping

So far we have assumed to be in the tight coupling regime. In reality photons have a finite mean free path given by $\lambda_{\rm T} = 1/n_e\sigma_{\rm T}$ before they scatter off an electron. In a Hubble time there are on average $N \sim \sigma_{\rm T}n_e/H$ such scatters, each randomizing the photon direction. Consequently, in a Hubble time the mean distance travelled by a photon is $\lambda_{\rm D} \sim \lambda_{\rm T}\sqrt{N} \sim 1/\sqrt{n_e\sigma_{\rm T}H}$. For $k \gtrsim k_{\rm D} \equiv 2\pi/\lambda_{\rm D}$ photons and electrons can't be considered as a single fluid: hot and cold regions mix, effectively cancelling out fluctuations. The result is an exponential suppression factor given by $e^{-k/k_{\rm D}}$. Notice that $\lambda_{\rm D}$ evolves with time and goes to infinity at recombination. By definition most of the photons last scatter before this happens, the scale actually important is the "thickness" of the last scattering surface, the average diffusion length at last scattering.

3.4 From inhomogeneities to anisotropies

In the previous section we have studied the evolution of the Fourier modes of the perturbations until the last scattering. In this section we illustrate how they translate into anisotropies.

One way of doing this is to perform the Legendre expansion of Eq. (3.46) to get an infinite number of coupled first order differential equations, one for each $\ell = 0, \ldots, \infty$. We can then let this infinite differential system evolve from the primordial perturbations deep inside the radiation dominated era until the present time. As explained in the previous section, as long as the plasma is in the tight coupling limit only Θ_0 and Θ_1 are not suppressed. As the differential optical depth decreases, eventually reaching zero after recombination, Θ_l with $\ell > 1$ are progressively excited. This was the approach adopted by early calculations (e.g. Stompor, 1994; Sugiyama, 1995).

Another way of computing anisotropies from primordial perturbations is to start by writing the formal solution of Eq. (3.46)

$$(\Theta + \Psi)(\tau_0, \mu) = \int_0^{\tau_0} [\dot{\kappa}(\Theta_0 + \Psi - i\mu v_{\rm b}) - \dot{\Phi} + \dot{\Psi}]e^{\kappa(\tau,\tau_0)}e^{ik\mu(\tau-\tau_0)}{\rm d}\tau. \tag{3.48}$$

The optical depth κ is computed from τ to today, τ_0 and the combination $g(\tau) \equiv \dot{\kappa}e^{-\kappa}$ is the *visibility function*, which gives the probability that an observed photon was last scattered at time τ, it has a sharp

peak at recombination τ_*. Taking the multiple moments of Eq. (3.48) we get the following approximate solution

$$
\begin{aligned}
\Theta_\ell(k, \tau_0) \simeq & [\Theta_0(k, \tau_0) + \Psi(k, \tau_0)] j_\ell[k(\tau_0 - \tau_*)] \\
& + 3\Theta_1(k, \tau_0) \left(j_{\ell-1}[k(\tau_0 - \tau_*)] - \frac{(l+1) \, j_\ell[k(\tau_0 - \tau_*)]}{j_\ell[k(\tau_0 - \tau_*)]} \right) \\
& + \int_0^{\tau_0} \mathrm{d}\tau \, e^{-\kappa} \left[\dot{\Psi}(k, \tau_0) - \dot{\Phi}(k, \tau_0) \right] j_\ell[k(\tau_0 - \tau)].
\end{aligned}
\tag{3.49}
$$

from which one can derive the angular power spectrum today from the initial power spectrum $\mathcal{P}_\Psi$

$$
C_\ell = (4\pi)^2 \int \mathrm{d}k k^2 \mathcal{P}_\Psi(k) |\Theta_\ell(k, \tau_0)|^2.
\tag{3.50}
$$

$j_\ell(x)$ are the Bessel functions. They express how a Fourier mode $\boldsymbol{k}$ contributes to a multipole ℓ: their peak at $x \sim \ell$ means that the $\boldsymbol{k}$ Fourier mode contributes predominantly to the $\ell = k(\tau_0 - \tau_*)$ multipole. However, they are not delta functions and consequently the angular power spectrum appears smoother than the power spectrum of the Fourier modes at the last scattering.

The most important term in Eq. (3.49) is the first row. It is the contribution from the oscillations of the effective temperature to the present multipoles. The second row represents the Doppler effect due to the line of sight velocity of the photon-baryon fluid at the moment of recombination. Its more complex geometrical term smooths the structure of the peaks compared to the one of the effective temperature. Since the velocity field is $\pi/2$ out of phase with the density field, at the power spectrum level the crest of the former corresponds to the trough of the latter. Consequently, the velocity field adds power on the top of the power spectrum of the density field but smooths its peaks.

The third raw in Eq. (3.49) is the *integrated Sachs-Wolfe*. It tells us that anisotropies can be sourced if, after the last scattering, the potential through which the photon is travelling evolves (Sachs and Wolfe, 1967). The contribution to this term is typically divided into two terms the *early ISW*, the *late ISW*. Moreover, the ISW has a non-linear analogous: the *Rees-Sciama effect*. Since they are produced after the last scattering we include them in the secondary anisotropies.

3.5 Secondary anisotropies

We discussed the physics of the primordial universe and how it leaves imprints in the CMB power spectrum at the last scattering surface. However, also the intervening universe produces anisotropies in the CMB, in this section we discuss very briefly the most relevant effects. For a review see Aghanim et al. (2008).

Integrated Sachs-Wolfe effect (Sachs and Wolfe, 1967). When a photon enters and then leaves a static gravitational potential well, it is first blue-shifted and then redshifted, with no net effect. However, if the potential evolves during the passage of the photon, the energy balance can be non-zero (see e.g. Mukhanov et al., 1992). This happens at early times, close to recombination, because the radiation energy density is not dynamically negligible (early ISW), and it also happens at late times (late ISW), when large-scale potentials decay because of the accelerated expansion during the dark energy dominated era.

Rees-Schiama effect. The ISW is a first order term, on small spatial scales these terms cancel out and thus second order effect can become relevant. In hierarchical structure formation the collapse of structure can produce an effect analogous to the ISW if the photon passing time is a relevant fraction of the collapse time-scale (Rees and Sciama, 1968; Dyer, 1976).

Gravitational lensing. The path of the CMB photons is deflected by the gravitational attraction exerted by intervening cosmic structures. Given the importance of this effect in the context of B-modes, and POLARBEAR in particular, we devote a dedicated section to the topic (see Sec. 3.8).

Sunyaev-Zeldovich effect. The optical depth of the ICM of galaxy clusters is typically 10^{-3}. Because of the high temperature of the ICM—typically 5-10 keV—the scattering of the photons injects energy in the CMB that cannot be thermalised. As discussed in Sec. 3.1 this produces a y-type distortion, known as thermal Sunyaev-Zeldovich (tSZ) effect (Sunyaev and Zeldovich, 1972, 1980). This signal has a singular property: its amplitude—the Compton parameter y, effectively the ICM pressure integrated along the line of sight—depend only on the cluster properties and not on its redshift, which gives to the SZ an edge on the X-ray on the detection of far clusters.

The development in the observation of large-area surveys by the ACT (Hincks et al., 2010; Marriage et al., 2011; Hasselfield et al., 2013), SPT (Staniszewski et al., 2009; Vanderlinde et al., 2010; Williamson et al., 2011; Reichardt et al., 2013) and *Planck* (Planck Collaboration, 2014c, 2015h,i) telescopes dramatically improved the quality and the number of the SZ cluster detections, allowing the extraction of cosmological information from the statistical properties of the SZ-cluster samples.

The abundance of clusters and its evolution with redshift are indeed sensitive to both the total matter density and the amplitude of the power spectrum of density fluctuations. SZ-cluster counts recently become a powerful probe of σ_8 and Ω_{m} (Planck Collaboration, 2014a, 2015e). In particular, these studies highlighted a tension with the determination of the same quantities from primary CMB anisotropies. This discrepancy is particularly interesting because the two measurements probe completely different cosmological epochs, and thus might imply that the current understanding of the evolution of the universe has to be improved.

The tSZ effect also provided other cosmological constraints, for example, on the redshift dependence of the CMB temperature (Hurier et al., 2014), on angular distance-redshift relation (Bonamente et al., 2006) and the distance duality relation (Uzan et al., 2004).

The tSZ effect has also emerged as a powerful probe of the physics of the ICM, especially combined with the traditional X-ray measurements. The two signals are indeed sourced by the same thermal gas of electrons but have a different dependence of the electron density: they are proportional to n_e and n_e^2 respectively. The tSZ signal enables to better characterise the outskirts of the clusters (e.g. Planck Collaboration et al., 2013), where the X-ray emission becomes too faint. The focus on the outskirts is also driven by the resolution of microwave-surveys that typically have arcmin-resolution—which is comparable with the angular size of the clusters and much larger than the arcsec-resolution of X-ray measurements. However, individual clusters have been observed at high resolution, e.g. by the MUSTANG (Mroczkowski et al., 2011) and NIKA experiments (Adam et al., 2015; Adam et al., 2015, 2016).

Until now we have discussed only the thermal motion of the electrons as a source of the SZ effect. However, it can be also sourced by the bulk motion of the ICM, the so-called kinematic SZ effect (kSZ). Detections of this effect have been reported either statistically (Hand et al., 2012) or from high-velocity mergers (Sayers et al., 2013).

More information about the SZ effect can be found in the review articles by Birkinshaw (1999), Carlstrom et al. (2002) and Kitayama (2014).

Reionisation. After recombination the universe is basically only composed by neutral hydrogen. About 400 Myr after, the first generations of galaxies start to form, setting the end of the so-called *dark ages*. The newly born astrophysical objects emit UV-radiation, which is capable of ionising the neutral hydrogen. Therefore, the first generation of stars and black hole driven sources (e.g. mini-quasars, X-ray binaries, etc.) starts to carve out ionised regions around them. The speed of this process increases with the number of sources and proceeds until the whole universe is ionised again. This process is called *reionisation* and is thought to take place within the redshift range $z \sim [15, 6]$.

Reionisation can be probed using the Lyman-α forest and future observations of the redshifted 21-cm emission line from neutral hydrogen will provide direct and detailed evidence of this process. In addition, reionisation affects also the CMB power spectrum in three ways. First, the re-scattering of CMB photons on free electrons causes the suppression of the angular anisotropies on all scales because it blends photons coming from initial different directions. Second, reionisation adds some small-scales power due to the inhomogeneous baryon density (*Ostriker-Vishniac* effect) and the patchy expected nature of reionisation. Third, the bulk flows of electrons at recombination create a quadrupole anisotropy that, as discussed later, produces a bump in the polarization power spectrum at very large scales, $\ell \lesssim 10$. Despite the challenges related to measuring such scales, this signal was detected (Kogut et al., 2003). The quantities that parametrise the impact of reionisation on the CMB power spectrum—the Thompson optical depth τ and the reionisation redshift—are difficult to constrain accurately because of their degeneracy with other parameters. Nevertheless, CMB observations start to provide relevant constraints on the recombination process (see Planck Collaboration, 2016a).
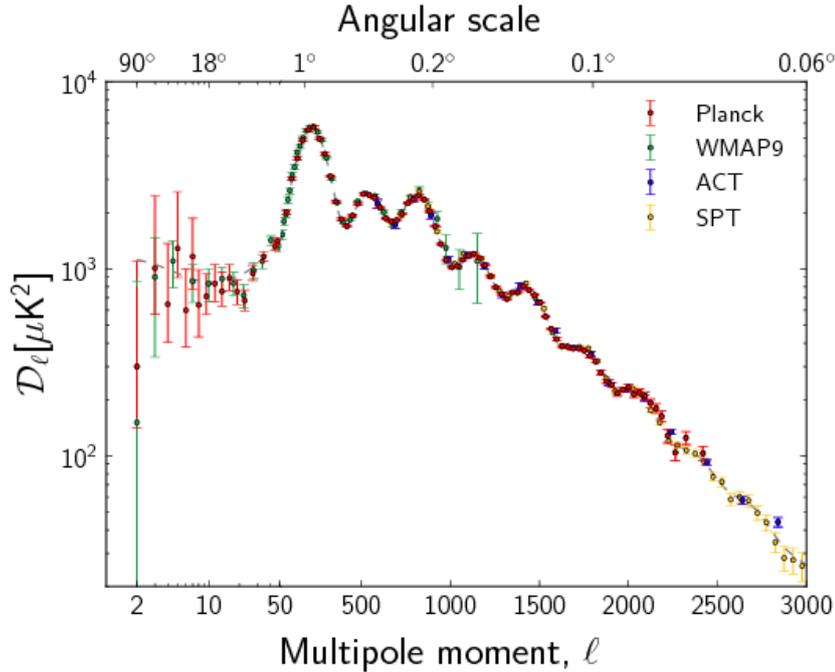
Figure 3.2: The TT spectrum of the *Planck* (Planck Collaboration, 2014b), ACT (Das et al., 2014) and SPT (George et al., 2015) experiments. The consistency with the WMAP9 measurement (Bennett et al., 2013) as well agreement between the other three experiments shows a sound control of the systematic effects. (Planck Collaboration, 2014b)

3.6   Constrains on the temperature power spectrum

Since the first detection of the CMB anisotropies by the COBE-DMR experiment (Smoot et al., 1992), the accuracy in their characterisation has been steadily increasing. The determination of the location of the first peak was a major landmark, after a tentative measurement by Torbet et al. (1999) the peak was detected by the BOOMERANG (Mauskopf et al., 2000; Melchiorri et al., 2000) and MAXIMA (Hanany et al., 2000; Stompor et al., 2001a) balloon experiments. These measurement constrained the flat geometry of the universe, indirectly confirming the dark energy fraction of the present universe. The second and third peak were detected by the WMAP satellite (Hinshaw et al., 2007). The most accurate measurement to date come from the *Planck* satellite (Planck Collaboration, 2015g) combined with the ACT (Das et al., 2014) and SPT (George et al., 2015) ground-based experiments, see Fig. 3.2. The former probes $\ell \leq 2500$ over the whole sky and the latter are sensitive to smaller scales, reaching $\ell = 11000$.

3.7   Including polarization and tensor perturbations

In this section we move our attention to polarization. Following Kosowsky (1996), we first illustrate how Compton scattering can produce polarization when the radiation being scattered off is unpolarised but not isotropic around the electron, more precisely when it has non-vanishing quadrupole moment. Then, we illustrate how the scalar perturbation underlying the temperature power spectrum also source polarization. We show that another possible source are gravitational waves and, unlike scalar perturbations, the polarization produced has a *B*-type component. This latter section is based on Hu et al. (1996).

3.7.1   Compton scattering: polarization from unpolarised light

We consider for simplicity an observation in the $-\hat{z}$ direction and suppose that the observed radiation all comes from a single electron that last scattered the CMB, supposed at the centre of the coordinate
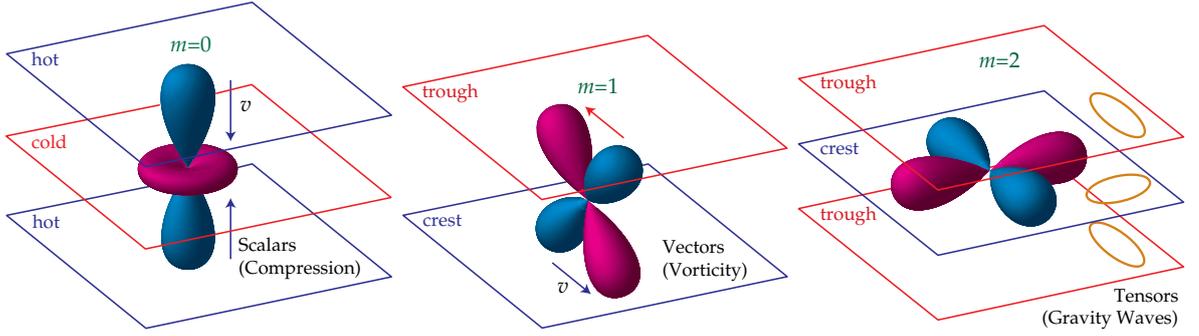
Figure 3.3: The quadrupole moments produce by scalar, vector and tensor perturbations. Assuming the $\hat{z}$ axis in the direction of the wave vector they correspond to $m = 0, \pm 1, \pm 2$ quadrupole moments respectively.

system. For a single photon, the Compton cross section is given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{3\sigma_\mathrm{T}}{8\pi} |\hat{\varepsilon}_\mathrm{in} \cdot \hat{\varepsilon}_\mathrm{out}|^2. \tag{3.51}$$

$\hat{\varepsilon}$ is the polarization versor of the radiation, orthogonal to the direction of propagation of the photon $\hat{n}$. $\hat{n}_\mathrm{out}$ is $\hat{z}$ and we chose $\hat{x}$ and $\hat{y}$ as a basis for $\hat{\varepsilon}_\mathrm{out}$. For the photon incoming from the (θ, ϕ) direction in spherical coordinates we have

$$\hat{n} = (\sin\theta \, \cos\phi, \sin\theta \, \sin\phi, \cos\theta) \tag{3.52}$$

and for $\hat{\varepsilon}_\mathrm{in}$ we choose the basis

$$\hat{\varepsilon}_1 = (\cos\theta \, \cos\phi, \cos\theta \, \sin\phi, -\sin\theta) \tag{3.53}$$
$$\hat{\varepsilon}_2 = (-\sin\phi, \cos\phi, 0). \tag{3.54}$$

Since we are assuming that the incoming radiation is unpolarised, the incoming intensity $I(\hat{n}_\mathrm{in})$ along $\hat{\varepsilon}_1$ and $\hat{\varepsilon}_2$ is on average the same (in the sense used in the definition of the Stokes parameters). We allow however this intensity to be depend on the direction $\hat{n}_\mathrm{in}$. Owing to Eq. (3.51), the contribution from the $\hat{n}_\mathrm{in}$ direction to I_x, the outgoing intensity on the $\hat{x}$ axis, is proportional to $I(\hat{n}_\mathrm{in})[|\hat{x} \cdot \hat{\varepsilon}_1|^2 + |\hat{x} \cdot \hat{\varepsilon}_1|^2]$ and a similar relation holds for I_y. By averaging over all the possible incoming directions, we get the total observed Stokes parameters

$$Q(\hat{n}_\mathrm{out}) = A \int \mathrm{d}\Omega \, I(\hat{n}_\mathrm{in}) \sin^2\theta \cos(2\phi)$$

$$\propto A \int \mathrm{d}\Omega \, I(\hat{n}_\mathrm{in})(Y_{2\,2} + Y_{2\,-2}) \tag{3.55}$$

$$U(\hat{n}_\mathrm{out}) = -A \int \mathrm{d}\Omega \, I(\hat{n}_\mathrm{in}) \sin^2\theta \sin(2\phi)$$

$$\propto A \int \mathrm{d}\Omega \, I(\hat{n}_\mathrm{in})(Y_{2\,-2} - Y_{2\,2}) \tag{3.56}$$

from which we learn that polarization is produced from unpolarised radiation when (and only when) the latter has non zero quadrupole moment ($\ell = 2$) around the scattering centre. Notice also that not all the quadrupole moments produce polarization, just the one with $m = \pm 2$ (loosely speaking, the ones with the anisotropies on the plane orthogonal to the line of sight $\hat{z}$).

3.7.2   Sources of polarization

In this section we qualitatively explain that scalar, vector and tensor perturbations source polarization and comment on the salient differences in the polarization patterns that they produce. We consider a single plane wave at the scattering surface with wave vector $\boldsymbol{k}$ parallel to the $\boldsymbol{z}$ axis.

If the perturbation is of scalar type, it causes bulk flows from hot to cold (effective) temperature regions. For scalar perturbation equal phase planes are equal temperature, so the flow is along $\boldsymbol{k}$ (orthogonal to equal phase planes). An electron that at last scattering is in a trough, a cold region, sees hot

radiation coming from $\hat{\boldsymbol{k}}$ or $-\hat{\boldsymbol{k}}$ and cold radiation coming from the orthogonal directions. It therefore observes a quadrupolar component of the radiation with $m = 0$, see Fig. 3.3. In positions different from the trough the $m = 0$ nature of the quadrupole moment is preserved, it is just modulated by the same phase of the oscillatory dynamics.

The vector modes source vorticity in the photon-baryon plasma. As a consequence the velocity field always satisfies $\nabla \cdot \boldsymbol{v} = 0$. For a single Fourier mode $\boldsymbol{k}$ it implies that the velocity is always orthogonal to $\boldsymbol{k}$. Equal phase planes glide on each other at different velocity, depending on their phase in the oscillatory dynamics. Because of this velocity, an electron at last scattering sees a dipole due to the velocity in its equal phase plane. Moreover, it sees a quadrupole moment due to the relative velocity between the equal phase planes "above" and "below" its own. The quadrupolar anisotropy sourced is of type $m = \pm 1$, see Fig. 3.3.

Tensor modes can be seen as gravitational waves. If we consider the equal phase plane of a Fourier mode $\boldsymbol{k}$, the space is periodically stretched in a direction and squeezed in the orthogonal direction, sourcing quadrupole anisotropy lying on the plane orthogonal to the $\boldsymbol{k}$ vector, thus of type $m = \pm 2$.

Summarising, the Fourier mode $\boldsymbol{k}$ of scalar, vector and tensor perturbations produce quadrupole anisotropies respectively with $m = 0, \pm 1, \pm 2$ at the last scattering surface. In the previous paragraph we said that, assuming the observation to take place in the $\hat{\boldsymbol{z}}$ direction, only the $m = \pm 2$ quadrupole produce polarization. So at the poles only tensor perturbations produce polarization. For any other direction, the amount of polarization produced by scalar, vector and tensor perturbations depends on how the $m = 0, \pm 1, \pm 2$ quadrupoles project on the $m = \pm 2$ quadrupoles assuming the $\hat{\boldsymbol{z}}$ axis in the observation direction. This geometrical dependence causes the different type of perturbation to produce different amounts of polarization.

The polarization pattern of the scalar case has a notable feature. Since the $m = 0$ quadrupole is symmetric around the z axis, the polarization pattern produced depend on θ but not ϕ. For the same reason, the polarization vector $\boldsymbol{P}$ can be oriented only either in the $\hat{\boldsymbol{e}}_\theta$ direction or in the $\hat{\boldsymbol{e}}_\phi$ direction. Summarising the polarization vector is either parallel or orthogonal to the gradient of its magnitude, implying that scalar perturbations produce a pure E-type polarization.

Tensor and vector perturbations, on the contrary, produce both E and B modes. One can show that, at small angular scales, the B/E ratio, equal to 0 for scalar perturbations, is 8/13 for tensor perturbations and 6 for vector perturbations.

Fig. 3.4 compares the expected power spectra produced by the scalar and tensor perturbations (the tensor-to-scalar ratio assumed, $r = 0.24$ is actually excluded to high significance nowadays). The scalar contribution dominates over the tensor one. Consequently, the cosmic variance of the former is an obstacle to the measurement of the latter. The only exception is the B-mode polarization, where the primordial scalar perturbations are expected to have no signal. As we will show in Sec. 3.8, primordial E modes are converted into B modes due to gravitational lensing. The lensing B signal dominates at small scales. Therefore, the large scales of the BB power spectrum is where the signal of the primordial gravitational waves is expected to be more visible.

3.7.3   Constraints on polarization power spectra

The first detection of CMB polarization was published in 2002 by the DASI experiment, which made statistical detection of both EE and TE.

TE and EE has been characterised accurately with a series of subsequent experiments: BICEP (Chiang et al., 2010), BOOMERANG (Piacentini et al.), CAPMAP (Bischoff et al., 2008), CBI (Sievers et al., 2005), DASI (Leitch et al., 2005), QUAD (Brown et al., 2009), WMAP (Bennett et al., 2013) and QUIET (Bischoff et al., 2011). The most stringent constraints on large and intermediate scales are set by Planck Collaboration (2015d), see Fig. 3.5, while the small scales are constrained to high accuracy by ACTPol (Naess et al., 2014a) and SPTPol (Crites et al., 2015).

The lensing B-modes were tentatively detected by POLARBEAR Collaboration (2014a). Soon after, BB power was detected to high significance by BICEP2 Collaboration (2014) at large scales. However, the contribution from galactic dust could not be reliably estimated and removed. This was performed in a BICEP2-*Planck* joint analysis (The BICEP2/*Keck Array* and *Planck* Collaborations, 2015), in which lensing B-modes were detected at 7σ and the most stringent constraints on the tensor-to-scalar ratio were set, $r < 0.07$ at 95% confidence level. ACTPol (Naess et al., 2014a) and SPTPol (Keisler et al., 2015) set further constraints at small scales and The BICEP2/*Keck Array* Collaboration (2016) improved their constraints, setting the most stringent upper bound on the tensor-to-scalar ratio, $r < 0.07$ at 95% confidence level, when combined with *Planck* and BAO measurements.
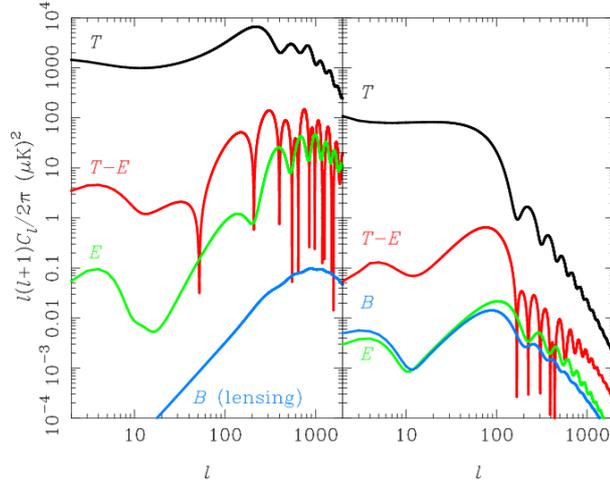
Figure 3.4: The power spectrum of scalar perturbations compared to the one of the tensor perturbations assuming $r = 0.24$. Large-scale B modes are the only probe for the primordial gravitational wave which is not dominated by the signal produced by the scalar perturbations. (Challinor, 2013)

3.8 Lensing

Gravity can bend the trajectory of light rays. While in the Newtonian theory this was obtained by supposing that the photons had an infinitesimal mass, in general relativity gravity bends the space-time itself and, consequently, the null geodesics followed by the light are not "straight". Gravitational structures act like lenses and the distortion of a background source due to an intervening structure is called *gravitational lensing*. This section describes this effect in the context of the CMB, see Hanson et al. (2010) for a concise and Lewis and Challinor (2006) for a comprehensive review.

The lensing of some source light allows probing the intervening structures. Many modern and forth-coming experiments probe the evolution of cosmic structures by studying the distortion of the ellipticity shape of distant galaxies. In the context of lensing, also the CMB is an invaluable probe. This backlight has indeed the unique properties of being extended across the full sky and being the furthest observable light, potentially enabling to probe the structure formation of the whole observable universe. Moreover, since the CMB is linearly polarized, two more Stokes parameters can be exploited in addition to the total intensity of the radiation. Last but not least, theoretical prediction and observations of the CMB anisotropies are compatible with Gaussian fluctuations, therefore the non-Gaussian component can be exploited to constrain the lensing effect.

In order to have an idea of the effect of weak lensing on the CMB, we start with an order of magnitude analysis. We consider some representative property of the cosmological structure. The characteristic comoving scale for the inhomogeneities of the gravitational potential is about $300 \, \text{Mpc}$ and a typical value for their depth is $\psi \sim 2 \times 10^5$. A photon approaching such a structure would be deflected by an angle $\delta\beta \sim 2\psi \sim 10^{-4}$. Since the comoving distance of the last scattering surface is about $14\,000 \, \text{Mpc}$, there are approximately $14000/300 \sim 50$ such deflections and the total deviation is about $\sqrt{50}\delta\beta \sim 7 \times 10^{-4} \sim 2 \, \text{arcmin}$. The coherence of this shift is given by the angular size of the structure. Supposing the typical structure to be halfway between us and the last scattering surface, the coherence scale is about $300/7000 \text{rad} \sim 2°$. Therefore lensing consists in a $\sim$arcminute displacement coherent over $\sim$degree scales.

3.8.1 The deflection angle and lensing potential

In this section we give a heuristic demonstration of how the displacement between the observed and original picture is related to the gradient of the gravitational potential integrated along the line of sight.

Consider the trajectory of a photon in a perturbed FLRW metric, schematically represented in Fig. 3.6 in comoving coordinates. The infinitesimal deflection $\delta\beta$ of the trajectory of the photon is related to the Newtonian potential ϕ (actually the Weyl potential). It can be shown that, for an infinitesimal displacement $\delta\chi$, the infinitesimal deflection is

$$\delta\beta = -2\delta\chi \nabla_\perp \Psi \tag{3.57}$$
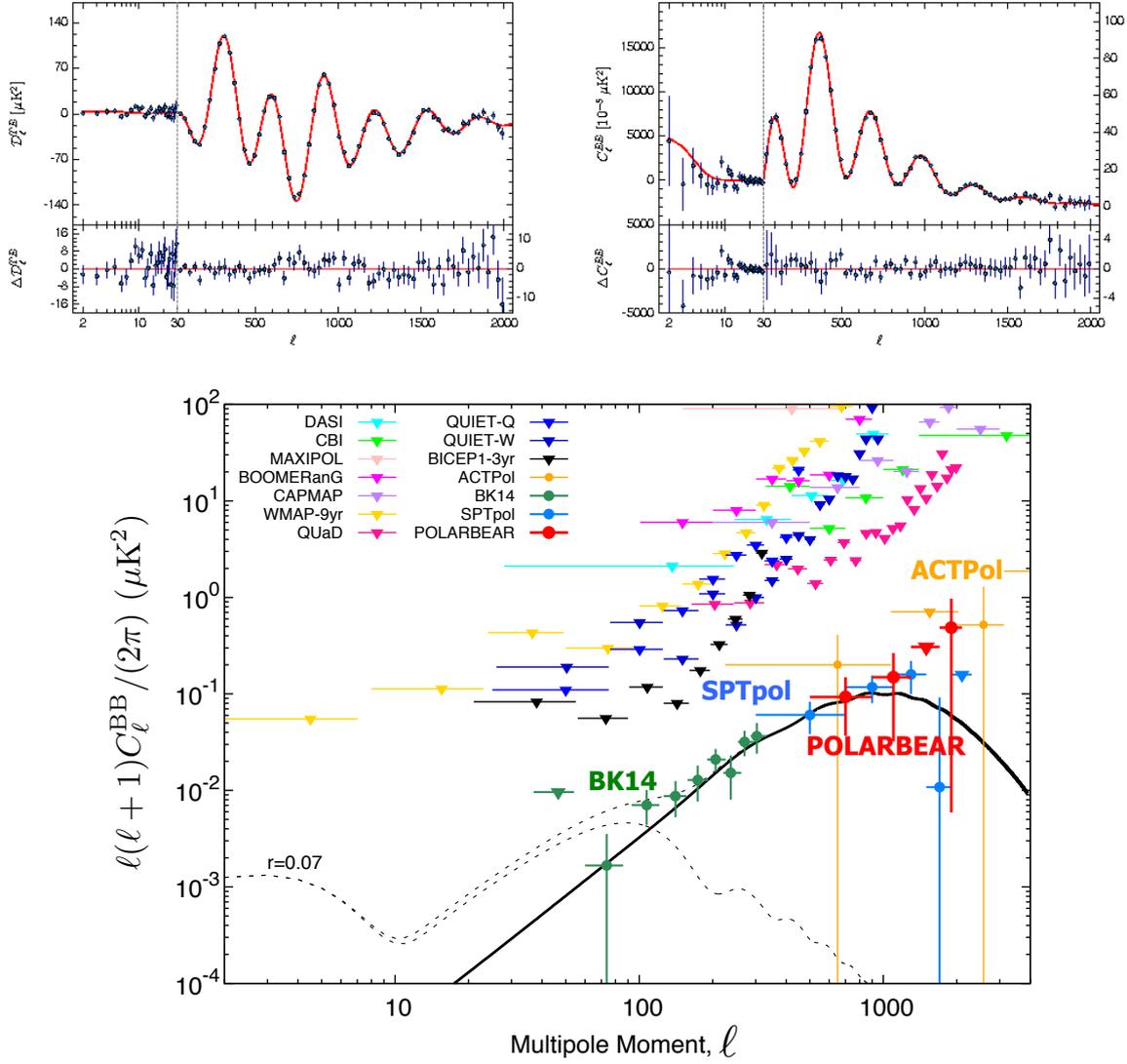
Figure 3.5: The top left and right panels show the TE and EE power spectrum from the *Planck* satellite ([Planck Collaboration, 2015d](#)). The bottom panel shows the current constraints on BB (Courtesy of Yuji Chinone).

where $\nabla_{\perp}\psi$ is the derivative of the potential in the direction orthogonal to the direction of movement of the photon. Since the $\delta\beta$ and $\delta\theta$ subtend the same comoving size, the following relation holds

$$S_k(\chi_* - \chi)\delta\beta = S_k(\chi_*)\delta\theta. \tag{3.58}$$

Thus the infinitesimal contribution to the deflection angle is

$$\delta\theta = -2\delta\chi \frac{S_k(\chi_* - \chi)}{S_k(\chi_*)}\nabla_{\perp}\Psi. \tag{3.59}$$

In principle, in order to reconstruct the total deflection angle, we should integrate all the infinitesimal deflections over the path of the photon. However, to first order, the quantities inside this path integral can be evaluated along the unperturbed path of the photon, this is called *Born approximation*. The gradient of the of the potential becomes orthogonal to the line of sight and it is thus related to the angular derivative (henceforth denoted simply by ∇) by $\nabla\Psi = \nabla_{\perp}\Psi S_k(\chi)$. The deflection field can be written as

$$\boldsymbol{\alpha}(\hat{\boldsymbol{x}}) = -2\int_0^{\chi_*} \mathrm{d}\chi \frac{S_k(\chi_* - \chi)}{S_k(\chi)S_k(\chi_*)}\nabla\Psi(\hat{\boldsymbol{x}}\chi, \tau_0 - \chi), \tag{3.60}$$
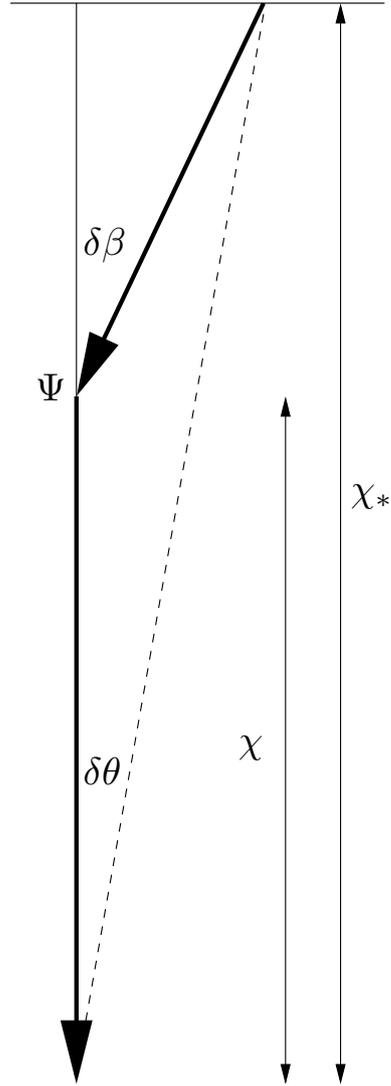
Figure 3.6: Geometry of a single deflection due to gravitational lensing. The plane at the top represents the last scattering surface while the vertical line is the line of sight. The deflection of the photon trajectory by $\delta\beta$ at a comoving distance χ causes a displacement between the observed and original direction of the photon of $\delta\theta$.

where the conformal time $\tau_0 - \chi$ is when the photon was at the comoving coordinates $\hat{\boldsymbol{x}}\chi$. The deflection field therefore can be written as

$$\boldsymbol{\alpha}(\hat{\boldsymbol{x}}) = \nabla\psi(\hat{\boldsymbol{x}}) \tag{3.61}$$

where we have defined the *lensing potential*

$$\psi(\hat{\boldsymbol{x}}) \equiv -2 \int_0^{\chi_*} \mathrm{d}\chi \frac{S_k(\chi_* - \chi)}{S_k(\chi)S_k(\chi_*)} \Psi(\hat{\boldsymbol{x}}\chi, \tau_0 - \chi), \tag{3.62}$$

which is nothing but the gravitational potential multiplied by a geometrical factor and integrated along the line of sight.
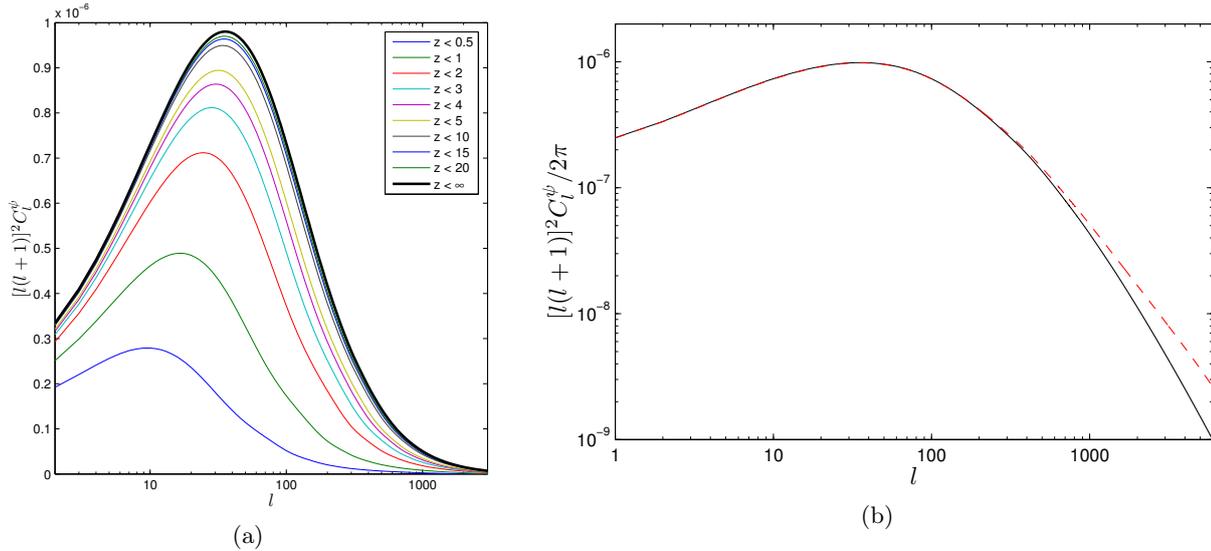
Figure 3.7: (a) The total lensing power spectrum (black) and cumulative contributions as function of the maximum redshift. While the galaxy surveys probe the lensing potential up to $z \sim 1$, the lensing potential to which the CMB is sensitive receive a substantial contribution from higher redshifts. (b) The angular power spectrum of the lensing potential from the linear theory compared to the one including the HALOFIT non-linear correction (Smith et al., 2003). (Lewis and Challinor, 2006)

3.8.2   Convergence and rotation fields

In this section we will work in the flat sky approximation and write the 3-dimensional versor $\hat{\boldsymbol{x}}$ as a 2-dimensional vector $\boldsymbol{x}$. The lensed CMB (denoted by a tilde) is related to the unlensed CMB by

$$\left(\begin{array}{c} \tilde{T}(\boldsymbol{x}) \\ \tilde{Q}(\boldsymbol{x}) \\ \tilde{U}(\boldsymbol{x}) \end{array} \right) = \left(\begin{array}{c} T(\boldsymbol{x} + \boldsymbol{\alpha}) \\ Q(\boldsymbol{x} + \boldsymbol{\alpha}) \\ U(\boldsymbol{x} + \boldsymbol{\alpha}) \end{array} \right) \tag{3.63}$$

where $\boldsymbol{\alpha}$, the *deflection angle*, is a 2-dimensional vector. Particularly important, the *magnification matrix*, related to the derivative of the deflection angle,

$$\boldsymbol{M} \equiv \frac{\partial(\boldsymbol{x} + \boldsymbol{\alpha})}{\partial \boldsymbol{x}} \tag{3.64}$$

is the Jacobian of the lensing remapping of the sphere. We always assume it to be invertible, this is the mathematical definition of the weak lensing regime. In practice, this is a necessary and sufficient condition to ensure that the remapping is a diffeomorphism, thus excluding the presence of caustics and multiple images of the same portion of the last scattering surface. These latter features define the *strong lensing* regime and can be seen only in the vicinity of galaxy clusters or compact objects.

The magnification matrix is usually parametrised as follows

$$\boldsymbol{M}(\boldsymbol{x}) = \left(\begin{array}{cc} 1 + \kappa + \gamma_Q & \gamma_U + \omega \\ \gamma_U - \omega & 1 + \kappa - \gamma_Q \end{array} \right)(\boldsymbol{x}). \tag{3.65}$$

The *convergence*, κ, magnifies the feature in the $\boldsymbol{x}$ direction by a factor $1 + \kappa$ while the rotation field, ω, will rotate it clock-wise by ω radians. The Q-shear stretch the feature in the x axis by $(1 + \gamma_Q)$ while squeezes it by $(1 - \gamma_Q)$ in the y axis. The U-shear does the same thing in the $x = y$ and $x = -y$ directions. The formalism has a complete analogy with the polarization field, with the shear being the equivalent of the Q and U Stokes parameters and the κ and ω being the analogous of the E and B modes respectively. The shear can be indeed derived from the convergence and rotation fields with Eq. (3.32).

As any vector field, $\boldsymbol{\alpha}$, can be decomposed into the gradient of a potential and the curl of the cross-potential,

$$\boldsymbol{\alpha}(\boldsymbol{x}) = \nabla \Psi(\boldsymbol{x}) + \star \nabla \Omega(\boldsymbol{x}), \tag{3.66}$$

where the $\star$ denotes the ninety-degree rotation operator. Propagating the decomposition through the magnification matrix we find

$$\kappa = -\frac{1}{2}\nabla^2\psi \tag{3.67}$$

$$\omega = -\frac{1}{2}\nabla^2\Omega. \tag{3.68}$$

Working with the potentials or with convergence and rotation fields is equivalent, the latter being equal to the former multiplied by $k^2/2$ in harmonic space.

The most important observation about the gradient-curl decomposition is that the rotation field is very small compared convergence field. Indeed, as it was shown in the previous section, in the Born approximation the rotation field is zero. Both recent post-Born semi-analytical work (Pratten and Lewis, 2016) and ray-tracing simulations (Calabrese and Fabbian, in preparation) confirm that the post-Born contribution to the rotation field is undetectable for any realistic experimental configuration. Even for the convergence field, the post-Born corrections don't give a relevant contribution to the power spectrum for any near future observation (Krause and Hirata, 2010; Calabrese et al., 2014; Pratten and Lewis, 2016) but they produce a relevant amount of non-Gaussianity that might be detectable within a decade (Pratten and Lewis, 2016).

The lensing potential ψ can be used to constrain the history of the gravitational potential Ψ, which is sensitive to both the primordial perturbations and their evolution. As long as the evolution is linear, the Fourier modes of the gravitational potential are given by

$$\Psi(\boldsymbol{k},\tau) = T_\Psi(k;\tau)\mathcal{R}_{\boldsymbol{k}} \tag{3.69}$$

where $\mathcal{R}$ is the primordial comoving curvature perturbation and T_Ψ is the transfer function that makes the Fourier modes evolve with the (conformal) time τ. The power spectrum of the lensing potential in Eq. (3.62) can be rewritten as (Lewis and Challinor, 2006)

$$C_\ell^{\psi\psi} = 16\pi \int \frac{\mathrm{d}k}{k}\mathcal{P}_\mathcal{R}(k)\left[\int_0^{\chi_*}\mathrm{d}\chi T_\Psi(k;\tau_0-\chi)j_\ell(k\chi)\left(\frac{\chi_*-\chi}{\chi_*\chi}\right)\right]^2 \tag{3.70}$$

where, for simplicity we have assumed a flat universe. This expression emphasises the three key elements that affect the lensing potential power spectrum: the primordial power spectrum of scalar fluctuations, $\mathcal{P}_\mathcal{R}$, its evolution through the transfer functions and the geometrical term. Because of these last two terms different distances contribute differently to the power spectrum (see Fig. 3.7a). In particular, the lensing potential sensitive to the late time non-linear evolution (see Fig. 3.7b). The non linear power spectrum is estimated through N-body simulations (Das and Bode, 2008; Fosalba et al., 2008; Carbone et al., 2009; Calabrese et al., 2014), though codes that compute theoretical CMB and matter power spectra (such as CAMB, Lewis and Challinor, 2006) use analytic fitting formulae (e.g. Smith et al., 2003), reaching per cent-level accuracy up to $\ell \simeq 2000$.

3.8.3 Reconstructing the lensing potential

In this section we illustrate the most popular technique to estimate the lensing potential from the lensed CMB sky (Hu and Okamoto, 2002). As in the previous section we adopt for simplicity the flat sky approximation but the full sky formalism (Okamoto and Hu, 2003) is very similar.

The lensing produces ∼arcmin deflections while CMB fluctuates on ∼degree-scale. Therefore, the following Taylor expansion is adequate

$$\tilde{X}(\boldsymbol{x}) = X(\boldsymbol{x}+\boldsymbol{\alpha}(\boldsymbol{x})) \tag{3.71}$$

$$= X(\boldsymbol{x}) + \boldsymbol{\alpha}(\boldsymbol{x})\cdot\nabla X(\boldsymbol{x}) + \mathcal{O}(\boldsymbol{\alpha}^2), \tag{3.72}$$

where X is any of the Stokes parameters. This equation expresses the lensed CMB $\tilde{X}$ as the unlensed CMB X plus a first order correction $\delta\tilde{X}$. This latter term involves the gradient of the of the lensing potential $\boldsymbol{\alpha} = \nabla\psi$ and the gradient of the Stokes parameter ∇X. Since it is a real space scalar product of these two functions, $\delta\tilde{X}$ can be equivalently expressed as a convolution of $\boldsymbol{l}\psi_{\boldsymbol{l}}$ and $\boldsymbol{l}'X_{\boldsymbol{l}'}$ in harmonic space

$$\delta\tilde{X}_{\boldsymbol{l}} = \int \frac{\mathrm{d}\boldsymbol{l}'}{(2\pi)^2}X_{\boldsymbol{l}'}\boldsymbol{l}'\cdot(\boldsymbol{l}-\boldsymbol{l}')\psi_{\boldsymbol{l}-\boldsymbol{l}'}. \tag{3.73}$$

It is insightful to express the polarization Stokes parameters in terms of E and B (see Eq. (3.32)),

$$\delta \tilde{E}_{\boldsymbol{l}} = \int \frac{\mathrm{d}\boldsymbol{l}'}{(2\pi)^2} \left[E_{\boldsymbol{l}'} \cos 2\varphi_{\boldsymbol{l}'\boldsymbol{l}} - B_{\boldsymbol{l}'} \sin 2\varphi_{\boldsymbol{l}'\boldsymbol{l}} \right] \boldsymbol{l}' \cdot (\boldsymbol{l} - \boldsymbol{l}') \psi_{\boldsymbol{l}-\boldsymbol{l}'} \tag{3.74}$$

$$\delta \tilde{B}_{\boldsymbol{l}} = \int \frac{\mathrm{d}\boldsymbol{l}'}{(2\pi)^2} \left[E_{\boldsymbol{l}'} \sin 2\varphi_{\boldsymbol{l}'\boldsymbol{l}} + B_{\boldsymbol{l}'} \cos 2\varphi_{\boldsymbol{l}'\boldsymbol{l}} \right] \boldsymbol{l}' \cdot (\boldsymbol{l} - \boldsymbol{l}') \psi_{\boldsymbol{l}-\boldsymbol{l}'}, \tag{3.75}$$

where the $\varphi_{\boldsymbol{l}'\boldsymbol{l}} \equiv \varphi_{\boldsymbol{l}'} - \varphi_{\boldsymbol{l}}$. We assume the Stokes parameters as well as the lensing potential to be Gaussian, so

$$\langle X_{\boldsymbol{l}'}^* X_{\boldsymbol{l}} \rangle = \delta(\boldsymbol{l} - \boldsymbol{l}') C_{\boldsymbol{l}}^{XX} \tag{3.76}$$

$$\langle \psi_{\boldsymbol{l}'}^* \psi_{\boldsymbol{l}} \rangle = \delta(\boldsymbol{l} - \boldsymbol{l}') l^2 C_{\boldsymbol{l}}^{dd}. \tag{3.77}$$

and, in addition,

$$\langle T_{\boldsymbol{l}'}^* E_{\boldsymbol{l}} \rangle = \delta(\boldsymbol{l} - \boldsymbol{l}') C_{\boldsymbol{l}}^{TE} \tag{3.78}$$

while $C_{\boldsymbol{l}}^{TB} = C_{\boldsymbol{l}}^{EB} = 0$. These equations together with Eq. (3.75) implies that the off-diagonal correlations of the harmonic modes can be used for estimating the displacement field (or equivalently the lensing potential). For $\boldsymbol{l} \neq \boldsymbol{l}'$ and any choice $X, Y \in \{T, E, B\}$

$$\langle \tilde{X}_{\boldsymbol{l}} \tilde{Y}_{\boldsymbol{l}'} \rangle_{\mathrm{CMB}} \simeq \langle X_{\boldsymbol{l}} Y_{\boldsymbol{l}'} \rangle_{\mathrm{CMB}} + \langle \delta \tilde{X}_{\boldsymbol{l}} Y_{\boldsymbol{l}'} \rangle_{\mathrm{CMB}} + \langle X_{\boldsymbol{l}} \delta \tilde{Y}_{\boldsymbol{l}'} \rangle_{\mathrm{CMB}} \tag{3.79}$$

$$= f^{\alpha}(\boldsymbol{l}, \boldsymbol{l}') \psi_{\boldsymbol{l}-\boldsymbol{l}'}^{\alpha} \tag{3.80}$$

where α denotes the XY combination, $\langle \cdots \rangle_{\mathrm{CMB}}$ means ensemble average over realisations of the CMB (not the lensing potential) and the proportionality constant $f^{\alpha}(\boldsymbol{l}, \boldsymbol{l}')$ depends only on the power spectra of the unlensed CMB and some known geometrical factor (see Hu and Okamoto, 2002, Table 1). Therefore, in order to estimate $d_{\boldsymbol{L}}$ trough the α cross correlations we can coadd all the products as the one in Eq. (3.80) such that $\boldsymbol{L} = \boldsymbol{l} - \boldsymbol{l}'$

$$d_{\boldsymbol{L}}^{\alpha} = \frac{A(L)}{L} \int \frac{\mathrm{d}\boldsymbol{l}}{(2\pi)^2} X_{\boldsymbol{l}} Y_{\boldsymbol{l}'} F^{\alpha}(\boldsymbol{l}, \boldsymbol{l}'). \tag{3.81}$$

Since for different $\boldsymbol{l}$ and $\boldsymbol{l}'$ the product $X_{\boldsymbol{l}} Y_{\boldsymbol{l}'}$ has typically different signal to uncertainty, the generic $F^{\alpha}(\boldsymbol{l}, \boldsymbol{l}')$ weights their coaddition. The factor

$$A(L) \equiv L^2 \left[\int \frac{\mathrm{d}\boldsymbol{l}}{(2\pi)^2} f^{\alpha}(\boldsymbol{l}, \boldsymbol{l}') F^{\alpha}(\boldsymbol{l}, \boldsymbol{l}') \right]^{-1} \tag{3.82}$$

ensures that the estimator is unbiased.

The power spectrum of the displacement field can be computed from the $d^{\alpha} d^{\beta}$ product for any α and β, it is therefore computed from the 4-point correlation function of the observed CMB maps. For example, in The POLARBEAR Collaboration (2014) the $\alpha, \beta = EE, EB$ and $\alpha, \beta = EB, EB$ were considered. In general one can combine all the d^{α} into a minimum variance coaddition d^{mv}. For current experiments the d^{mv} is dominated by the αs that involve temperature, TT in particular. However, as the sensitivity of the experiments increases, polarization will take over and eventually the minimum variance estimator will be dominated by the EB estimator (Hu and Okamoto, 2002).

3.8.4 The lensed CMB power spectra

Similarly to the previous section, the lensed CMB power spectra can be estimated by making a lowest order expansion that expresses the lensed Stokes parameters in terms of the unlensed CMB and the lensing potential. The only major difference is that in order to get per cent-level accuracy up to $\ell \sim 2000$ we have to include second order terms,

$$\tilde{X} = X + \boldsymbol{\alpha} \cdot \nabla X + \sum_{ij} \alpha^j \alpha^i \nabla_i \nabla_j X + \mathcal{O}(\boldsymbol{\alpha}^3). \tag{3.83}$$

This expression is then written in Fourier domain and the power spectrum is computed up to the second order. Making the usual Gaussianity assumptions about the primordial CMB and assuming
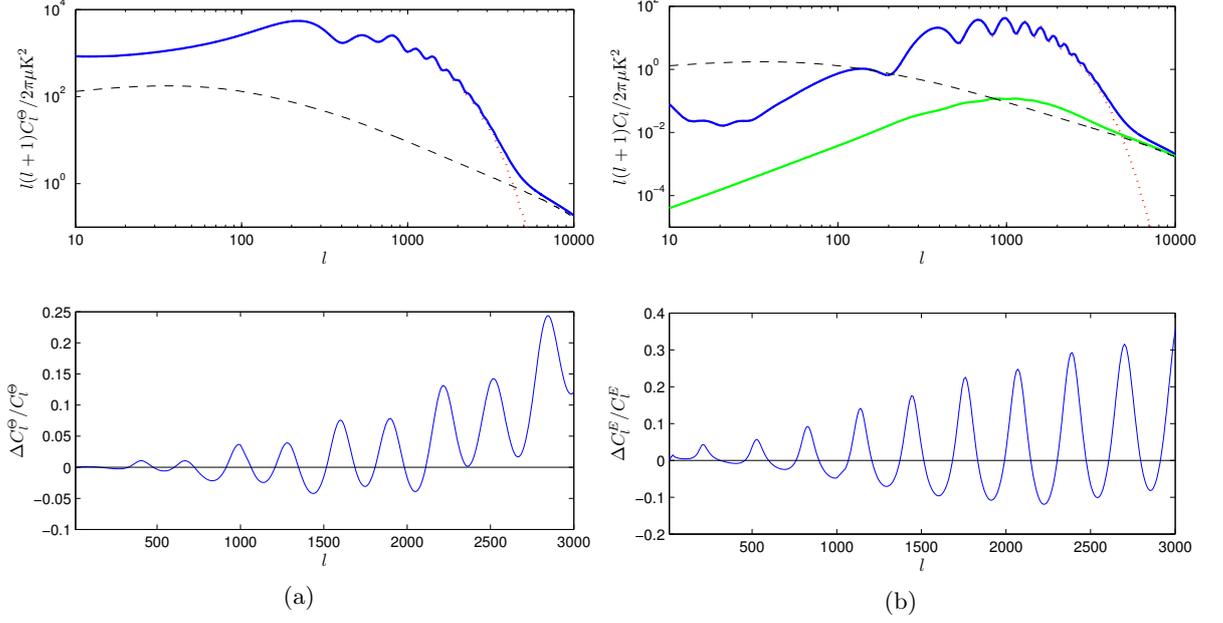
Figure 3.8: Temperature (a) and polarization (b) lensed power spectra (*top*) and their difference with respect to the unlensed spectra (bottom). The dashed line is the power spectrum of the lensing potential. (Lewis and Challinor, 2006)

$C_l^{TB} = C_l^{EB} = C_l^{BB} = 0$, the lensed power spectra are (Hu, 2000)

$$C_l^{\tilde{T}\tilde{T}} \simeq (1 - l^2 R^\psi)C_l^{TT} + \int \frac{d\boldsymbol{l'}}{(2\pi)} \left[\boldsymbol{l'} \cdot (\boldsymbol{l} - \boldsymbol{l'})\right]^2 C_{|\boldsymbol{l}-\boldsymbol{l'}|}^{\psi\psi} C_{l'}^{TT} \tag{3.84}$$

$$C_l^{\tilde{E}\tilde{E}} \simeq (1 - l^2 R^\psi)C_l^{EE} + \int \frac{d\boldsymbol{l'}}{(2\pi)} \left[\boldsymbol{l'} \cdot (\boldsymbol{l} - \boldsymbol{l'})\right]^2 C_{|\boldsymbol{l}-\boldsymbol{l'}|}^{\psi\psi} C_{l'}^{EE} \cos^2 2(\phi_{\boldsymbol{l'}} - \phi_{\boldsymbol{l}}) \tag{3.85}$$

$$C_l^{\tilde{T}\tilde{E}} \simeq (1 - l^2 R^\psi)C_l^{TE} + \int \frac{d\boldsymbol{l'}}{(2\pi)} \left[\boldsymbol{l'} \cdot (\boldsymbol{l} - \boldsymbol{l'})\right]^2 C_{|\boldsymbol{l}-\boldsymbol{l'}|}^{\psi\psi} C_{l'}^{TE} \cos 2(\phi_{\boldsymbol{l'}} - \phi_{\boldsymbol{l}}) \tag{3.86}$$

$$C_l^{\tilde{B}\tilde{B}} \simeq \int \frac{d\boldsymbol{l'}}{(2\pi)} \left[\boldsymbol{l'} \cdot (\boldsymbol{l} - \boldsymbol{l'})\right]^2 C_{|\boldsymbol{l}-\boldsymbol{l'}|}^{\psi\psi} C_{l'}^{EE} \sin^2 2(\phi_{\boldsymbol{l'}} - \phi_{\boldsymbol{l}}), \tag{3.87}$$

where we used half the total deflection angle power

$$R^\psi \equiv \frac{1}{4\pi} \int \frac{dl}{l} l^4 C_l^{\psi\psi}. \tag{3.88}$$

The integral term is basically a convolution of an unlensed CMB spectrum with the spectrum of the lensing potential. Notice that for BB, we are assuming no primordial power and the unlensed spectrum convolved is actually EE: these are the so-called *lensing B modes*. The lensed spectra are reported in Fig. 3.8.

At large scales and intermediate scales, for everything but BB, the result of the convolution is a blurring of the acoustic peaks with a characteristic kernel width of $\delta l \sim 50$. Consequently, the peaks of the lensed spectra are smoother compared to the unlensed one. This effect is more visible on EE then TT because the former has a sharper structure of peaks (compare the a and b bottom panels of Fig. 3.8).

As far as the large scales of BB are concerned, since EE has little power at small l, most of the contribution of the convolution comes from $l \ll l'$ and thus we can assume $C_{|\boldsymbol{l}-\boldsymbol{l'}|}^{\psi\psi} \sim C_{l'}^{\psi\psi}$. The integral then becomes a constant contribution, independent of l. Consequently, at large scales the power spectrum of the lensing B modes is approximately white, for standard ΛCDM $C_l^{\tilde{B}\tilde{B}} \sim 2 \times 10^{-6}\,\mu\text{K}^2$, which corresponds to the power of a $\sim 5\,\mu\text{Karcmin}$ noise level.

At small scales, primordial spectra fall sharply because of the diffusion dumping, therefore the convolution term of both temperature and polarization spectra results in an increase of power at $l > 3000$. Moreover, since the primordial term (first term) of the lensed spectra is negligible, the lensed EE and BB converge toward the same spectrum (see the top panel of Fig. 3.8b).
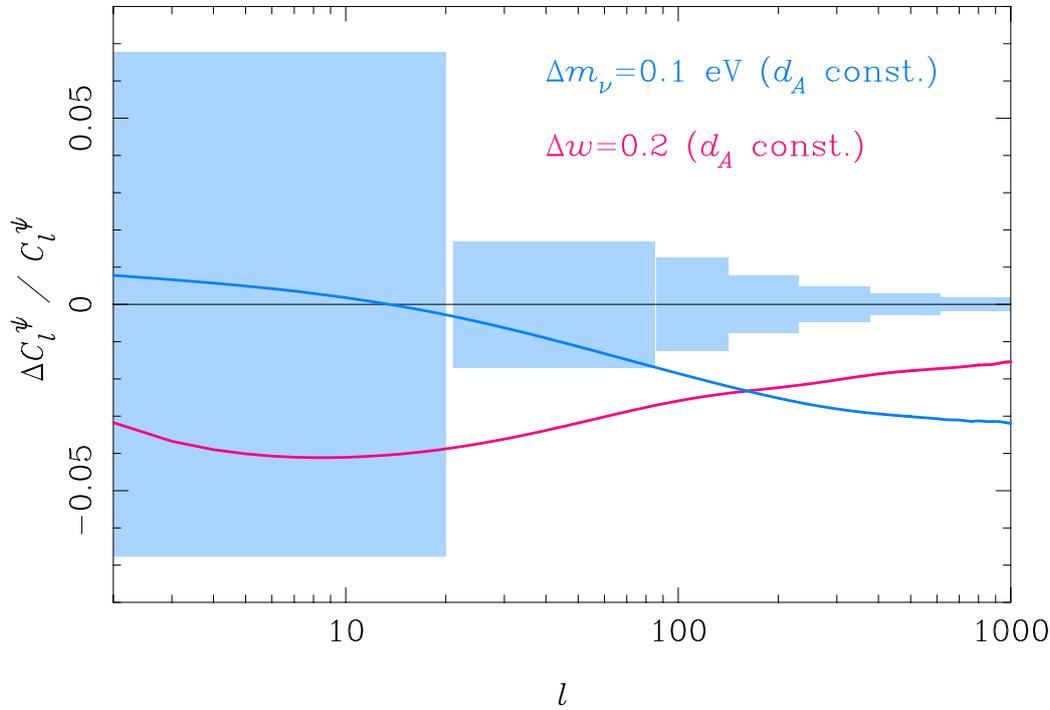
Figure 3.9: Effect on the lensing potential of a neutrino mass $\sum_\nu m_\nu = 0.1\,\mathrm{meV}$ and an equation of state of dark energy $w = .8$, compared to the standard massless neutrino and $w = -1$ case. The error boxes are the cosmic variance. (Lewis and Challinor, 2006)

3.8.5   Neutrino mass and dark energy from CMB lensing

The neutrino mass and dark energy are main drivers of the quest for lensing B-modes. The physics behind the primordial anisotropies is basically insensitive to these two components. At recombination the standard dark energy was negligible and neutrinos were relativistic, consequently their number of degrees of freedom mattered but not their mass. On the contrary, lensing is sensitive to late times, when the mass of the neutrinos becomes relevant and dark energy starts dominating the universe. Therefore, it is more effective in probing the total mass of the neutrinos and the properties of dark energy. In Fig. 3.9 and 3.10 we show the effect of a neutrino mass $\sum_\nu m_\nu = 0.1\,\mathrm{meV}$ and an equation of state of dark energy $w = 0.8$, compared to the standard massless neutrino and $w = -1$ case. These scenarios affect both the lensing potential and the CMB power spectrum. BB is particularly sensitive because it is not affected by the cosmic variance of the unlensed spectrum, which constitutes the dominant contribution to the TT and EE spectrum.
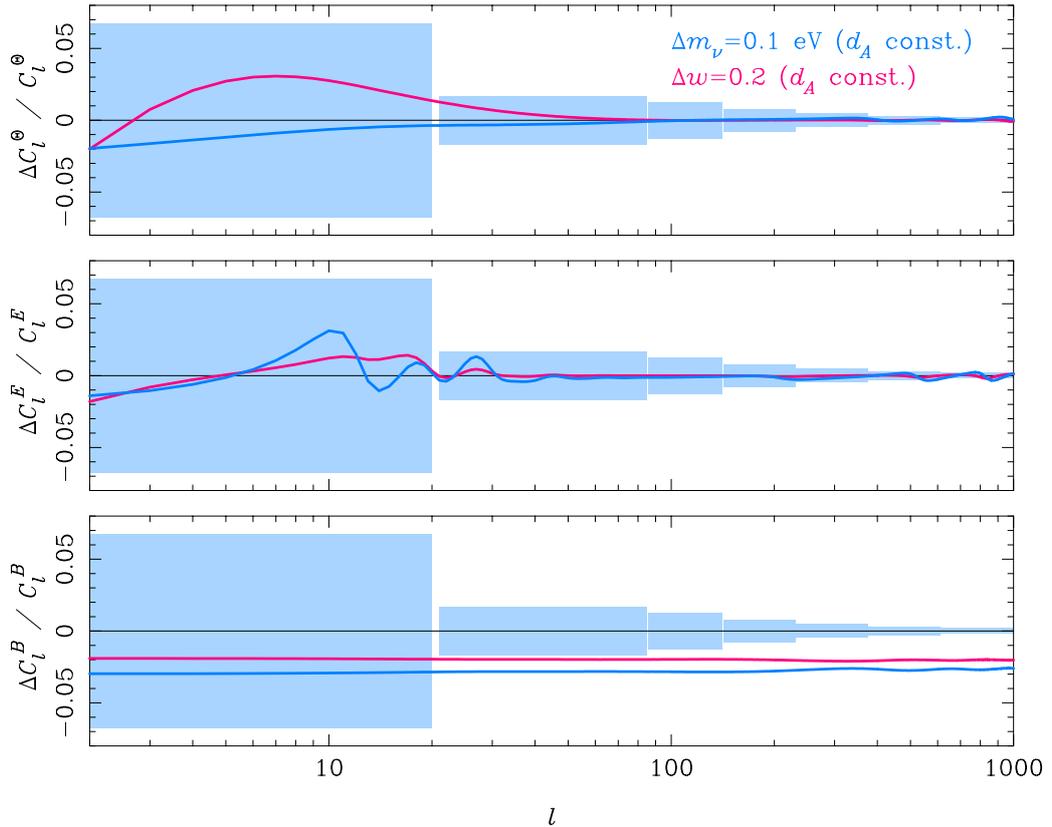
Figure 3.10: Effect on the CMB power spectrum of a neutrino mass $\sum_\nu m_\nu = 0.1\,\mathrm{meV}$ and an equation of state of dark energy $w = .8$, compared to the standard massless neutrino and $w = -1$ case. The error boxes are the cosmic variance. (Lewis and Challinor, 2006)

3.8.6   CMB lensing measurements

In the last decade we have seen tremendous progress in the measurements of CMB lensing, moving from first detections to stringent cosmological constraints. Measurements are of two types. The lensing potential can be reconstructed from CMB data only, as explained in Sec. 3.8.3. Alternatively, the lensing potential estimated with the quadratic estimators (Eq. 3.81) can be correlated with tracers of the gravitational potential, like the Cosmic Infrared Background (CIB) or galaxy and radio surveys: since they are sensitive to the same gravitational potential (at least for some redshift range) they are expected to have a positive correlation.

The first detection was made by Smith et al. (2007) through cross-correlation of the quadratic estimator applied on WMAP maps and radio galaxy counts from the NVSS survey. Similarly, Hirata et al. (2008) correlated the WMAP-estimated potential with the NVSS and SDSS surveys. Using ACT data and data from the Canada-France-Hawaii Telescope, Hand et al. (2015) provided the first cross-correlation between CMB lensing and galaxy lensing.

The first detection of the lensing potential from CMB only–through the four point correlation– was done by ACT (Das et al., 2011). Soon after, the constraints were improved by SPT (van Engelen et al., 2012) with a similar measurement. This type measurements started to really deliver cosmological constraints with the *Planck* satellite (The Planck Collaboration, 2014): compared to the CMB power spectrum constraints, lensing—detected at 25σ—allowed to double the constraints on curvature and partially break the τ-A_s degeneracy.

When polarization data became available, polarization lensing estimators were used as a mean of detecting B modes. Cross-correlating SPTpol data with the CIB map from Herschel-SPIRE Hanson et al. (2013) obtained the first B-mode detection. Later the same year, POLARBEAR Collaboration (2014b) performed a cross-correlation of its polarization quadratic estimators with the Herschel-ATLAS
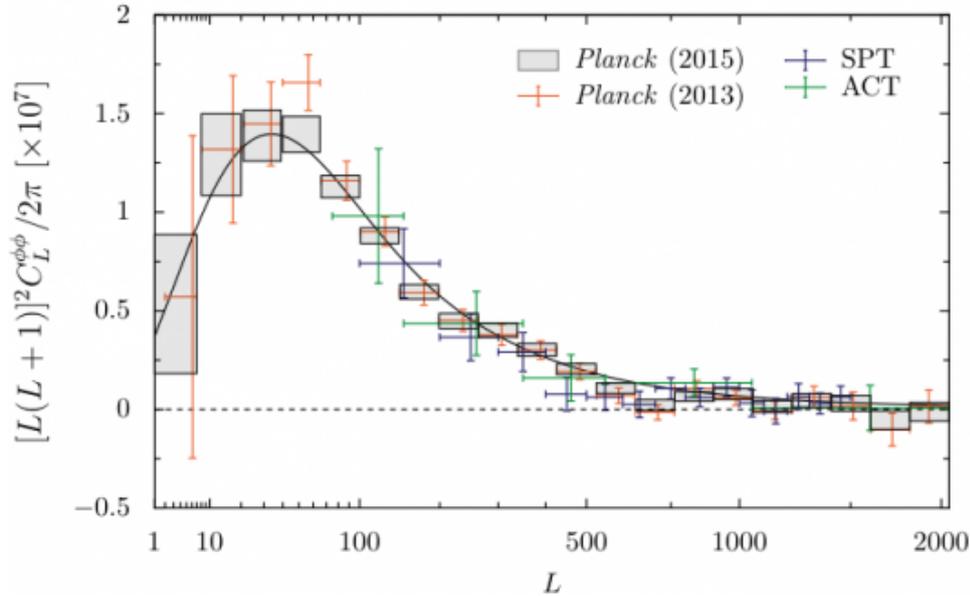
Figure 3.11: Lensing power spectrum reconstructed by *Planck* (The Planck Collaboration, 2014; Planck Collaboration, 2015d), SPT (van Engelen et al., 2012) and ACT (Das et al., 2014). (Image credits ESA/*Planck*)
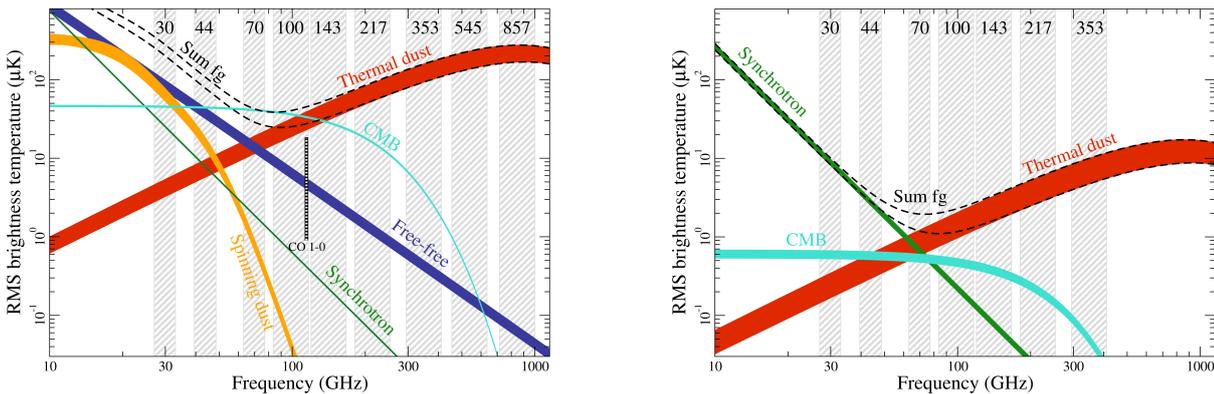
Figure 3.12: Intensity (left) and polarised (right) spectrum of the CMB and the foregrounds. Figure credits: *Planck*.

map. Also ACTpol performed a similar analysis by cross-correlating its temperature- and polarization-derived convergence fields with the CIB maps from *Planck* (van Engelen et al., 2015). In addition, the CMB lensing from galaxy clusters has been recently measured by ACT (Madhavacheril et al., 2015) and SPT (Baxter et al., 2015).

The first lensing reconstruction using polarization data alone was performed by The Polarbear Collaboration (2014). A similar analysis was recently published by The BICEP2/*Keck Array* Collaborations (2016). As previously mentioned, polarization data will eventually dominate the constraints on the lensing potential. However, we are not at that stage yet: the most stringent constraints on the lensing potential are given by Planck Collaboration (2015d), see Fig. 3.11, and are dominated by the TT quadratic estimator.

3.9   CMB foregrounds

One of the reasons why the CMB has been so powerful in constraining cosmology comes from the fact that observations did not encounter fundamental obstacles: improving constraints required mainly increasing the sensitivity of the experiments. Nowadays the situation is changing. For example, the constraints from the temperature power spectrum are cosmic variance limited. Another major obstacle comes from

the foreground emission by other astrophysical sources. Indeed, before the era of B-mode observation started, it was always possible to find a region in the sky and an observational frequency in which the amplitude of the CMB signal at large and intermediate scales was dominating over the foreground emission. However, the primordial B-mode signal (if any) is weaker than the foreground emission and properly disentangling between CMB and foreground has become more and more challenging.

Foregrounds are mostly of galactic origin and thus their intensity is typically more intense close to the galactic plane and fainter at galactic poles. Nevertheless some of them are extragalactic and, consequently, their signal is more uniform across the sky.

We now summarise the principal CMB foregrounds, emphasizing those relevant to polarization measurements.

Thermal dust. The relevance of thermal dust for B-mode measurements became strikingly clear after BICEP2 Collaboration (2014) and the follow-up study The BICEP2/*Keck Array* and *Planck* Collaborations (2015), which assessed that the detected large scales B-mode signal interpreted as a detection of $r = 0.2^{+0.07}_{-0.05}$ was in fact due to the foreground emission of galactic dust.

Part of the interstellar medium is indeed in the form of dust grains with size ranging from few nm to several μm. They are heated by ultraviolet and visible light and re-emit thermal radiation in the far-infrared. However, the tail of their emission law is relevant also in the microwaves.

Dust grains can have an asymmetric shape and emit preferably along their shortest axis, producing a polarized emission. The orientation of the grains in a region of the galaxy can be correlated due to the presence of a galactic magnetic field. Consequently, the emission of thermal dust can be polarized, with a typical polarization fraction ranging from few per cent up to $\sim 20\%$. Remarkably, directions with a high dust column density are less polarized due to the line-of-sight integration or, in other words, the polarization fraction can be stronger where the dust intensity signal is fainter.

The latest constraints from the *Planck* satellite (Planck Collaboration, 2016b) show that the dust emission law is consistent with a modified black-body with temperature $T_\mathrm{d} = 19.6\,\mathrm{K}$ and a spectral index of $\beta_\mathrm{d} = 1.59$, while the polarised angular power spectrum is consistent with a power law $C_\ell \propto \ell^\alpha$ with $\alpha = -2.45 \pm 0.02$ (interestingly, the BB power spectrum has approximately half the amplitude of the EE one). Therefore, dust emission is particularly important at high frequencies and large angular scales.

Synchrotron. At low frequencies another foreground obstruct CMB measurements: the synchrotron emission produced by cosmic-ray electrons spiraling around the galactic magnetic field. Its emission law is well described by a power law ν^{β_s} steeply decreasing, $\beta_\mathrm{s} = -3.16 \pm 0.40$ (Planck Collaboration, 2015d). Also the angular power spectrum follows a power law $\propto \ell^\alpha$, with $\alpha \in [-3.0, -2.6]$, depending on the galactic latitude (La Porta et al., 2008). At high galactic latitudes the emission is polarised at $10-40\%$ level (Planck Collaboration et al., 2015a) and has to be taken into account for B-mode studies (see e.g. Krachmalnicoff et al., 2015)

Point sources Some galaxies host an active galactic nucleus, which is thought to be the accretion of a massive black hole. These galaxies have a strong synchrotron emission and are loud in the radio frequencies. These sources have a Poissonian statistics and therefore have white angular power spectrum. At small scales the signal from unresolved point sources dominates over the CMB signal, which is suppressed by Silk dumping. Point sources can be polarized, typically between 1 and 5%, which makes them a potential foreground for small scales CMB polarization measurements.

Carbon monoxide The spectrum of the Milky Way interstellar medium contains line emissions. The most relevant for CMB observational frequencies is the rotational line emission from carbon monoxide (CO). The most significant transitions are $J = 1 \rightarrow 0$, $J = 2 \rightarrow 1$ and $J = 3 \rightarrow 2$ at approximately $115\,\mathrm{GHz}$, $230\,\mathrm{GHz}$ and $345\,\mathrm{GHz}$ respectively. CO lies mostly in the denser parts of the interstellar medium and thus its signal is concentrated at low and intermediate galactic latitudes. For more details and a full sky map of CO emission see Planck Collaboration (2014d).

CO emission is not expected to be significantly polarized but any line emission is potentially pernicious for polarization measurements. At different frequency in its band-pass, each detector has a different response. If the incoming signal has a smooth spectrum the detector-specific response is integrated over the band-pass. However, the signal recorded from line emission is very sensitive to the to specific response of different detectors. Therefore, when taking the difference between the signal of two orthogonal detectors the CO contribution does not perfectly cancel out and some intensity-to-polarization leakage is produced, typically of the order of few per cent.
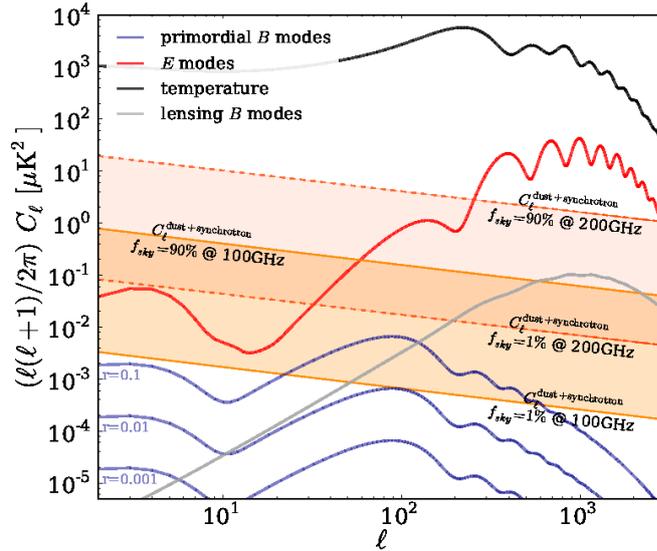
Figure 3.13: Estimated angular power spectrum of the polarised foregrounds at different frequencies and for different sky fractions. For comparison, the theoretical CMB power spectrum is also reported, including both the lensing B modes and primordial B modes for different values of the tensor to scalar ratio r. (Errard et al., 2016)

Cosmic infrared background   The infrared emission by the interstellar medium of all galaxies in the universe forms a background called cosmic Infrared background (CIB). The CIB carries information about the galaxies at all epochs and therefore it enables to probe objects that are too distant to be resolved. The CIB was detected for the first time by Puget et al. (1996) using data from the FIRAS instrument onboard the COBE satellite. Important constraints on the CIB spectrum were set in the past years thanks to the data from the Hershel Space Observatory (Béthermin et al., 2012).
For reviews on the CIB see Hauser and Dwek (2001) and Kashlinsky (2005).

Free-free   Free electrons interact with ions in our galaxy and produce bremsstrahlung radiation. This component is relevant at low frequencies and has a steep spectrum. Due to the incoherent nature of the process, the signal produced is not polarized to relevant levels and is negligible in CMB polarization measurements.

Others.   For completeness we mention, without elaborating on, other foreground signals which are not expected to be relevant to CMB polarization studies. The Zodiacal light is produced by the emission from interplanetary dust. Spinning dust signal is produced by dust grains with a magnetic dipole spinning at GHz frequencies. Magnetic dust emission is caused by changes in the magnetization of dust grains due to thermal fluctuation.

3.9.1   Foregrounds in future B-mode measurements.

Compared to previous CMB measurements, the detection of primordial B-modes is particularly challenging because the target signal is covered by foregrounds.

Foregrounds can be distinguished from the CMB signal owing to the different frequency dependence of their spectrum. The process of reconstructing the CMB from a mixture of CMB and foregrounds, measured at multiple frequencies, is called *component separation*.

Until a few years ago the approach of most of the ground-based B-mode experiments was to observe in one frequency band only a region of the sky known to have negligible foreground emission. Recent studies have demonstrated that no sky area can be safely considered foreground-free, especially if the large scales are to be probed. All the CMB experiments are thus upgrading their hardware, in order to have the multi-frequency coverage needed for component separation.

In the design of forthcoming experiments, it is crucial to assess how a given configuration will perform in terms of component separation. Errard et al. (2011) developed an efficient framework for accomplishing this task, enabling the quick exploration of large sets of experimental configurations.

Such techniques have been heavily used in designing and optimising the future instalments of the POLARBEAR observatories, including POLARBEARII and Simons Array (see e.g. Fig. 5.9c and Sec. 5.7). As they have not been applied to the POLARBEAR data sets, I do not elaborate on them in this thesis, the details are available in the quoted paper.

A key assumption in this method is that the sky model adopted in the component separation process matches the actual sky emission. During my thesis I worked on extending this method to the case in which the sky is more complex then the model assumed. The method is capable of efficiently evaluating the consequent bias and uncertainty of the cosmological parameter estimation, as described in Stompor et al. (2016).

3.10 Summary

The CMB anisotropies are a goldmine of cosmological information. Measuring B-modes is the most powerful probe we have for primordial gravitational waves and, consequently, it is a unique window on inflation and the physics of the early universe. On the other hand, the lensing B modes are very sensitive to structure formation and consequently to the properties of dark energy and the neutrino mass.

A successful detection and characterization of the CMB B-mode polarization will have an enormous impact not only on the modern cosmology but also on fundamental physics. However, extracting this invaluable information requires accurate measurements and sophisticated data analysis techniques. This topic is the subject of the rest of the present work.

Chapter 4

MAP-MAKING AND POWER SPECTRUM ESTIMATION FOR THE CMB EXPERIMENTS

Nowadays, in the quest for B-modes, the time-ordered data (TOD) of CMB experiments consist of tens of billions of samples. This volume is impractical for cosmological studies (but see e.g. Wandelt et al., 2004, who estimate cosmological parameters directly from the TOD), therefore data are compressed in a sequence of intermediate steps that aim at reducing the data size while preserving the relevant cosmological information. This operation is typically done in two steps: map-making and power-spectrum estimation. The former reconstructs a map of the observed sky signal, the latter use this map to estimate the power spectrum of the CMB across the sky. These problems are well understood and, under broad assumptions, these operations can be lossless (Tegmark, 1996). However, such algorithms are not applicable in modern data analysis. The growing size and the complexity of present and forthcoming CMB datasets poses serious computational challenges and require the design of new algorithms capable of extracting the cosmological information within accessible computational costs.

In this chapter, I introduce the map-making formalism as used later in the POLARBEAR data analysis, Sec. 4.1. The major focus is the inclusion of the time-domain filtering in the formalism. In Sec. 4.2 we describe in detail what are the consequences of this filtering and, more specifically, when it can prevent a faithful reconstruction of the sky map. We also discuss the typical filters applied in the contest of ground-based CMB experiments. This part is based on Poletti et al. (2016). Then, in Sec. 4.3 we discuss the second step of the information compression, the power-spectrum estimation. We focusing on pseudo-power estimation methods in particular, as they are the only computationally accessible option for high resolution experiments.

4.1 Map-making

Map-making aims at compressing the volume of the data—typically by 3 to 6 orders of magnitude—while preserving the cosmological information (e.g., Janssen and Gulkis, 1992; Wright et al., 1996; Tegmark, 1997). Map-making is usually phrased as a linear problem and the generalized least square (GLS) technique gives a closed form solution. In this section we illustrate this formalism. Then, we amend the formalism to accommodate the presence of the time domain filtering, always required in the analysis. Finally, we explain the possible degeneracies in the formalism and how to deal with them.

4.1.1 The standard map-making

The starting point of map-making is the calibrated time ordered data recorded by the detectors. We collect all these time samples in a vector, $\boldsymbol{d}$, which contains thus $\mathcal{N}_t$ elements. The scanning strategy of the telescope and the polarization modulation defines how the sky signal contributes to each measured sample, d_t, which can be then represented as,

$$d_t = I_{p_t} + \cos(2\varphi_t)Q_{p_t} + \sin(2\varphi_t)U_{p_t} + n_t. \tag{4.1}$$

Here n_t is the noise, I, Q and U are the Stokes parameters of the incoming light for the sky pixel p being observed at time t. φ is the orientation of the linear polarization sensitive detector projected on the sky.

We assume in this work that the instrumental beams are axially symmetric and are therefore convolved with the sky signals, which we aim at estimating. There are two important, specific cases of (4.1), which we will find useful further in this work. One is the total intensity measurement,

$$d_t = I(p_t) + n_t, \tag{4.2}$$

and the other is the polarization-only measurement,

$$d_t = \cos(2\varphi_t)Q_{p_t} + \sin(2\varphi_t)U_{p_t} + n_t. \tag{4.3}$$

All the Stokes parameters characterizing the sky signal in $\mathcal{N}_p$ observed sky pixels can be arranged in a single signal $\mathcal{N}_s$-vector, $\boldsymbol{s}$. For concreteness, we sort the vector in such a way that the Stokes parameters for one pixel are followed by the Stokes parameter for a subsequent one. The entire data vector, $\boldsymbol{d}$, can be then represented in a compact way as,

$$\boldsymbol{d} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{n}. \tag{4.4}$$

Here, noise vector $\boldsymbol{n}$ is assumed to have zero expected value and covariance matrix $\boldsymbol{C_n}$. The *pointing matrix* $\boldsymbol{A}$ is a $\mathcal{N}_t$ by $\mathcal{N}_s$ known matrix. Each row of $\boldsymbol{A}$ defines a linear combination of the signal, which contributes to the measurement at the time corresponding to the row. A column of $\boldsymbol{A}$ is a "time domain signature" of the corresponding entry of $\boldsymbol{s}$, telling us when a given sky pixel was observed and with what weight it contributed to the measured signal.

The least squares estimator

Map-making is a linear statistical problem. The generalised least square (GLS) method provides the following estimator of the signal vector

$$\hat{\boldsymbol{s}} = (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{d}, \tag{4.5}$$

for any choice of the positive definite weight matrix, $\boldsymbol{M}$. An intuitive geometrical interpretation of the estimator is provided at the end of Sec. 4.1.4. The estimator is unbiased: $\langle \boldsymbol{d} \rangle = \boldsymbol{A}\boldsymbol{s}$ and therefore $\langle \hat{\boldsymbol{s}} \rangle = (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A} \boldsymbol{s} = \boldsymbol{s}$.

It is derived as minimisation of the quadratic form

$$g(\boldsymbol{s}) = (\boldsymbol{A}\boldsymbol{s} - \boldsymbol{d})^\top \boldsymbol{M}(\boldsymbol{A}\boldsymbol{s} - \boldsymbol{d}), \tag{4.6}$$

which can be analytically computed as

$$0 = \frac{\partial \mathrm{g}(\boldsymbol{s})}{\partial \boldsymbol{s}} = 2\left(\frac{\boldsymbol{A}\boldsymbol{s}}{\partial \boldsymbol{s}}\right)^\top \boldsymbol{M}(\boldsymbol{A}\boldsymbol{s} - \boldsymbol{d}) = 2\boldsymbol{A}^\top \boldsymbol{M}(\boldsymbol{A}\boldsymbol{s} - \boldsymbol{d}). \tag{4.7}$$

Since $\boldsymbol{M}$ is positive definite, $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$ is non-singular and thus invertible if and only if $\boldsymbol{A}$ is an injection, i.e. if its column are linearly independent. For time being, we assume that $\boldsymbol{A}$ is indeed full rank, we discuss this issue later on. The covariance of the matrix of the estimator is

$$\boldsymbol{C}_{\hat{\boldsymbol{s}}} = (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{C_n} \boldsymbol{M} \boldsymbol{A})(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \tag{4.8}$$

Maximum likelihood estimator

If the noise vector is assumed to have a Gaussian distribution, the likelihood function takes the following form

$$\begin{aligned} -2\ln \mathcal{L}(\boldsymbol{s}) &\equiv \boldsymbol{n}^\top \boldsymbol{C_n}^{-1} \boldsymbol{n} \\ &= (\boldsymbol{A}\boldsymbol{s} - \boldsymbol{d})^\top \boldsymbol{C_n}^{-1}(\boldsymbol{A}\boldsymbol{s} - \boldsymbol{d}) + \mathrm{const.} \end{aligned} \tag{4.9}$$

By a change of the constant in the previous equation, the likelihood can be expressed in the following form

$$-2\ln \mathcal{L}(\boldsymbol{s}) = [(\boldsymbol{A}^\top \boldsymbol{C_n}^{-1} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{C_n}^{-1} \boldsymbol{d} - \boldsymbol{s}]^\top \boldsymbol{A}^\top \boldsymbol{C_n}^{-1} \boldsymbol{A}[(\boldsymbol{A}^\top \boldsymbol{C_n}^{-1} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{C_n}^{-1} \boldsymbol{d} - \boldsymbol{s}] + \mathrm{const.} \tag{4.10}$$

Since $\boldsymbol{A}^{\top}\boldsymbol{C}_{\boldsymbol{n}}^{-1}\boldsymbol{A}$ inherits the positive definiteness from $\boldsymbol{C}_{\boldsymbol{n}}$, the RHS clearly has a minimum when the $(\boldsymbol{A}^{\top}\boldsymbol{C}_{\boldsymbol{n}}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{C}_{\boldsymbol{n}}^{-1}\boldsymbol{d} = \boldsymbol{s}$. Therefore, for $\boldsymbol{M} = \boldsymbol{C}_{\boldsymbol{n}}^{-1}$ the GLS estimator in Eq. (4.5) coincide with the maximum likelihood estimator. In this specific case, the covariace matrix of the estimator reduces to

$$\boldsymbol{C}_{\hat{\boldsymbol{s}}} = (\boldsymbol{A}^{\top}\boldsymbol{M}\boldsymbol{A})^{-1} \tag{4.11}$$

This estimator not only is minimum variance—the Fisher information of the estimator and the TOD is the same—but also saturates the Creamér-Rao bound, i.e. the variance of the estimator is the inverse of its Fisher information. Consequently, for any mode (component of $\boldsymbol{s}$), any other unbiased estimator has larger variance. In a more picturesque way, this means that for any other unbiased estimator $\hat{\boldsymbol{s}}'$, the ellipsoid defined by the vectors $\boldsymbol{u}$ such that $(\boldsymbol{u} - \boldsymbol{s})^{\top}\boldsymbol{C}_{\hat{\boldsymbol{s}}'}^{-1}(\boldsymbol{u} - \boldsymbol{s}) = \text{const}$ lies everywhere outside $(\boldsymbol{u} - \boldsymbol{s})^{\top}(\boldsymbol{A}^{\top}\boldsymbol{C}_{\boldsymbol{n}}^{-1}\boldsymbol{A})(\boldsymbol{u} - \boldsymbol{s}) = \text{const}$.

4.1.2   The need for filtering

The analysis of cosmic microwave background (CMB) datasets typically requires some filtering of the raw TOD. The purpose of this operation can be removing from the TOD some signal different from the CMB, for example the ground pickup. Another reason can be that some component of the noise have a very large variance. It is the case of the low-frequency noise, which typically arises either from the atmospheric contamination or instrumental thermal drifts. Often, exploiting the corresponding time-domain information does not provide significant constraints on the signal but complicates the implementation and increases the computational cost of the estimator. Consequently, such modes are filtered, effectively discarding any information about the signal they might contain. Both of these examples will be studied in more detail in Sec. 4.2.4.

In the following section we explain how to build a filtering operator that properly removes all the unwanted components. We show that a heuristic unbiased map estimator can be built by substituting the filtering operator to the weight matrix in Eq. (4.5).

We continue with the meta-pixel approach and show that the same estimator can be equivalently obtained by introducing extra components in Eq. (4.4) and solving for the parameters of this extended data model.

4.1.3   The filtering operator

The specific filters employed are adapted to the nature and the amount of the contaminating signals. However, filtering generally works as follows. The contaminating signal can be written as an unknown amplitude y multiplying a known time domain template $\boldsymbol{t}$. Filtering sets to zero the amplitude of the $\boldsymbol{t}$ component of the timestream:

$$\boldsymbol{d}' = \left(\boldsymbol{1} - \boldsymbol{t}(\boldsymbol{t} \cdot \boldsymbol{t})^{-1}\boldsymbol{t}^{\top}\right)\boldsymbol{d} \equiv \boldsymbol{F}_{\boldsymbol{t}}\boldsymbol{d}, \tag{4.12}$$

so that $\boldsymbol{t} \cdot \boldsymbol{d}' = 0$.

In the rest of this section we refine this elementary filtering operation in order to properly handle an arbitrary number of templates $\boldsymbol{t}_i$ that we collect in the columns of a single *template matrix* $\boldsymbol{T}$. Our filtering operator will be

$$\boldsymbol{F}_{\boldsymbol{T}} = \boldsymbol{M} - \boldsymbol{M}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{M}\boldsymbol{T})^{-1}\boldsymbol{T}^{\top}\boldsymbol{M} \tag{4.13}$$

where $\boldsymbol{M}$ can be any full rank matrix.

The simplified approach, actually adopted by many of the current CMB data analysis pipelines (e.g. POLARBEAR Collaboration, 2014a; BICEP2 Collaboration, 2014; Keisler et al., 2015), is to sequentially apply the filters $\boldsymbol{F}_{\boldsymbol{t}_i}$. This approach has two drawbacks. First, the result depends on the order of the filters because in general

$$\boldsymbol{F}_{\boldsymbol{t}_i}\boldsymbol{F}_{\boldsymbol{t}_j} = \boldsymbol{1} - \boldsymbol{t}_i\boldsymbol{t}_i^{\top} - \boldsymbol{t}_j\boldsymbol{t}_j^{\top} + \boldsymbol{t}_i(\boldsymbol{t}_i^{\top}\boldsymbol{t}_j)^{-1}\boldsymbol{t}_j^{\top} \neq \tag{4.14}$$

$$\neq \boldsymbol{1} - \boldsymbol{t}_i\boldsymbol{t}_i^{\top} - \boldsymbol{t}_j\boldsymbol{t}_j^{\top} + \boldsymbol{t}_j(\boldsymbol{t}_j^{\top}\boldsymbol{t}_i)^{-1}\boldsymbol{t}_i^{\top} = \boldsymbol{F}_{\boldsymbol{t}_j}\boldsymbol{F}_{\boldsymbol{t}_i}, \tag{4.15}$$

where, for simplicity, we have assumed that all the templates are normalised, $\boldsymbol{t}_i^{\top}\boldsymbol{t}_i = 1$. Second, the templates are not completely filtered out since, in general

$$\boldsymbol{F}_{\boldsymbol{t}_i}\boldsymbol{F}_{\boldsymbol{t}_j}\boldsymbol{t}_j = 0 \quad \text{but} \quad \boldsymbol{F}_{\boldsymbol{t}_i}\boldsymbol{F}_{\boldsymbol{t}_j}\boldsymbol{t}_i = -\boldsymbol{t}_j\boldsymbol{t}_j^{\top}\boldsymbol{t}_i + \boldsymbol{t}_i(\boldsymbol{t}_i^{\top}\boldsymbol{t}_j)^{-1}\boldsymbol{t}_j^{\top}\boldsymbol{t}_i \neq 0. \tag{4.16}$$

Both these properties of the sequential filtering are caused by the fact that in general the templates are not orthogonal, $\boldsymbol{t}_i^\top \boldsymbol{t}_j \neq 0$.

The remedy consists in orthogonalising the templates before building the filtering operator. We build an orthonormal version of the template matrix as

$$\tilde{\boldsymbol{T}} = \boldsymbol{T}\,(\boldsymbol{T}^\top \boldsymbol{T})^{-1/2}, \tag{4.17}$$

where $(\boldsymbol{T}^\top \boldsymbol{T})^{-1/2}$ is any matrix that once squared gives $(\boldsymbol{T}^\top \boldsymbol{T})^{-1}$. We note that the number of columns of $\tilde{\boldsymbol{T}}$ may be smaller than that of $\boldsymbol{T}$, this happens when some of the original templates are not linearly independent. In such cases, the matrix $(\boldsymbol{T}^\top \boldsymbol{T})^{-1}$ has to be interpreted as the pseudo-inverse. The fact that the columns of $\tilde{\boldsymbol{T}}$ are orthonormal is apparent from

$$\tilde{\boldsymbol{T}}^\top \tilde{\boldsymbol{T}} = (\boldsymbol{T}^\top \boldsymbol{T})^{-1/2\,\top}\,\boldsymbol{T}^\top\,\boldsymbol{T}\,(\boldsymbol{T}^\top \boldsymbol{T})^{-1/2}\,=\,\boldsymbol{1}. \tag{4.18}$$

They are also a basis for the columns of $\boldsymbol{T}$, the space of the modes that we want to filter out. The fact that $(\boldsymbol{T}^\top \boldsymbol{T})^{-1/2}$ is not unique reflects the fact that the choice of the basis of the columns of $\boldsymbol{T}$ is not unique.

Now that we have an orthonormal version of the templates we can build the filtering operator

$$\begin{aligned}
\boldsymbol{F_T} &= \boldsymbol{1} - \tilde{\boldsymbol{T}}\tilde{\boldsymbol{T}}^\top \\
&= \prod_i \boldsymbol{F}_{\tilde{\boldsymbol{t}}_i} \tag{4.19} \\
&= \boldsymbol{1} - \boldsymbol{T}(\boldsymbol{T}^\top \boldsymbol{T})^{-1}\boldsymbol{T}^\top. \tag{4.20}
\end{aligned}$$

Notice that the order in the product in Eq. (4.19) does not matter because the $\tilde{\boldsymbol{t}}_i$ are orthogonal. Since

$$\boldsymbol{F_T}\boldsymbol{T} = 0, \tag{4.21}$$

any possible unwanted contribution $\boldsymbol{T}\boldsymbol{y}$ is correctly filtered out by $\boldsymbol{F_T}$. This property persists if we generalise Eq. (4.20) to Eq. (4.13) by adding an arbitrary full rank weight matrix $\boldsymbol{M}$.

With help of this operator we can now generalize the map-making equation, (4.5), as

$$\hat{\boldsymbol{s}} = (\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{d}, \tag{4.22}$$

where $\boldsymbol{F_T}$ acting upon $\boldsymbol{d}$ removes all the unwanted modes and weights the others as requested, while the matrix operator $(\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1}$ ensures that the estimator is unbiased. Indeed, inserting (4.4) for $\boldsymbol{d}$ we get,

$$\hat{\boldsymbol{s}} = \boldsymbol{s} + (\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{n}, \tag{4.23}$$

and thus,

$$\langle \hat{\boldsymbol{s}} - \boldsymbol{s} \rangle = \langle (\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{n} \rangle = 0, \tag{4.24}$$

where the average is taken over the statistical ensemble of instrumental noise realizations and $\langle \boldsymbol{n} \rangle = 0$. The map domain noise covariance matrix of the unbiased map estimator is (see, e.g., Tegmark (1997) or Stompor et al. (2001b))

$$\begin{aligned}
\boldsymbol{C}_{\hat{\boldsymbol{s}}} &\equiv \langle (\hat{\boldsymbol{s}} - \boldsymbol{s})(\hat{\boldsymbol{s}} - \boldsymbol{s})^\top \rangle \\
&= (\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{F_T}\,\langle \boldsymbol{n}\boldsymbol{n}^\top \rangle\,\boldsymbol{F_T}\boldsymbol{A}(\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1} \\
&= (\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{F_T}\,\boldsymbol{C_n}\,\boldsymbol{F_T}\boldsymbol{A}\,(\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1} \tag{4.25}
\end{aligned}$$

If $\boldsymbol{M} = \boldsymbol{C_n}^{-1}$ then this can be rewritten in a compact way as,

$$\boldsymbol{C}_{\hat{\boldsymbol{s}}} = (\boldsymbol{A}^\top \boldsymbol{F_T}\boldsymbol{A})^{-1}, \tag{4.26}$$

owing to the fact that, $\boldsymbol{F_T}\,\boldsymbol{M}^{-1}\,\boldsymbol{F_T} = \boldsymbol{F_T}$.

4.1.4 Meta-pixel approach

We note that an equation analogous to (4.22) can be derived from the meta pixel approach (e.g., Stompor et al., 2001b; Cantalupo et al., 2010). In order to do so, let us generalize the data model in Eq. (4.4) to incorporate the effects discussed in Sec. 4.1.3. This can be done as follows,

$$\boldsymbol{d} = \boldsymbol{As} + \boldsymbol{Ty} + \boldsymbol{w}. \tag{4.27}$$

where $\boldsymbol{T}$ is in general a set of time domain templates that we want to filter out, $\boldsymbol{y}$ is their unknown amplitude and $\boldsymbol{w}$ is the noise. Rewriting Eq. (4.27) as,

$$\boldsymbol{d} = [\boldsymbol{A}, \boldsymbol{T}] \left[\begin{smallmatrix} \boldsymbol{s} \\ \boldsymbol{y} \end{smallmatrix}\right] + \boldsymbol{w}. \tag{4.28}$$

Estimating the unknown parameters for this data model or for the one in (4.4) is formally equivalent. Similarly to Sec. 4.1.1, we assume that $[\boldsymbol{A}, \boldsymbol{T}]$ is full rank: the time domain signature of the unknown parameters are all linearly independent, postponing a more pertinent discussion on degeneracies to Sec. 4.2. The GLS estimator for Eq. (4.27), written in a block wise fashion is

$$\hat{\boldsymbol{s}} = (\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d} \tag{4.29}$$

$$\hat{\boldsymbol{y}} = (\boldsymbol{T}^\top \boldsymbol{F_A} \boldsymbol{T})^{-1} \boldsymbol{T}^\top \boldsymbol{F_A} \boldsymbol{d} \tag{4.30}$$

where operators $\boldsymbol{F_T}$ and $\boldsymbol{F_A}$ are filtering operators as in (4.13). Given the relevance of Eq. (4.29) in this work, we provide a full derivation in Appendix A.

As (4.29) shows, $\hat{\boldsymbol{s}}$ can be computed without explicitly solving for the amplitude of the templates $\hat{\boldsymbol{y}}$. We call this the *direct* approach. Moreover, this is exactly the same expression we derived in the previous section, (4.22). There are however several advantages to this derivation. For instance, it shows that if the weights are chosen appropriately, i.e., $\boldsymbol{M} = \boldsymbol{C_n}^{-1}$, then the map estimated via (4.22) is both maximum likelihood and minimum variance. Moreover if the Bayesian perspective is adopted, then the posterior probability distribution for the combined vector of unknowns in (4.28), i.e., $\left[\begin{smallmatrix} \boldsymbol{s} \\ \boldsymbol{y} \end{smallmatrix}\right]$, is Gaussian and the second term in the expression for the filtering operator, $\boldsymbol{F_T}$, (4.13), arises simply as a result of a marginalization of this posterior over the unknowns contained in $\boldsymbol{y}$ assuming flat priors.

This manner of deriving the estimator also suggests that the map can be solved with a *two-step* approach. If instead of using (4.29) we first estimate $\hat{\boldsymbol{y}}$, then $\hat{\boldsymbol{s}}$ can be estimated using the following formula (see Appendix A for the proof),

$$\hat{\boldsymbol{s}} = (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{M} (\boldsymbol{d} - \boldsymbol{T}\hat{\boldsymbol{y}}), \tag{4.31}$$

which, for simple $\boldsymbol{M}$ may be much easier to solve than (4.29).

This latter approach provides a basis for the destriping technique (e.g., Poutanen et al., 2004; Keihänen et al., 2004, 2010). We emphasize that the direct or two-step methods are formally equivalent: they lead to the same estimate of the sky signal and they filter the same components of the TOD vector $\boldsymbol{d}$. Eq. (4.31) is usually easy and cheap to implement as the weight matrix, $\boldsymbol{M}$, is typically taken to be diagonal in this context. Choosing between the direct or two-step approach is therefore driven by the cost of (4.29) compared to (4.30). These two equations require handling dense algebraic objects of a dimension equal to the number of columns of $\boldsymbol{A}$ and $\boldsymbol{T}$ respectively. The former is proportional to the number of pixels observed while the latter to the number of temporal templates that have to be filtered. As this last number typically grows with the number of detectors and the presence of spurious signals in the observations, the direct method provides a potentially attractive approach for modern experiments, which employ kilo-pixel arrays of detectors. This includes ground-based observatories, such as POLARBEAR (The POLARBEAR Collaboration, 2014) but also planned CMB satellite missions, e.g., LITEBIRD (Matsumura et al., 2016) or CORE (The COrE Collaboration, 2011). In contrast, the two-step approach was very well adapted to experiments with a limited number of detectors observing big sky areas, which therefore produced maps with a large number of pixels but needed few templates to be removed. For these reasons, it has played a particularly prominent role in the analysis of the Planck data (Burigana et al., 1999; Maino et al., 1999; Tristram et al., 2011; Planck Collaboration, 2015a,b).

If $\boldsymbol{M}$ is diagonal then $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$ and $\boldsymbol{T}^\top \boldsymbol{M} \boldsymbol{T}$ are usually sparse and structured and therefore easy to compute and invert. On the contrary $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ and $\boldsymbol{T}^\top \boldsymbol{F_A} \boldsymbol{T}$ are dense and potentially big. The size of the former matrix is proportional to the number of pixels, ranging from 10^5 from ground-based experiments to 10^9 for satellites covering the whole sky. The size of the latter matrix also typically exceeds 10^8 for kilo-pixels arrays. These matrices are not only computationally-heavy to perform algebraic operations

on but often prohibitively large to permit their explicit computation and/or storage. As we discuss in more detail in Sec. 4.1.5, computing the inverse of dense matrices is computationally intensive and, consequently, the solution of the inverse problem either in (4.29) or in (4.30) would ideally be found using some iterative linear equations solvers such the preconditioned conjugate gradient method (PCG) (e.g., de Gasperis et al., 2005; Cantalupo et al., 2010; Szydlarski et al., 2014). However, the convergence of these iterative solvers may be hard to attain if the matrices are not well conditioned. We discuss the PCG approach in the context of the extended map-making equation, (4.22), in Sec. 4.2.3, from the formal point of view, and in Sec. 6.6.2 in the specific context of POLARBEAR.

Biased estimator

Performed directly or iteratively, the inversion of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ (or $\boldsymbol{T}^\top \boldsymbol{F_A} \boldsymbol{T}$) is the bottleneck of both the development and the computational time. This is why many experiments preferred the use of the following biased map estimator

$$\hat{\boldsymbol{s}} = (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d}. \tag{4.32}$$

As $\boldsymbol{F_T}$ still acts upon the data vector, $\boldsymbol{d}$, the templates are still explicitly filtered out of the data. However, this is not accounted for in the system matrix, $(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}$, which therefore does not correct for the filtering but only for the weighting. If $\boldsymbol{M}$ is diagonal, this choice simplifies enormously the implementation and drastically reduces the computational cost. The price to pay is a bias in the estimator. This bias is usually evaluated and removed at the power spectrum level using Monte Carlo simulations and typically requires some additional assumptions (e.g., Hivon et al., 2002). This approach is thus most frequently considered to be a part of the power spectrum estimation pipeline.

The presence of the bias on the map level is apparent as we have,

$$\text{Bias} = \langle \hat{\boldsymbol{s}} - \boldsymbol{s} \rangle = \left[(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A} - \boldsymbol{1} \right] \boldsymbol{s}, \tag{4.33}$$

which does not vanish in general. The covariance of the estimated map reads,

$$\begin{aligned} \boldsymbol{C}_{\hat{\boldsymbol{s}}} &\equiv \langle (\hat{\boldsymbol{s}} - \langle \hat{\boldsymbol{s}} \rangle)(\hat{\boldsymbol{s}} - \langle \hat{\boldsymbol{s}} \rangle)^\top \rangle \\ &= (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{F_T} \langle \boldsymbol{n} \boldsymbol{n}^\top \rangle \boldsymbol{F_T} \boldsymbol{A} (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \\ &= (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{C_n} \boldsymbol{F_T} \boldsymbol{A} (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}. \end{aligned} \tag{4.34}$$

Therefore, if the filters are chosen in the way that matrix $\boldsymbol{F_T} \boldsymbol{C_n} \boldsymbol{F_T}$ is nearly diagonal for any diagonal weights, $\boldsymbol{M}$, then we can take them to be, $\boldsymbol{M} = \text{diag}(\boldsymbol{F_T} \boldsymbol{C_n} \boldsymbol{F_T})$, yielding,

$$\boldsymbol{C}_{\hat{\boldsymbol{s}}} \approx (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}, \tag{4.35}$$

and therefore the covariance can have a particularly simple structure. However, generally this is rarely the case and instead even if the true noise is uncorrelated and $\boldsymbol{M} = \boldsymbol{C_n}^{-1}$, the expression for the covariance reads

$$\boldsymbol{C}_{\hat{\boldsymbol{s}}} = (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A} (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}, \tag{4.36}$$

and thus it is a dense matrix with potentially non-negligible, off-diagonal correlations.

We note that the informational content of both, the unbiased and biased, maps is the same as one can be derived from the other via an invertible linear operation. Indeed,

$$\hat{\boldsymbol{s}}_{\text{unbiased}} = \boldsymbol{R} \, \hat{\boldsymbol{s}}_{\text{biased}} \tag{4.37}$$

where $\boldsymbol{R} \equiv (\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A})^{-1} (\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})$ is an invertible matrix. However, the keystone of the biased approach is that no attempt is ever made to estimate the matrix $\boldsymbol{R}$. Consequently though the same information is contained in both maps, it is encoded differently in each of them and whenever the biased and unbiased maps are further processed in a similar way, the information is compressed differently in both cases, giving rise to different statistical properties of the resulting estimator.

Geometrical interpretation of the estimators

In this section I re-examine the key map-making equations from the geometric perspective and shows how the estimators can be interpreted as orthogonal projections in the vector space of time domain vectors. We start from the standard map-making and then move to the meta-pixel approach, both direct and two-step. For simplicity, but without losing generality, we will consider $M = 1$.

Before starting, we describe briefly the orthogonal projector (for ease of reading we will drop "orthogonal" from now on). Consider the columns of $\boldsymbol{X}$ to be the basis of the subspace $\mathrm{span}(\boldsymbol{X})$. The projector on $\mathrm{span}(\boldsymbol{X})$ is

$$\boldsymbol{P_X} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top. \tag{4.38}$$

For any $\boldsymbol{z} \in \mathrm{span}(\boldsymbol{X})$, $\boldsymbol{P_X z} = \boldsymbol{z}$. For any $\boldsymbol{z}$ orthogonal to $\mathrm{span}(\boldsymbol{X})$, $\boldsymbol{P_X z} = \boldsymbol{0}$. Any projector is idempotent: $\boldsymbol{P_X P_X} = \boldsymbol{P_X}$. It is true also for the projector on the space orthogonal to $\mathrm{span}(\boldsymbol{X})$: $\boldsymbol{P_X^\perp} = \boldsymbol{1} - \boldsymbol{P_X}$ (notice the similarity with the filtering operator). If $\boldsymbol{X}$ is not full rank and $\boldsymbol{X}^\top \boldsymbol{X}$ is not invertible one can still build the orthogonal projector by using the pseudoinverse in place of the inverse.

Forgetting for a second about $\boldsymbol{s}$, the data model Eq. (4.4) can be written as $\boldsymbol{d} = \boldsymbol{a} + \boldsymbol{n}$, with $\boldsymbol{a} \in \mathrm{span}(\boldsymbol{A})$. We now build the estimator $\hat{\boldsymbol{a}}$. The RHS of Eq. (4.9) is the squared of the length of $\boldsymbol{d} - \boldsymbol{a}$. Therefore the likelihood is maximized by the vector $\hat{\boldsymbol{a}}$ the closest to $\boldsymbol{d}$:

$$\hat{\boldsymbol{a}} = \boldsymbol{P_A d} \tag{4.39}$$

which is nothing but the estimator Eq. (4.5) multiplied by $\boldsymbol{A}$ from left. The map-making estimator Eq. (4.5) is the $\boldsymbol{s}$ such that $\boldsymbol{As}$ is the projection of $\boldsymbol{d}$ on $\mathrm{span}(\boldsymbol{A})$. Notice that this projection can always be determined (no assumption on $\boldsymbol{A}$ was made in the derivation of (4.39)). However, if $\boldsymbol{A}$ is not full rank (its columns are not linearly independent) there is not a one-to-one correspondence between the $\boldsymbol{a}$ and the $\boldsymbol{s}$ such that $\boldsymbol{a} = \boldsymbol{As}$. Consider $\boldsymbol{a}^{(1,2)} \equiv \boldsymbol{As}^{(1)} = \boldsymbol{As}^{(2)}$ for two linearly independent modes $\boldsymbol{s}^{(1)}$ and $\boldsymbol{s}^{(2)}$, there is no problem in constraining $\boldsymbol{a}^{(1,2)}$ but it is impossible to get separate constraints on $\boldsymbol{s}^{(1)}$ and $\boldsymbol{s}^{(2)}$.

We now consider the case of the data model containing time domain templates Eq. (4.27). We remind that it is the very same problem as before with $[\boldsymbol{A}, \boldsymbol{T}]$ in place of $\boldsymbol{A}$. The arguments on the degeneracies are the very same. We just emphasize that now $\boldsymbol{s}^{(1)}$ can be the sky mode and $\boldsymbol{s}^{(2)}$ a "template mode" $\boldsymbol{y}^{(2)}$. This said, we will now assume $[\boldsymbol{A}, \boldsymbol{T}]$ to be full rank.

As explained previously in this section, mapmaking means finding the $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{y}}$ such that $\boldsymbol{A}\hat{\boldsymbol{s}} + \boldsymbol{T}\hat{\boldsymbol{y}} = \boldsymbol{P}_{[\boldsymbol{A},\boldsymbol{T}]}\boldsymbol{d}$. The only difficulty comes from the fact that $\boldsymbol{A}$ and $\boldsymbol{T}$ are not orthogonal.

The following steps might not be as straightforward as the previous ones and it is convenient to use the low dimensional example in Fig. 4.1 to guide the intuition. $\boldsymbol{A}$ and $\boldsymbol{T}$ are still general, despite they are both depicted as single vectors. The following relations hold (only the second one is represented in the plot)

$$\boldsymbol{P_T^\perp A}\hat{\boldsymbol{s}} = \boldsymbol{P}_{\boldsymbol{P_T^\perp A}}\boldsymbol{d} \tag{4.40}$$

$$\boldsymbol{P_A^\perp T}\hat{\boldsymbol{y}} = \boldsymbol{P}_{\boldsymbol{P_A^\perp T}}\boldsymbol{d}. \tag{4.41}$$

They are Eq. (4.29) and (4.30) multiplied from left by $\boldsymbol{P_T^\perp A}$ and $\boldsymbol{P_A^\perp T}$ respectively. Eq. (4.40) shows that the direct estimation is equivalent to finding the projection of $\boldsymbol{d}$ on the component of $\boldsymbol{A}$ orthogonal to $\boldsymbol{T}$. The two-step approach can be interpreted by noting that

$$\boldsymbol{P_A d} = \boldsymbol{A}\hat{\boldsymbol{s}} + \boldsymbol{P_A T}\hat{\boldsymbol{y}} \tag{4.42}$$

is nothing but a rearrangement of Eq. (4.31) after the multiplication by $\boldsymbol{A}$ from left.

4.1.5 Solving the map-making inverse problem

The inverse problem is the main obstacle to the implementation and cost in the solution of the unbiased map-making estimator, Eq. (4.29). In this section we summarise how this linear algebra problem is typically tackled.

We want to find the $\boldsymbol{x}$ satisfying

$$\boldsymbol{Bx} = \boldsymbol{b} \tag{4.43}$$

where, in the case of Eq. (4.29), $\boldsymbol{B} = \boldsymbol{A}^\top \boldsymbol{F_T A}$ and $\boldsymbol{b} = \boldsymbol{AF_T d}$.
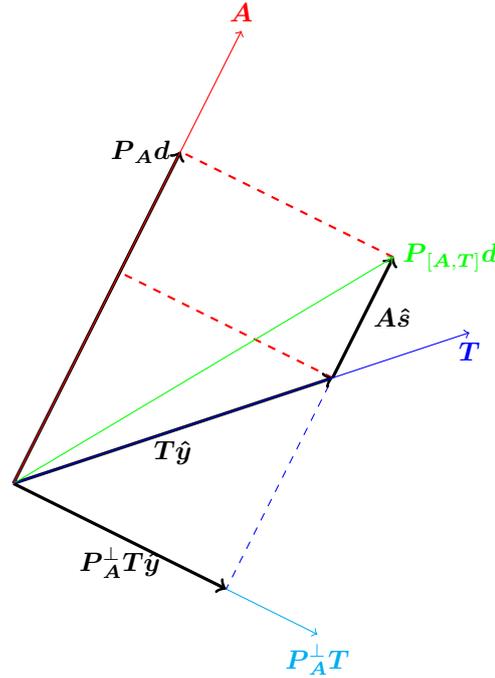
Figure 4.1: Geometric visualisation of the map-making operations. In particular, the figure highlights the equivalence between the direct and the two-step approach (see Eqs. (4.40-4.42) and the surrounding text)

Explicit method

It consists in actually computing $\boldsymbol{B}^{-1}$ and applying it on $\boldsymbol{b}$.

The advantages of this approach are

- the solution is numerically accurate;

- once $\boldsymbol{B}^{-1}$ is computed, solving $\boldsymbol{B}\boldsymbol{x}' = \boldsymbol{b}'$ is straightforward and efficient;

- the computation of the inverse can be performed by highly-optimised numerical libraries;

while the disadvantages are

- the scaling with the size of the system n is steep: the cost is $\mathcal{O}(n^3)$;

- it requires the explicit computation of $\boldsymbol{B}$, not only the capability of applying $\boldsymbol{B}$ on a vector.

If the matrix is ill-conditioned, the inversion requires some care. In order to have the numerical control of the process, the eigendecomposition of $\boldsymbol{B}$ has to be computed in order to identify and handle the degeneracies. In addition to the above list of pro and cons, the eigendecomposition adds valuable knowledge and control of the solution of the system, while preserving the $\mathcal{O}(n^3)$ scaling of the computational cost.

Iterative method

All the present mega-pixel map-makers cannot afford the direct solution and implement an iterative one: the preconditioned conjugate gradient (PCG). The idea of the PCG is the following. The $\boldsymbol{B}$ matrix is symmetric and positive definite, therefore it can be used to define a scalar product in the vector space. This scalar product choice has an advantage: the component of the solution along a normalised vector $\boldsymbol{p}$ is equal to $\boldsymbol{p}^\top \boldsymbol{B}\boldsymbol{x} = \boldsymbol{p}^\top \boldsymbol{b}$, thus computable without actually having $\boldsymbol{x}$.

The iterative process finds the projection of the solution over the space spanned by a growing set of orthonormal vectors. The first basis vector is typically $\boldsymbol{p}_0 = \boldsymbol{b}/\sqrt{\boldsymbol{b}^\top \boldsymbol{B}\boldsymbol{b}}$, the approximate solution is thus $\boldsymbol{x}_0 = (\boldsymbol{p}_0^\top \boldsymbol{b})\boldsymbol{p}_0$. The next dimension of the basis is chosen to be $\boldsymbol{r}_0 = \boldsymbol{b} - \boldsymbol{B}\boldsymbol{x}_0$, $\boldsymbol{p}_1$ is defined by removing the component of $\boldsymbol{r}_0$ along the previous basis vector and normalising the result. It is not guaranteed

that $\boldsymbol{p}_1$ actually exists since $\boldsymbol{r}_0$ could be proportional to $\boldsymbol{p}_0$. Assuming it is not the case, the solution can now be refined as $\boldsymbol{x}_1 = (\boldsymbol{p}_0^\top \boldsymbol{b})\boldsymbol{p}_0 + (\boldsymbol{p}_1^\top \boldsymbol{b})\boldsymbol{p}_1$. The process continues analogously by choosing as next dimension $\boldsymbol{r}_i = \boldsymbol{b} - \boldsymbol{B}\boldsymbol{x}_i$. The iterations are stopped either after some predefined number of iterations or when a convergence criterion is met, e.g. $|\boldsymbol{r}_i|^2$ is lower than some threshold.

However, the convergence cannot be guaranteed. Typically, the convergence is not achieved when the $\boldsymbol{B}$ matrix is ill-conditioned. Its condition number can be improved by using a preconditioner but no general prescription exists: preconditioners are typically *ad-hoc* and require some knowledge on the specific system to be solved.

Summarising the advantages of the iterative solution are

- the speed: it scales as the cost of the application of $\boldsymbol{B}$ times the number of iterations;

- $\boldsymbol{B}$ does not have to be explicitly computed;

while the downsides are

- no guarantee that the convergence criteria are met and

- when it is not the case, the control of the solution is limited;

- solving the system $\boldsymbol{B}\boldsymbol{x} = \boldsymbol{b}$ does not give any advantage in the solution of $\boldsymbol{B}\boldsymbol{x}' = \boldsymbol{b}'$

4.2 Degeneracies

So far we have assumed that the generalized map-making equation can be robustly solved, implicitly assuming that the system matrix, $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$, is invertible. This may not always be however the case, in this section we elaborate on those cases and discuss what can be done in such circumstances.

We first note that invertibility of the system matrix is ensured if matrix $[\boldsymbol{A}, \boldsymbol{T}]$ is full column rank. This can be seen immediately by noting that matrix $(\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A})^{-1}$ stands for an upper left diagonal block of the inverse of the matrix

$$\left[\begin{array}{c} \boldsymbol{A}^\top \\ \boldsymbol{T}^\top \end{array} \right] \boldsymbol{M} \left[\boldsymbol{A}, \boldsymbol{T}\right] = \left[\begin{array}{cc} \boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A} & \boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{T} \\ \boldsymbol{T}^\top \boldsymbol{M} \boldsymbol{A} & \boldsymbol{T}^\top \boldsymbol{M} \boldsymbol{T} \end{array} \right], \tag{4.44}$$

which is invertible only when $[\boldsymbol{A}, \boldsymbol{T}]$ is full column-rank. This may not be the case for three different reasons, leading to three possible types of degeneracies. We discuss each of them in turn in the following three paragraphs. Obviously, in practice, as all the operations have to be performed numerically, what really matters is not the strict linear independence in the mathematical sense but rather linear independence sufficient to ensure stable and robust, finite-precision numerical calculations, as exemplified by Eqs. (4.29), (4.30), (4.13), (4.31) and (4.32).

In general, given a matrix $\boldsymbol{B}$ and a vector $\boldsymbol{z}$ of laying in its range, if $\boldsymbol{B}$ is singular then we can solve equation $\boldsymbol{B}\boldsymbol{x} = \boldsymbol{z}$ for $\boldsymbol{x}$ only down to an unknown contribution from the nullspace of $\boldsymbol{B}$. Typically, the component of the solution parallel to the nullspace will be arbitrarily set to zero and its true value unavoidably lost. This in practice could be obtained with help of a regularization of the matrix and involves first calculating its inverse via computing and inverting its eigenvalues. The regularization is applied to all eigenvalues that are smaller than some predefined threshold and as the corresponding eigenvalues of the inverse are set directly to zero.

4.2.1 The columns of $\boldsymbol{A}$ are not independent

This degeneracy could affect as much the standard as the extended map-making, but we include it here for completeness. In such a case, the scanning strategy does not allow for the reconstruction of some sky mode. A typical example is a polarization pixel that was not observed with sufficient redundancy in the polarization angle. This case is easy to avoid because $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$ is easy to build. If $\boldsymbol{M}$ is diagonal the cost is $\mathcal{O}(\mathcal{N}_t)$ operations and $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$ is block diagonal (a block for each pixel), its eigendecomposition ($\mathcal{O}(\mathcal{N}_p)$ operations) enables to evaluate the condition number of each of the diagonal blocks. After a pixel selection based on these condition numbers (and removal from the TOD of the corresponding samples) the new $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$ can be safely inverted.

4.2.2   The columns of T are not independent

This reflects the fact that there is redundancy in the templates. This case basically corresponds to an attempt to filter the same template twice. For example, this happens in practice when two different sets of templates, e.g. the polynomial and the ground template filters, both remove the global offset of the TOD.

Since the final goal is to estimate $\hat{s}$ and not $\hat{y}$, Eqs. (4.29) and (4.30), this degeneracy does not pose any fundamental problem. We merely need to construct a basis of span(T) and use it to define a new (smaller) set of independent templates $\hat{T}$ as it is described in Sec. 4.1.3.

In practice, the situation is also quite straightforward. By construction, there are usually many known orthogonality relations between the templates. As a consequence, $T^\top M T$ is typically reasonably easy to compute and is sparse and structured. Its inverse can be then computed explicitly using standard matrix inversion techniques. The condition number of this matrix provides an easy test of the linear independence of the templates. If it is too high the matrix has to be regularized while being inverted. Once such a regularized inverse, $(T^\top M T)^{-1}$, is computed it should be used in the projector of Eqs. (4.29) and (4.32), taking care of the redundancies and therefore the degeneracies.

4.2.3   Some columns of A are not independent from the columns of T

This is the most insidious type of possible degeneracy of the map-making problem. Its presence highlights existing degeneracy between the sky signal and the templates. As a consequence the reconstruction of some sky component is not possible if the templates have been filtered. As a trivial example, by systematically filtering the mean of the total intensity TOD we create a degeneracy with the global offset of the temperature map.

This type of the degeneracy manifests itself as a singularity of both $A^\top F_T A$ and $T^\top F_A T$. This can be seen immediately by noting that, if the columns of A and T are not independent, then there exits at least two modes, one in the map domain, $\tilde{s}$, and one in the template domain, $\tilde{y}$, such as,

$$A\tilde{s} = T\tilde{y}, \tag{4.45}$$

and therefore

$$A^\top F_T A\tilde{s} = A^\top F_T T\tilde{y} = 0, \tag{4.46}$$
$$T^\top F_A T\tilde{y} = T^\top F_A A\tilde{s} = 0, \tag{4.47}$$

given that $F_T T = F_A A \equiv 0$, (4.21). The two modes, $\tilde{s}$ and $\tilde{y}$, constitute a pair of degenerate modes, residing in different domains but leading to the same effects in the time domain and therefore indistinguishable from each other.

Consequently, the very best one can do to solve the map-making problem in such cases is to regularize the inversion of the singular matrix and compute all the modes of the map for which the solution can be obtained and determine those modes for which this can not be done. These last modes will be missing from the reconstructed map. We note that this is not due to the regularization procedure but because these modes are removed from the data by the filters. Indeed,

$$A F_T d = A F_T A s + A F_T n = A F_T A s_\perp + A F_T n, \tag{4.48}$$

where the subscript $\perp$ denotes the part of the sky signal orthogonal to the nullspace of $A F_T A$, as the part parallel to it is irrecoverably lost. The information about these removed modes needs then to be propagated to next steps of the analysis and properly taken into account to ensure that the final results are statistically meaningful.

This route is only straightforward in practice if all the matrices appearing Eq. (4.44), can be constructed and decomposed explicitly. This however may be a formidable task, because of their sizes and complexity, and often unfeasible even if the help of the largest massively parallel platforms and parallel software packages is enlisted. If this is indeed the case and the solution can be only derived using some iterative technique, then the singular modes may not only be impossible to compute and correct for, but indeed it may not be clear from the onset whether the matrices are singular or not. In such cases this may need to be inferred from the solver behavior after the computation. We discuss these issues in the context of POLARBEAR map-making in Sec. 6.6.2.

We also emphasize that in the presence of this kind of degeneracies the maps computed by the direct and two-step approaches may not be identical. Indeed, in the direct case the unconstrained sky modes

will be missing in the estimated map. In the two-step case the situation is different as the regularized inversion has to happen when the estimate of the template amplitudes is performed. Consequently, these will be degenerate "modes" of the templates, which will be missing in $\hat{\boldsymbol{y}}$, while the template-corrected data vector, $\boldsymbol{d} - \boldsymbol{T}\hat{\boldsymbol{y}}$, will retain a time domain component that should have been filtered out. Hence, the map estimated on the second step, (4.31), will contain some of the degenerate sky modes, which will be set automatically to zero in the direct approach. Obviously, if the fact that these modes cannot be estimated is properly taken into account in the maps covariance, both maps will be statistically fully equivalent. We also note that to compute what map modes are non-constrainable, one may use the singular modes of $\boldsymbol{T}^\top \boldsymbol{F_A} \boldsymbol{T}$ found, while performing the first step of the two-step method and use them on the second step, by replacing the template-corrected data vector, $\boldsymbol{d} - \boldsymbol{T}\hat{\boldsymbol{y}}$ by $\boldsymbol{T}\boldsymbol{z}$, where $\boldsymbol{z}$ stands for one of the singular eigenvectors. Map-domain templates resulting from this calculation will have to be then orthogonalized using, e.g., the Gram-Schmidt procedure.

In the case of the biased map estimator, (4.32), this degeneracy does not pose any numerical issue. As in the unbiased cases the map estimate will have no contribution from the sky signal modes residing in the nullspace of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$, due to the filtering applied on the time-domain level, (4.48). However, these modes may, and usually will, appear in the biased map itself. Nevertheless, their amplitude will not reflect that of the true, underlying sky signal.

4.2.4 Case of ground-based experiments

As a specific application of the above formalism let us consider a map-making problem for a modern, ground-based, CMB experiment, which scans the sky with a kilo-pixel array of polarization sensitive detectors in the presence of both atmospheric and ground emissions. Commonly, in order to minimize the impact of the atmospheric contributions, the scans are performed at constant elevations for relatively short periods of time. The elevation is then changed in order to catch up with the sky patch, which moves away due to the Earth rotation. The scan amplitude, specific choices of the scan elevations, as well as elevation dwelling time are specific to a specific observation. In the following, we consider only cases, which conform to this general description.

The data of such an experiment are typically more complex than the simple model exemplified by (4.4), be it due to the atmospheric, instrumental, or ground contaminations. Below we describe common contributions of this type and explain how they fit within the extended data model, (4.27), and how they can be treated with the map-making technique introduced earlier.

Noise correlations

The noise is usually correlated between different detectors and typically displays significant low-frequency excess, dubbed typically as $1/f$ noise, which arises either due to instrumental reasons, the emission of the atmosphere or some other effects. As a result, the time domain noise correlation matrix, $\boldsymbol{C}_n$, is dense, making its inversion and multiplication computationally demanding. Moreover, the matrix, $\boldsymbol{C}_n$, is unknown and has to be estimated from the data themselves (e.g., Ferreira and Jaffe, 2000). This procedure is usually difficult, especially as far as the low-frequency modes and detector-detector correlated modes are concerned. In addition, the low-frequency contributions may not be even stationary or Gaussian, and thus cannot be properly described merely by a covariance. Consequently, using the optimal weight, $\boldsymbol{C}_n^{-1}$, in the map-making process may be difficult. Using a simpler weight matrix, $\boldsymbol{M}$, does not lead to a bias in the standard map estimator, Eq. (4.5), however its choice does affect the quality of the final map. This is because for different weights, the time-domain data, $\boldsymbol{d}$, are coadded differently on the first step of the map-making procedure, when the so-called right hand side of the map-making equation, i.e., $\boldsymbol{A}^\top \boldsymbol{M}\boldsymbol{d}$, is calculated. For instance, the diagonal weights (in map domain) cannot suppress selectively some temporal frequency bands over the others. Consequently, if the $1/f$ noise is present, even if it is Gaussian and stationary, the low-frequency modes will not be properly down-weighted as compared to the high-frequency ones. This may result in striped features appearing in the direction of the scans (see, e.g., Maino et al., 2002). The effect is particularly apparent if the scanning strategy does not provide a good cross-linking. Non-Gaussian/non-stationary features can be even more difficult to deal with.

Instead of down-weighting such long modes, one may prefer to filter them out as it is done in the filtered map-making, Eq. (4.22). The time-domain data are then explicitly filtered, when compressed to the pixel-domain object, $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d}$. The long-term trends present in the time-ordered data can be removed at this stage. Such removal is blind to the origin and nature of the trends, potentially removing the true sky signal together with the noise and some unwanted spurious contributions. However, the

signal-to-noise ratio for these modes is usually very low, as the $1/f$ noise quickly dominates, and the information loss due to the potential removal of the sky signal is typically negligible. If this is the case, the filtering can provide a useful alternative to weighting in this context.

The modes to be filtered out are typically assumed to be arbitrary linear combinations of some sufficiently rich family of temporal templates, which has to be defined case-by-case. For the long-term trends, the templates are often taken as piece-wise low order polynomials or harmonic functions defined for all samples of the data and are represented by columns of some template matrix, $\boldsymbol{B}$. If the templates are well matched to the problem at hand, then the residual noise, $\boldsymbol{w}$, defined as,

$$w \equiv n - Bx, \tag{4.49}$$

is expected to be 'whiter' than the actual time domain noise, $\boldsymbol{n}$, and thus potentially, sufficiently well, albeit approximately, characterized by a diagonal noise covariance matrix. Here, $\boldsymbol{x}$ stands for a set of a priori arbitrary parameters to be determined, similar to the sky signal, $\boldsymbol{s}$, and, which define amplitudes of the corresponding templates (Stompor et al., 2001b).

We point out that in general one could introduce some additional information about the long-term modes of the noise by setting some constraints on $\boldsymbol{x}$. As such constraints are typically hard to identify and might not lead to a significant improvement in the sensitivity, we will not consider them in this work (see, e.g., Keihänen et al., 2010, for an implementation of this idea in the context of the two-step map-making). As we will see below, in their absence, the map-making will discard all the information in the data that matches the time domain signature of the templates, effectively filtering all template-like modes out of the data. As previously explained, this causes a loss of a sky mode, $\tilde{\boldsymbol{s}}$, only if its time domain projection $\boldsymbol{A}\tilde{\boldsymbol{s}}$ is a template-like mode. In general, subsequent observations brake the degeneracy that a template might have on a subset of the dataset. There are two well-known exceptions. First, the global offset of the time streams is always filtered, as a consequence the global offset of the temperature map is unconstrained. Second, when observing from the Earth's pole, the scans at constant elevation always probe constant declination stripes, with different elevations corresponding to different declinations. Filtering the offset from the time streams of each of these scans prevents the reconstruction of the offset of each of these constant declination stripes. These can be partially recovered if the stripes share some of the sky pixels owing to the assumption that the sky signal is constant across a pixel. The resulting constraints would be typically weak, in particular for the relative offset of two stripes that are not directly adjacent, which could be constrained only thanks to the intermediary ones. Consequently, one should expect in such cases poorly constrained long modes in declination.

Ground pickup

Ground-based experiments usually have non-negligible ground-synchronous signal contaminating their TOD. A possible source of this signal is the response of the detectors to the local magnetic field. Also emission from the ground can be registered by the instruments due to the far side lobes of the beam. These are designed to be very low but, as the ground is very bright by the CMB measurement standards, the contribution cannot be typically neglected. This ground-synchronous signal could be thought of as a two-dimensional template in Earth-bound coordinates. However, in practice, the situation is more complex as the signal can vary in time and may be different for different detectors, as they may have a different, and usually poorly known, level and structure of the side lobes. Therefore, typically one models the ground signal as a one-dimensional template specific to each constant elevation scan and to each detector, or at the very least a group of those located sufficiently close to each other in the focal plane.

Such a one-dimensional ground template can be parametrized with the azimuth of the observation and represented by one dimensional discretized map with entries standing for the amplitude of the ground signal in each of the disjoint, consecutive azimuth bins. Following our previous argument, we need to introduce such a template for each detector and each constant elevation scan separately. These concatenated together are then denoted as $\boldsymbol{g}$. The time domain data, in the presence of the ground pickup, can be the modelled as a sum of three terms the sky signal term, $\boldsymbol{As}$, the ground-pickup term, $\boldsymbol{Gg}$, and the noise, $\boldsymbol{n}$, assuming, as we do for the sake of this section, that no other contributions are present. We can write then,

$$d = As + Gg + n. \tag{4.50}$$

This is merely a specialized version of (4.27). The role played by the matrix $\boldsymbol{G}$ is analogous to that played by the sky pointing matrix, $\boldsymbol{A}$. Adopting the simple binned ground template model introduced
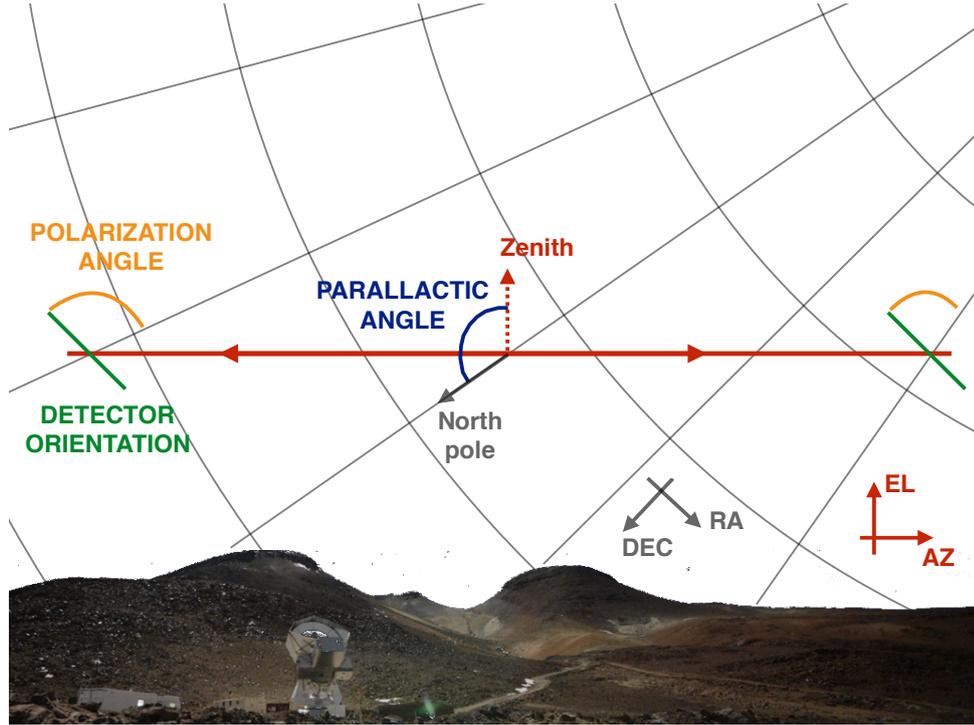
Figure 4.2: Geometry of a constant elevation observation. Gray lines represent the equatorial coordinate system at a fixed time instant. The red line represents the scan at constant elevation: it is the (moving) pointing direction in the ground-bound horizontal coordinate system. The orientation of the detector projected on the sky (green), assumed constant in horizontal coordinates, varies from left to right because the parallactic angle (blue) also changes from left to right.

above, each column of $\boldsymbol{G}$ is associated with some specific azimuthal bin assigned to some specific detector and some constant elevation scan. This column will be composed of ones and zeros, with 1 indicating that the given measurement was made by the specific detector and was performed within the specific scan, when the observations azimuth fallen with the specific azimuthal bin. Therefore applying $\boldsymbol{G}$ to the template, $\boldsymbol{g}$, will add the same value of the ground pick up to all these measurements. In this section we summarize the effects the modeling of the ground pick-up may have on the quality of the map recovered with the unbiased map estimator, leaving a more detailed discussion to Appendix B.

We start by considering a site far from the Earth poles. For a single constant elevation scan, the samples having the same ground pick will correspond to the measurements taken by a single detector with the azimuthal position falling into one of the azimuthal bins. As time progresses and the sky rotates in the Earth-bound coordinates (see Fig. 4.2 and 4.3 for a sketch of the geometry of the problem) the measurements will cover the sky area extending along the RA direction in the sky coordinates. The size in declination of the area depends on the scan elevation and the azimuth of the bin. The sky areas covered by measurements corresponding to two different azimuthal bins will be disjoint and as both these subsets are affected by a different ground pickup amplitude, their relative offsets will be unconstrained. As a consequence, a single detector map will have multiple degenerate modes, each corresponding to a sky patch swept by a different azimuthal bin. The degeneracies can be in part removed if data of another detector are used, but only if the azimuthal bins of the latter are shifted with respect to those of the first detector in such a way that their corresponding sky patches are also shifted and each patch of the second detector includes pixels from two adjacent patches of the first one.

Yet another factor breaking the degeneracy here is the sky pixelization, as the sky pixels crossing the boundary of two adjacent bin sky patches will constrain their relative amplitude. However, in all these cases the breaking may be quite weak because the overlaps typically involve only a limited number of neighboring sky patches corresponding to single azimuthal bins.

For the observations taken from the poles the situation is different. Since the sky movement is in the azimuthal direction the same sky can be measured in different azimuthal bins. Consequently, there is a significant constraining power on the relative offsets, leaving the overall offset of the observed constant
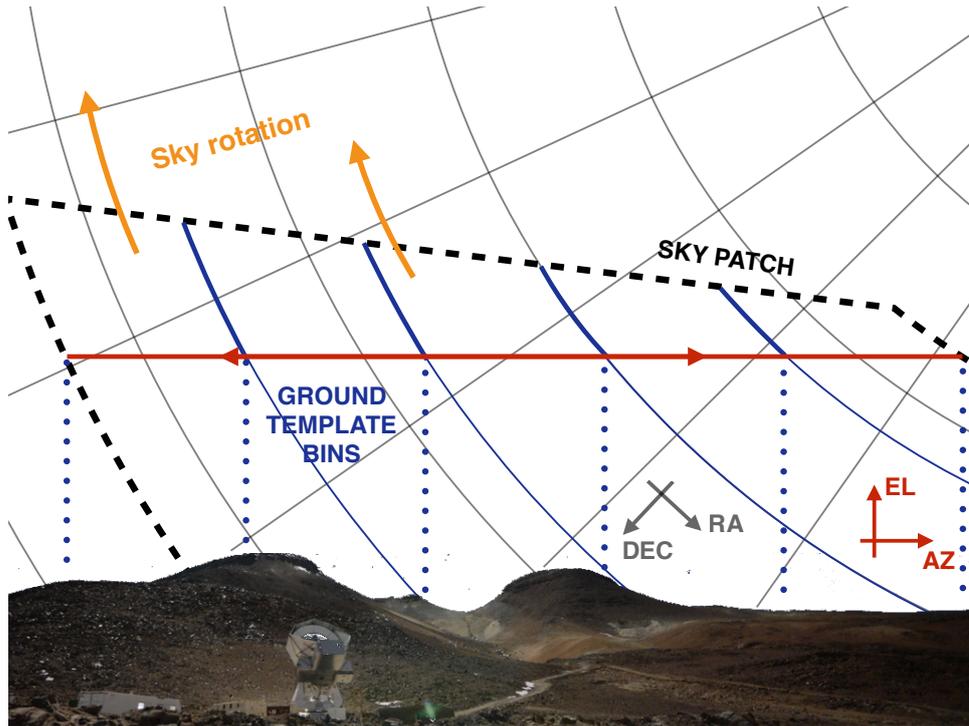
Figure 4.3: The picture illustrates two aspects of the interplay of constant elevation observation and sky rotation. First, while in horizontal coordinates the observation consists of just one line, because of the sky rotation an entire area is observed in equatorial coordinates: it is the target patch. The area above the red line has been already observed, while the part below the red line is still to be observed. Second, disjoint azimuthal segments are in correspondence with disjoints regions in equatorial coordinates. Notice that the declination interval of the leftmost region is larger than the one of the rightmost, because the parallactic angle gets close to π. Filtering the constant ground signal inside an azimuthal bin also filters the offset of the corresponding sky region. As explained in the text, this holds for both temperature and polarization, if the polarization angle does not vary appreciably in the azimuthal range of the bin.

declination stripe as the only truly degenerate mode. We already encountered this degenerate mode in Sec. 4.2.4: the offset of the TOD is filtered also when removing correlated noise. The consequences of this degenerate mode and the possible degeneracy braking effects were already discussed at the end of Sec. 4.2.4.

Recap

The following useful conclusions have been drawn out in this section (and are borne out further by a more detailed analysis in Appendix B). In the presence of time domain filtering of the kinds discussed here, only relatively few sky modes are expected to be genuinely degenerate. Nevertheless, a few additional ill-conditioned modes with large variance should also be expected. They are mainly related to the filtering of the ground signal, which cause poorly constrained modes in the declination direction. The details of these modes will depend on choices made regarding pixelization, definition of the ground template bins and observational strategy but they will be more prominent for observations conducted from the poles. Given this, the biased map-making, by construction oblivious to the presence or absence of such modes, may be seen as providing a more convenient and adaptable way to perform the analysis. Indeed, it has been the approach of choice for multiple past analysis of this kind of data sets (e.g., Culverhouse et al., 2010; Schaffer et al., 2011; BICEP2 Collaboration, 2014). However, a relevant question, which we discuss in more detail in Chapter 6, is whether it is feasible at all for a ground-based experiment to produce an unbiased estimate of the sky signal and with what fidelity.

4.3 Power spectrum estimation

In Sec. 3.2.2 we introduced the power spectrum as the harmonic counterpart of the two-point correlation function of a statistically isotropic field. It is particularly important for Gaussian fields because it fully characterises the statistical properties of the field. In Eq. (3.36) we introduced the simplest example of power spectrum estimator. However, it is not suitable for the power spectrum estimation of real observation because it assumes the true CMB a_{lm}s and thus perfect knowledge CMB over the full sky.

The power spectrum estimation can be formulated as a maximum likelihood problem (e.g. Bond et al., 1997). Suppose that the available sky map is

$$\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{n} \tag{4.51}$$

where $\boldsymbol{s}$ is the actual sky signal and $\boldsymbol{n}$ is the noise contribution, we denote its covariance matrix as $\boldsymbol{N}$. Consistently with Eq. (3.33), we call the covariance matrix of the signal $\boldsymbol{C}$.

Assuming signal and noise to be uncorrelated, the probability of a realisation of $\boldsymbol{x}$ and likelihood function for the power spectrum is (e.g. Tegmark, 1996)

$$P(\boldsymbol{x}|C_\ell) = \mathcal{L}(C_\ell) = \frac{\exp\{\frac{1}{2}\boldsymbol{x}^\top (\boldsymbol{N} + \boldsymbol{C})^{-1}\boldsymbol{x}\}}{\sqrt{(2\pi)^{\mathcal{N}_p} \det(\boldsymbol{N} + \boldsymbol{C})}} \tag{4.52}$$

where the dependence on the power spectrum C_ℓ is inside $\boldsymbol{C}$, see Eq. (3.34). Evaluating this likelihood function is expensive because it involves the inversion of the covariance matrix, which is different at each evaluation. Estimating C_ℓ though likelihood maximisation therefore requires the number of sky pixels to be limited or, equivalently, only low multipoles can be probed since $\mathcal{N}_p \sim \ell_{\max}^2 f_{\rm sky}$. Maximum likelihood estimation is indeed only used for probing the low multipoles of the power spectrum, for example $\ell < 40$ in Hinshaw et al. (1996) and $\ell < 30$ in Planck Collaboration (2015d).

4.3.1 Pseudo-power spectrum estimation

The cost of the maximum likelihood power spectrum estimation scales as $\ell_{\max}^6$ while the one of a spherical harmonics transform and, consequently, of the idealistic power spectrum estimator in Eq. (3.36) is typically of order $\ell_{\max}^3$. The pseudo-power spectrum methods fill this gap in the scaling, providing a method for estimating power spectra under realistic assumption at a $\ell_{\max}^3$ computational cost (but with a large pre-factor compared to the naive spectrum estimator). They were first introduced in the context of galaxy surveys (Hauser and Peebles, 1973) and then applied to the CMB (Hivon et al., 2002; Kogut et al., 2003). In the next paragraph we explain the basic idea using temperature while two different ways of including polarization are presented in dedicated subsections. More details can be found in e.g. Brown et al. (2005).

We first outline the key idea of the pseudo-power spectrum estimators by considering a signal only reconstructed map, $\boldsymbol{x} = \boldsymbol{s}$. For each ℓ multipole there is an associated subspace of functions: the one generated by the set of spherical harmonics $\{Y_{lm}\}_{\ell \in [-\ell, \ell]}$. The estimator Eq. (3.36) first evaluates the scalar product of the CMB sky with each basis functions, $a_{\ell m} = \int Y_{\ell m}^* s \mathrm{d}\Omega$, and then sums their squared absolute values. Similarly, the pseudo-power spectrum estimators first define a pseudo basis, $\tilde{Y}_{\ell m}$, then compute the pseudo-multipoles

$$\tilde{a}_{\ell m} \equiv \int \tilde{Y}_{\ell m}^* s \mathrm{d}\Omega \tag{4.53}$$

and the pseudo-power spectrum

$$\tilde{C}_\ell = \sum_{m=-\ell}^{\ell} \frac{\tilde{a}_{\ell m}^* \tilde{a}_{\ell m}}{2\ell + 1}. \tag{4.54}$$

Finally, they estimate the power spectrum from the pseudo-power spectrum. Since

$$\tilde{a}_{\ell m} = \int \tilde{Y}_{\ell m}^* \sum_{\ell' m'} a_{\ell' m'} Y_{\ell' m'} \mathrm{d}\Omega = \sum_{\ell' m'} K_{\ell m \ell' m'}^0 a_{\ell' m'}, \tag{4.55}$$

where we have defined the *convolution kernel*,

$$K_{\ell m \ell' m'}^0 \equiv \int \tilde{Y}_{\ell m}^* Y_{\ell' m'} \mathrm{d}\Omega, \tag{4.56}$$

the pseudo-power spectrum is

$$\tilde{C}_\ell = \sum_{m=-\ell}^{\ell} \frac{\tilde{a}_{\ell m}^* \tilde{a}_{\ell m}}{2\ell + 1} \tag{4.57}$$

$$= \sum_{m=-\ell}^{\ell} \frac{1}{2\ell + 1} \sum_{\ell' m' \ell'' m''} a_{\ell' m'}^* a_{\ell'' m''} K_{\ell m \ell' m'}^0 (K_{\ell m \ell'' m''}^0)^* \tag{4.58}$$

By using the power spectrum we are implicitly assuming statistical isotropy, consequently $\langle a_{\ell' m'}^* a_{\ell'' m''} \rangle = C_{\ell'} \delta_{\ell' \ell''} \delta_{m' m''}$ and the expected value of the previous equation simplifies to

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell \ell'} C_{\ell'} \tag{4.59}$$

where the *mixing kernel* is defined as

$$M_{\ell \ell'} \equiv \sum_{m=-\ell}^{\ell} \sum_{\ell'} \sum_{m'=-\ell'}^{\ell'} \frac{2\ell' + 1}{2\ell + 1} K_{\ell m \ell' m'}^0 (K_{\ell m \ell' m'}^0)^*. \tag{4.60}$$

The mixing kernel is then inverted and the power spectrum estimated.

One way of understanding the pseudo-power spectrum estimation is the following. The actual realisation of the sky has a component in the ℓ subspace, the vector subspace of the functions on the sphere spanned by $\{Y_{\ell m}\}_{m \in [-\ell, \ell]}$. The assumption of statistical isotropy means that, on average, the power of this component is the same, C_ℓ, in all the directions of the ℓ subspace. Therefore, we can measure this power as the average power over some set of functions belonging to the ℓ subspace. In Eq. (3.36) this functions are the $\{Y_{\ell m}\}_{m \in [-\ell, \ell]}$, an orthonormal basis of the ℓ subspace. For this choice, the evaluation of the sky power on each function is independent and the information in the sky realisation is fully exploited. If we give up on these attractive properties, the set of functions can be completely arbitrary. In the formalism described above the set of functions, called pseudo-basis, is even more general. As the actual basis, the functions of the pseudo-basis are grouped in subsets indexed by ℓ and $m \in [-\ell, \ell]$, but this does not require the functions to belong to the respective ℓ subspace: it is just an arbitrary indexing choice related to the way these functions are computed in practice. Any given function of the pseudo-basis, $\tilde{Y}_{\ell m}$, can be decomposed in the components belonging to the different ℓ' subspaces. The sky power evaluated on the function is the sum of the power of each of these components. On average, the power along these components is given by the power spectrum $C_{\ell'}$ and therefore the expected value of the whole power along $\tilde{Y}_{\ell m}$ is a known linear combination of the true power spectrum. By doing this study on the $\tilde{Y}_{\ell m}$ for all the ms, we are able to say how much power (on average) $\tilde{C}_\ell$ gets from each ℓ' multipole of the true power spectrum $C_{\ell'}$, as expressed in Eq. (4.59). Typically, the mixing matrix has diagonal structure just approximately, actually coupling neighbouring multipoles, but the fact that each ℓ of the pseudo-power spectrum is a mixture of possibly all the ℓ of the actual power spectrum does not pose any problem to the estimation of the latter–as long as the mixing matrix is invertible. However, suppose that C_ℓ contributes to $\tilde{C}_{\ell'}$ with $\ell' \neq \ell$, if $C_{\ell'}$ has a cosmic variance much bigger than C_ℓ, the former might dominate the uncertainty budget of $\tilde{C}_\ell$ and, consequently, of the estimation of C_ℓ. (Also) for this reason, both pseudo and actual power spectrum are multiplied by $\ell(\ell + 1)/(2\pi)$ because

$$\mathcal{D}_\ell \equiv \ell(\ell + 1)C_\ell/(2\pi) \tag{4.61}$$

is roughly constant at large scales–the ones affected the most by the cosmic variance–and therefore its cosmic variance is proportional to a gentle $\ell^{-1/2}$.

Now that the main idea of the pseudo-power spectrum estimators is explained, we illustrate how the formalism easily accommodates the presence of noise in the maps, their convolution with the experimental beam as well as their partial sky coverage. The pseudo-basis is usually built as

$$\tilde{Y}_{\ell m} \equiv W Y_{\ell m} \tag{4.62}$$

where the *window function* W is a function on the sphere that is different from zero only inside the observed region. Similarly to how sharp features in a function $f : \mathbb{R} \to \mathbb{R}$ cause ringing in its Fourier transform, a smooth apodisation of the window function mitigates the off-diagonal correlations in the mixing matrix. Moreover, if the window function is somewhat inversely proportional to the noise variance

map, the definition of the pseudo-basis in Eq. (4.62) naturally downweights the noise pixels when computing the pseudo-power spectrum. Taking into account the noise and the beam of the instrument, the multipole of the observed map are

$$\int Y_{\ell m}^* x = B_\ell a_{\ell m} + n_{\ell m} \tag{4.63}$$

where the B_ℓ function models the effect of the beam. For a Gaussian beam $B_\ell = e^{\ell^2 \sigma_b^2}$, where σ_b is related to the FWHM by FWHM $= 2\sqrt{2 \ln 2} \sigma_b \sim 2.4 \sigma_b$. So far we have assumed that the mixing matrix is invertible. However, this is typically not the case, the limited observed sky area prevents the reconstruction of the lowest multipoles, for example. Therefore, the spectrum is usually binned, with the mutipoles bins spanning a sufficiently large ℓ interval. As mentioned earlier the use of the pseudo-basis also correlates neighbouring multipoles and thus the size of the bins is usually chosen to be large enough so that the measurements of the different bins are sufficiently uncorrelated. The binning can be represented as a binning operator that averages over $\Delta \ell_b$, the multipole range of the bin b,

$$C_b \equiv \sum_\ell P_{b\ell} C_\ell \equiv \sum_{\ell \in b} \frac{C_\ell}{\Delta \ell_b} \tag{4.64}$$

If we work with $\mathcal{D}_\ell$s we average an almost constant quantity, rather then the steeply decreasing C_ℓ. This is another reason why the former is typically preferred. Eq. (4.59) now reads

$$\langle \tilde{\mathcal{D}}_b \rangle = \sum_{b'} M_{bb'} \mathcal{D}_{b'} + \langle \tilde{\mathcal{D}}_b^N \rangle. \tag{4.65}$$

where the binned mixing matrix reads

$$M_{bb'} = \sum_{\ell \ell'} P_{b\ell} M_{\ell \ell'} B_{\ell'}^2 P_{b'\ell'}. \tag{4.66}$$

The second term, the noise bias, is the binned pseudo-$\mathcal{D}_\ell$ of the noise:

$$\tilde{\mathcal{D}}_b^N \equiv \sum_{\ell \in b} \frac{\ell(\ell+1)}{2\pi \Delta \ell_b} \sum_m \frac{\tilde{n}_{\ell m}^* \tilde{n}_{\ell m}}{2\ell + 1} \tag{4.67}$$

The power spectrum estimator can thus be estimated as

$$\hat{\mathcal{D}}_b = \sum_{b'} M_{bb'}^{-1} \left(\tilde{D}_{b'} - \langle \tilde{\mathcal{D}}_{b'}^N \rangle \right). \tag{4.68}$$

The estimation of the expected value of the noise bias is usually a difficult task. For this reason, many experiments adopt mainly cross-spectra: the pseudo-cross-spectrum of two maps containing the same sky but uncorrelated noise is noise bias-free (see, e.g., Hinshaw et al., 2003; Tristram et al., 2005).

Polarization: standard estimator

In analogy with the temperature case, the formalism can be extended to polarization introducing a spin 2 pseudo-basis,

$$_{\pm 2}\tilde{Y}_{\ell m} \equiv W _{\pm 2}Y_{\ell m}. \tag{4.69}$$

Given Eq. (3.20), we define the E and B pseudo-multipoles,

$$_E\tilde{a}_{lm} = -\frac{1}{2} \int {}_{+2}\tilde{Y}_{\ell m}^*(Q - iU) + {}_{-2}\tilde{Y}_{\ell m}^*(Q + iU) \tag{4.70}$$

$$_B\tilde{a}_{lm} = \frac{i}{2} \int {}_{+2}\tilde{Y}_{\ell m}^*(Q - iU) - {}_{-2}\tilde{Y}_{\ell m}^*(Q + iU). \tag{4.71}$$

Writing the polarization field as $Q \mp iU = \sum_{\ell' m'} ({}_E a_{\ell' m'} \pm i {}_B a_{\ell' m'}) {}_{\pm 2}Y_{\ell' m'}$ and defining

$$K_{\ell m \ell' m'}^{\pm} \equiv \frac{1}{2} \left(\int {}_{+2}\tilde{Y}_{\ell m}^* {}_{+2}Y_{\ell' m'} d\Omega \pm \int {}_{-2}\tilde{Y}_{\ell m}^* {}_{-2}Y_{\ell' m'} d\Omega \right), \tag{4.72}$$

we can rewrite the E and B pseudo-multipoles in terms of the E and B multipoles

$$_E\tilde{a}_{lm} = \sum_{\ell'm'} \left(K^+_{\ell m\ell'm'} {}_E a_{lm} + K^-_{\ell m\ell'm'} {}_B a_{lm} \right) \tag{4.73}$$

$$_B\tilde{a}_{lm} = \sum_{\ell'm'} \left(-K^-_{\ell m\ell'm'} {}_E a_{lm} + K^+_{\ell m\ell'm'} {}_B a_{lm} \right). \tag{4.74}$$

Defining the pseudo-power spectrum as

$$\tilde{C}^{XY}_\ell \equiv \sum_{m=-\ell}^{\ell} \frac{_X\tilde{a}^*_{\ell m} {}_Y\tilde{a}_{\ell m}}{2\ell+1}, \tag{4.75}$$

its expected value is related to the power spectrum by the mixing kernel

$$\left\langle \begin{pmatrix} \tilde{C}^{TT}_\ell \\ \tilde{C}^{TE}_\ell \\ \tilde{C}^{TB}_\ell \\ \tilde{C}^{EE}_\ell \\ \tilde{C}^{EB}_\ell \\ \tilde{C}^{BB}_\ell \end{pmatrix} \right\rangle = \begin{pmatrix} M^{00}_{\ell\ell'} & & & & & \\ & M^{0+}_{\ell\ell'} & M^{0-}_{\ell\ell'} & & & \\ & -M^{0-}_{\ell\ell'} & M^{0+}_{\ell\ell'} & & & \\ & & & M^{++}_{\ell\ell'} & (M^{-+}_{\ell\ell'} + M^{+-}_{\ell\ell'}) & M^{--}_{\ell\ell'} \\ & & & -M^{+-}_{\ell\ell'} & (M^{++}_{\ell\ell'} - M^{--}_{\ell\ell'}) & M^{-+}_{\ell\ell'} \\ & & & M^{--}_{\ell\ell'} & -(M^{-+}_{\ell\ell'} + M^{+-}_{\ell\ell'}) & M^{++}_{\ell\ell'} \end{pmatrix} \begin{pmatrix} C^{TT}_\ell \\ C^{TE}_\ell \\ C^{TB}_\ell \\ C^{EE}_\ell \\ C^{EB}_\ell \\ C^{BB}_\ell \end{pmatrix} \tag{4.76}$$

where the blocks of the mixing matrix are defined as

$$M^{YZ}_{\ell\ell'} = \sum_{m=-\ell}^{\ell} \sum_{\ell'} \sum_{m'=-\ell'}^{\ell'} K^Y_{\ell m\ell'm'} (K^Z_{\ell m\ell'm'})^*. \tag{4.77}$$

with $Y, Z \in \{0, +, -\}$.

The presence of many zero blocks is due to the fact that the pseudo basis does not mix temperature and polarization. However, the polarization pseudo-basis mixes E and B modes whenever the M^{--} block is not zero, causing a problem analogous to the mixing of power from different ℓs with very different C_ℓs. The mixing matrix quantifies which fraction of the E power leaks into the pseudo-B and effectively corrects for it when applying the inverted mixing matrix on the pseudo-power spectrum. However, the cosmic variance of EE severely affects the BB uncertainty, especially at large scales. This problem is overcome by using the pure estimators described in the next section.

Polarization: pure estimator

As said earlier, a polarization field $\boldsymbol{P}$ is type $E(B)$ if $\boldsymbol{D}^\dagger_B \cdot \boldsymbol{P} = 0$ ($\boldsymbol{D}^\dagger_E \cdot \boldsymbol{P} = 0$). A polarization field that satisfies both is said to be *ambiguous*. On the sphere (and actually on any manifold without boundary) if $\boldsymbol{D}^\dagger_B \cdot \boldsymbol{D}_B f = 0$ (or $\boldsymbol{D}^\dagger_E \cdot \boldsymbol{D}_E f = 0$) then the polarization field $\boldsymbol{D}_B f$ (or $\boldsymbol{D}_E f$) itself is zero, implying that the space of E and B modes are orthogonal. However, the observed sky region typically have boundaries. In this case there exist a set of non-trivial functions f that satisfy $\boldsymbol{D}^\dagger_B \cdot \boldsymbol{D}_B f = 0$ (or $\boldsymbol{D}^\dagger_E \cdot \boldsymbol{D}_E f = 0$) and thus are ambiguous modes. In these cases it is useful to define the *pure B modes* as those B modes that are orthogonal to any E mode, *pure E modes* are similarly defined. A polarization field can thus be unambiguously decomposed into pure E, pure B and ambiguous modes (Bunn et al., 2003). The power in the ambiguous modes cannot be unambiguously assigned to E or B. However, in the CMB case this power (and its related cosmic variance) is dominated by the E mode contribution. In order to prevent the E-type cosmic variance from affecting the B mode estimation, the pure-pseudo-power spectrum estimators adopt a pure B-mode pseudo-basis (Smith, 2006; Smith and Zaldarriaga, 2007).

Any E mode can be written as $\boldsymbol{D}_E f$, for some function f, therefore having a pure B pseudo-basis means that for any f and (ℓ, m)

$$\int_{4\pi} {}_B\tilde{\boldsymbol{Y}}^\dagger_{\ell,m} \cdot \boldsymbol{D}_E f = 0 \tag{4.78}$$

A natural choice is (c.f. Eq. (4.69))

$$_B\tilde{\boldsymbol{Y}}_{\ell m} \equiv \boldsymbol{D}_B(WY_{\ell m}), \tag{4.79}$$

which indeed fulfils Eq. (4.78),

$$\int_{4\pi} (_B\tilde{\boldsymbol{Y}}_{\ell m})^\dagger \cdot \boldsymbol{D}_E f = \int_{4\pi} [\boldsymbol{D}_B(WY_{\ell m}))]^\dagger \cdot \boldsymbol{D}_E f \tag{4.80}$$

$$= \int_{4\pi} WY_{\ell m} \boldsymbol{D}_B^\dagger \cdot \boldsymbol{D}_E f = 0. \tag{4.81}$$

The integration by partition in the second row does not produce any boundary term because the sphere has no boundaries while the last equality comes form Eq. (3.28). The second order differentiation in the definition (4.79) implicitly assumes some regularity conditions on W. In order to account for the unobserved region, W is equal to zero in the unobserved region and consequently at least the value at the boundary of both W and its derivative in the outward direction have to be zero (the Dirichlet and Neumann boundary conditions). Another way of understanding these condition is the following. Since the pseudo-basis is zero outside of the observed patch, all the integrals in this section can be restricted to the observed area. The only relevant consequence is that the integration by partition in the last equations produces boundary integrals containing either W or its derivative in the outward direction. The Neumann-Dirichlet conditions ensure that these terms are equal to zero.

Analogously, we can define the pure E pseudo-basis and refer to this case as the *pure formalism*. However, the B-to-E leakage does not pose problems for the estimation of the E spectrum and therefore the use of the pure pseudo-basis is not required. On the contrary, using the standard E pseudo-basis and the pure B pseudo-basis (*hybrid formalism*) exploits the information contained in the ambiguous modes for the E mode estimation while prevents the E-modes cosmic variance from affecting the B-modes estimation.

In general, the mixing matrix reads

$$M_{\ell\ell'} = \begin{pmatrix} M_{\ell\ell'}^{TT,TT} & & & \\ & M_{\ell\ell'}^{TE,TE} & M_{\ell\ell'}^{TE,TB} & \\ & M_{\ell\ell'}^{TB,TE} & M_{\ell\ell'}^{TB,TB} & \\ & & & M_{\ell\ell'}^{EE,EE} & M_{\ell\ell'}^{EE,EB} & M_{\ell\ell'}^{EE,BB} \\ & & & M_{\ell\ell'}^{EB,EE} & M_{\ell\ell'}^{EB,EB} & M_{\ell\ell'}^{EB,BB} \\ & & & M_{\ell\ell'}^{BB,EE} & M_{\ell\ell'}^{BB,EB} & M_{\ell\ell'}^{BB,BB} \end{pmatrix}. \tag{4.82}$$

Without getting too much into details, in the pure formalism the $M^{EE,BB}$ and $M^{BB,EE}$ are zero by construction and therefore the EE power is prevented from contributing to the BB pseudo-power and vice versa. In the hybrid formalism $M^{BB,EE}$ is still zero but $M^{EE,BB}$ can be different from zero.

4.3.2 Computing the spectrum of biased maps

The previous section assumed that the map available $\boldsymbol{x}$ is an unbiased rendition of the sky signal. When it is not the case the formalism of the pseudo-power spectrum estimators has to be amended. In this section we will illustrate an approximate quick method called Monte Carlo Apodized Spherical Transform Estimator (MASTER), very popular in current CMB B-mode experiments.

The most common bias in the reconstructed maps comes from the use of the biased estimator in Eq. (4.32). The map is thus related to the unbiased estimate by some linear relation $\boldsymbol{x}^{\text{biased}} = \boldsymbol{R}\boldsymbol{x}^{\text{unbiased}}$. The bias is due to the fact that the TOD are filtered but the filtering is not taken into account when estimating the map. The filtering is (or at least involve) a linear operation diagonal in the frequency domain, e.g. high pass filtering. Since the observations consist of constant velocity scans, there is a qualitative correspondence between frequencies in the time domain and angles (and thus ℓ) on the sky. Following this heuristic argument, the bias is modelled as a systematic multiplicative bias in the ℓ-space, F_ℓ, completely analogous to the one of the beam. The relation between the CMB power spectrum and the pseudo-power spectrum therefore has an extra term

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell\ell'} M_{\ell\ell'} F_{\ell'} B_{\ell'}^2 C_\ell + \langle \tilde{N}_\ell \rangle, \tag{4.83}$$

where $\tilde{N}_\ell$ is the pseudo-power spectrum of the noise bias. The *transfer function* F_ℓ are evaluated using Monte Carlo simulations. Noise-free simulations of the CMB are performed and a TOD is generated using the actual scanning strategy of the data set being analysed. A map is produced using the same data processing of the real data and its pseudo-power spectrum is evaluated. The average of these spectra

over the set of Monte Carlos is then plugged in Eq. (4.83) and F_ℓ can be estimated either directly, if $M_{\ell,\ell'}$ is invertible, of iteratively (Hivon et al., 2002). Once the transfer functions are computed the estimation of the power spectrum of the real data can be performed. Since Eq. (4.83) is an ansatz, the procedure assumes that the estimation of the transfer functions does not depend on the sky signal and, therefore, the cosmology assumed in the simulations does not bias the power spectrum estimation of the real data. Hivon et al. (2002) have shown that, in the specific case they studied, the dependence of their transfer functions on the assumed cosmology is very weak.

4.4   Summary

The standard map-making problem can be extended to include and account for time-domain filtering in statistically sound manner. The general framework, introduced in this chapter, includes as special cases some of the most common map-making approaches and thus sheds new light on them. This formalism is a keystone of one of the POLARBEAR map-making pipelines as it is elaborated on in Chapter 6.

Chapter 5

The Polarbear experiment

5.1 The project

In the 2000s the scientific potential of the B modes of CMB polarization was well understood and recognised (e.g. Zaldarriaga and Seljak, 1997; Kamionkowski, 2007). On the experimental level, several polarization-sensitive instruments didn't encounter any fundamental observational obstacle when setting upper bounds to the BB spectrum (e.g. Kovac et al., 2002; Page et al., 2007; Pryke et al., 2009; Hinderks et al., 2009). In this period Polarbear was devised as B-mode dedicated experiment (Lee et al., 2008). By constraining both the lensing and the primordial B modes, the project aims at probing both the energy scale of inflation and the total mass of the neutrinos as well as dynamical dark energy scenarios.

The constraining power of the experiment will increase over time as the project evolves in its different stages: Polarbear, PolarbearII and Simons Array. After the 2010 engineering campaign in the Cedar Flat in the Inyo Mountains of California, Polarbear was deployed in Chile in 2011 and saw first light on January 10th 2012. In summer 2013 I started my PhD and joined the Polarbear effort, to which I contributed at the data analysis level. The results of the first observational campaign were published in December 2013 and March 2014.

5.2 The site

The Polarbear experiment is installed at the James Ax Observatory, located in Chilean Atacama desert on Cerro Toco, at West longitude $67°47'10.40''$ and South latitude $22°57'29.03''$. Being at 5200 m above the sea level, it is one of the observatories located at the highest elevation. The main obstacle of for microwave observations is the water vapour in the atmosphere. The dryness of the atmosphere makes Polarbear site one of the best microwave observational site in the world after the south pole. During Polarbear first observational campaign, the median precipitable water vapour (PWV) measured by the nearby APEX water vapour radiometer was 1.0 mm. This corresponds to a sky brightness in the Polarbear design band of 12 K at an elevation angle of $60°$. In addition to this very good observational condition, the site has a considerable logistical advantage. It is indeed accessible throughout the whole year: the site team drives from San Pedro de Atacama (2400 m of elevation) to the observational site on a daily basis.

5.3 The instrument

The Polarbear experiment is composed of a cryogenic receiver mounted on the Huan Tran Telescope (HTT). A more detailed overview of the instrument can be found in Kermish et al. (2012).

5.3.1 The Huan Tran Telescope

HTT is a two-mirror refractive telescope with an off-axis Gregorian design. In on-axis telescopes the secondary mirror is necessarily supported by some structure that obstructs the beam of the telescope.
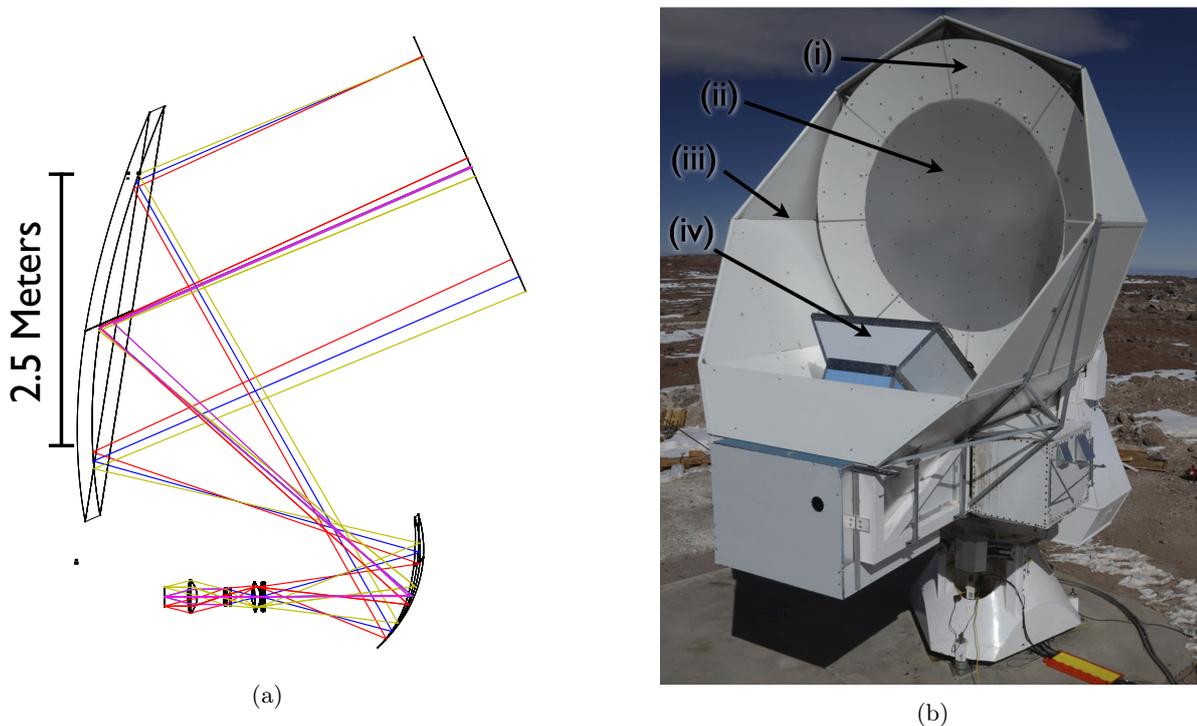
Figure 5.1: (a) Ray tracing scheme of the optics of the telescope. (b) The Huan Tran Telescope fully assembled at the James Ax observatory in Chile. (i) Primary guard ring. (ii) Monolithic precision primary mirror. (iii) Comoving shields (iv) Primary focus baffle.

These structures can scatter or diffract signal from the ground into the main beam. HTT avoids the consequent limitation in the performances by adopting an off-axis design. The angle between the symmetry axis of primary and secondary mirrors satisfies the Mizuguchi-Dragone condition (Mizugutch et al., 1976; Dragone, 1978). Consequently, the resulting antenna has a rotationally symmetric equivalent paraboloid, ensuring low cross polarization and astigmatism over a large diffraction-limited field of view. This condition can be satisfied both with a Gregorian and with a crossed Dragone design. One major difference between the two designs is the size of the secondary mirror, typically comparable with the one of the primary for the crossed configuration. HTT adopts a Gregorian configuration, which allows easier baffling to prevent far sidelobes due to scattering at the receiver window. The scheme of the optics of the telescope is schematically reported in 5.1a. In addition to the supporting mechanics, the telescope is endowed with comoving baffles that shield the detectors from stray radiation. The size of the primary mirror is 3.5 m. However, because of the 4 K aperture stop in the receiver, only the inner 2.5 m is actually illuminated, in an inverse ray tracing sense. This inner part is a monolithic aluminium dish with 53 µm r.m.s. surface accuracy. The 2.5-meter primary gives a 3.5 arcminute beam at 150 GHz, setting an upper bound of $\ell \sim 2500$ to the multipoles range to which the experiment is sensitive. The size of the beam is the most important element of the B-mode dedicated design: is sufficiently small to fully characterise the lensing B-modes—which peak at $\ell \sim 1000$—but not too small because it would impede the measurement of large scales.

5.3.2   The POLARBEAR receiver

The cryogenic receiver contains a cold half-wave plate, a re-imaging optics and a focal plane. The key design goal is to keep the emission from the Chilean atmosphere the main noise contribution. Controlling the noise of the instrument itself below this level requires a refrigerated focal plane and optics.

The optics

The cryogenic system has a large aperture, seeking high optical throughput. In order to preserve the cryogenic environment, filters have to reject the optical power outside of the band of interest. For this reason the thermal filters are located right after the vacuum window. They are composed by several
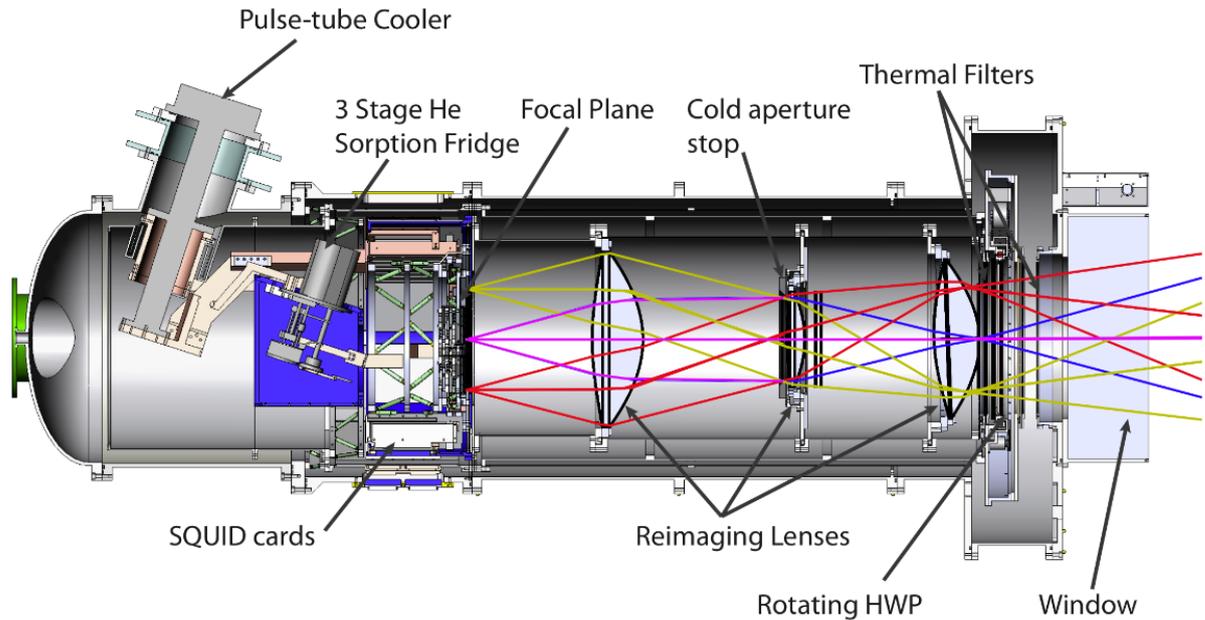
Figure 5.2: Scheme of the cryostat of POLARBEAR.

hot-pressed multilayer low-pass filters, single layer IR shaders and an IR-absorbing sheet of porous teflon.

A cold half-wave plate is located close to the aperture and refrigerated at $\sim 80\,\mathrm{K}$, for reducing its in-band emission. It is made of a single crystal of sapphire with an anti-reflection coating. The half-wave plate can be stepped, providing a way of modulating the sky signal in addition to the one given by the sky rotation. This extra modulation helps in mitigating systematics related to the non-idealities of the instrument.

A set of lenses reimage the focus of the telescope—located just in front of the first lens—on the focal plane. The lenses are made of ultra-high molecular weight polyethylene and endowed with an anti-reflection coating. They are cooled at $\sim 6\,\mathrm{K}$, to reduce the optical loading on the detectors. Together with the reimage lenses, a cold aperture stop creates the $2.4°$ diffraction-limited field of view.

The focal plane

POLARBEAR focal plane houses 637 pixels organised in 7 hexagonal wafers. Each pixel hosts two antenna-coupled transition edge sensor (TES) detectors. The two dipole antennas are sensitive to orthogonal polarization states of the incoming photons. The planar antenna is in direct contact with a silicon lenslet. The lenslet increases the directionality of the beam inherent to the antenna. Moreover, the lenslet covering the pixel magnifies the active size of the antenna, leaving some room under the lenslet, away from the centre of the antennas. The bolometers and the microstrips can be accommodated in this region of the pixel, effectively without losing precious focal plane area coupled to the diffraction limited field of view. The antennas couple the incoming radiation onto superconducting Niobium microstrips. Along the lines an on-chip bandpass filter is applied. Incoming light has been already filtered by the intervening optics, but it is the microstrip filters that effectively define the bandpass of the detectors around 150 GHz (see Fig. 5.3d). After filtering, the radiation is dissipated in a load resistor located in a thermally isolated island. A TES thermistor converts changes in the absorbed optical power into a modulated current. An image of the focal plane and its parts is reported in Fig. 5.3, more details can be found in Arnold et al. (2012).

Readout

One of the main difficulties in deploying kilo pixel arrays is how to read the signal from the bolometers without disturbing the sub-Kelvin cryogenic environment of the focal plane. In POLARBEAR this task is accomplished by using a digital frequency domain multiplexer (DfMUX) (Dobbs et al., 2008).

Each bolometer is biased with a constant amplitude sinusoidal voltage, defined by am LC tank circuit in series with the bolometer. The circulating current is converted to magnetic flux by an inductor coil

(a)

(b)

(c)

(d)
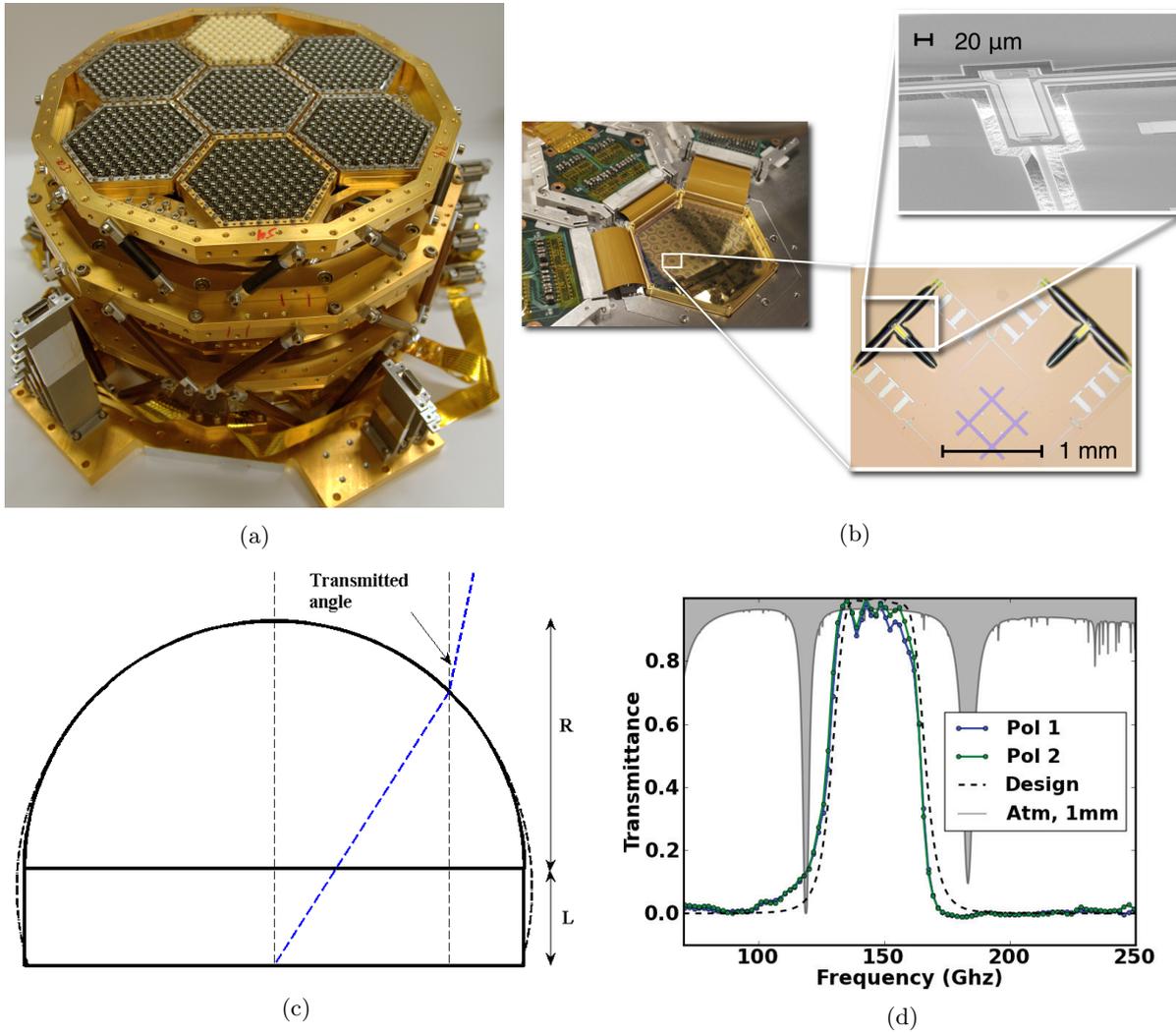
Figure 5.3: (a) The fully assembled POLARBEAR focal plane. The photo also shows the support structure and the milliKelvin wiring. In Fig. (b) the left panel shows the back of a wafer, the unfolded LC boards are visible around of the hexagonal support. The bottom-left inset is an image of the focal plane pixel underneath a lenslet, showing the orthogonal antennas, the microstrips (the microstrip filters are also visible) and the two bolometers. The top-right panel zooms on the thermally isolated island of the bolometer.
(c) Schematic representation of a lenslet. The extended hemisphere approximates an elliptical shape (dashed line). In the latter case, for a specific value of the eccentricity defined by the optical index of the silicon, the transmitted angle would be zero. Since the values of the eccentricity would be small, in POLARBEAR the elliptical lenslet is approximated with an extended hemisphere, with a value of L/R ranging between 0.36 and 0.39. Fig. (d) shows the $150\,\mathrm{GHz}$ bandpass fitting the atmospheric transmittance window.

Patch	R.A.	Decl.	Effective Area
RA4.5	$4^{\mathrm{h}}40^{\mathrm{m}}12^{\mathrm{s}}$	$-45°$	$7.0\,\mathrm{deg}^2$
RA12	$11^{\mathrm{h}}53^{\mathrm{m}}0^{\mathrm{s}}$	$-0°30'$	$8.7\,\mathrm{deg}^2$
RA23	$23^{\mathrm{h}}1^{\mathrm{m}}48^{\mathrm{s}}$	$-32°48'$	$8.8\,\mathrm{deg}^2$

Table 5.1: The three POLARBEAR patches observed during the first and the second observational campaigns

and measured by a superconducting quantum interference device (SQUID) amplifier. The optical power absorbed by the bolometer changes the resistance of the bolometer and therefore the resonant frequency of the circuit. Consequently, part of the signal is moved to the sidebands adjacent to the carrier, from which the optical energy absorbed by the bolometer is estimated, after demodulation at room temperature.

Each detector has a different bias frequency, ranging from 300 kHz to 1 MHz. This allows POLAR-BEAR's 8x DfMUX to pack the signal from 8 bolometers on the same wire and measure them with one single SQUID. This cutting-edge technology allowed to dramatically reduce the volume of the readout electronics as well as the heat it produces.

The LC circuits are lithographed on boards folded under the bolometer array. In order the make the noise of the carrier negligible, the LC boards are cooled at $\sim 0.25\,\mu$K. The SQUIDs operate at ~ 4 K, the temperature is needed not only for their functioning but also for achieving high gain in the readout.

5.4 First and second season observation

Between May 2012 and June 2016 the first two POLARBEAR observational campaigns took place. They focused on three small patches integrated deeply (see Table 5.1). Their locations were chosen on the basis of three criteria: low dust intensity, availability throughout the day and overlap with other observations. RA23 and RA4.5 overlap with QUIET observations (Araujo et al., 2012), RA12 and RA23 overlap with Hershel ATLAS fields (Eales et al., 2010), which was exploited in the cross-correlation study published in December 2013 (POLARBEAR Collaboration, 2014b). The observation included also astrophysical sources, used as on-sky calibrator. Tau A, in the crab nebula, has an intrinsic polarization and it was used in order to calibrate the polarization angle of the detectors. Jupiter and Saturn were also observed and they were used for characterizing the beam of the experiment.

The observations are scheduled according to the cryostat fridge cycle, normally 36 hours long. The typical schedule starts with calibration observations, in which the telescope tracks either planets or Tau A. Afterwards, the CMB observation stars with one of the patches. The fundamental unit of the POLARBEAR CMB observation is the *constant elevation scan* (CES). A CES consists of a pointing phase, which brings the boresight to where the centre of the patch will be in 7.5 minutes, and the observation itself. The observation consists of approximately 8°-wide sweeps in azimuth while keeping the elevation fixed, each of them is called *subscan*. During a subscan the telescope moves at the constant velocity of 0.75 deg /s, afterwards it slows down in order to invert the azimuthal direction for the subsequent subscan. Only the data taken in the constant velocity portion are retained for the CMB analysis, 74% of the observing time. After 15 minutes (corresponding to roughly 150 subscans) the patch leaves the field of view of the telescope and a new CES starts. Each patch is available for 6-8 hours per day. The minimum elevation for CMB observation was 40° for the first three months and 30° for the rest, while the maximum observing elevation is 80°. As the elevation changes the atmospheric optical loading varies. In order to optimise the linearity and sensitivity of the detectors, the detectors are re-biased every hour. Before and after re-bias the relative gain between detectors is assessed using both elevation nods and stimulator observations. The former consists of 2° elevation movements of the telescopes. During the latter, the detectors observe a thermal source located behind the secondary mirror. The source is kept at 700 deg C and is modulated at a frequency ranging from 4 to 44 Hz by a chopper. Stimulator observations also enable to estimate the time constants of the detectors–which were found to be of the order of the mHz and thus negligible compared to our science band (1.04 to 3.13 Hz).

5.5 Overview of the low-level data processing

My main contribution to the POLARBEAR analysis was the unbiased map-making, extensively discussed in Chapter 6. It requires the timestreams to be prepared: the measurements of the detectors have to be calibrated and, for each time sample, the observed sky pixels and polarization angle are also required. For completeness, this section presents an overview of how this information is extracted from the observations. More details can be found in POLARBEAR Collaboration (2014a).

5.5.1 Pointing

The boresight pointing is reconstructed using a five-parameter pointing model (Magnum, 2014). The parameters are azimuth and elevation zero encored offset, IA and IE, the collimation error of the electromagnetic axis, CA, and the azimuth axis offset/misalignment (north-south and east-west), AN and AW. The parameters are estimated using point sources selected from Wrobel et al. (1998) and Murphy et al. (2010). The robustness of the pointing model was tested by estimating the parameters only from a subset of the data. Some deviations were found and propagated to the systematics pipeline in order to asses their impact on the BB spectrum.

The beam offset of the individual detectors with respect to the boresight is estimated using Jupiter and Saturn scans. These offsets are measured several times per weeks and the fluctuations over time were less than $6''$.

5.5.2 Beam

The beam is reconstructed using Jupiter observations. For each observation a map is produced after a masked filtering of a zero and first order polynomial for each sweep over the planet. It is masked in the sense that the samples within a $50''$ radius from the planet are excluded from the evaluation of the amplitude of the polynomial to be removed. Single detector maps are then coadded using the rms outside of the radius as weight. The beam is calculated as the azimuthal average of the two-dimensional Fourier transform dinned in $\Delta\ell = 80$ wide bins. The so-obtained B_ℓ is then corrected for the pixel window function of the observation map and the finite size of the planetary disk.

We assume the inaccuracy of the pointing model to act as a blurring of the coadded CMB map. This is modelled as a σ_p r.m.s. pointing error in map space. The resulting effective beam is then assumed to be $B_\ell^{\text{eff}} = B_\ell \times e^{-\ell(\ell+1)\sigma_p^2/2}$ and σ_p is estimated patch by patch by fitting the point sources present in the observed field.

5.5.3 Gain

The raw TOD measured by the detectors are currents. These values have to be converted into K_{CMB}, this process is done in three steps.

Before and after every period of constant bias of the detectors (one hour of observations) the relative gain between the detectors is assessed by performing an observation of the stimulator, $700 \deg \text{C}$ thermal source behind the secondary mirror observable through a $6\,\text{mm}$ light-pipe. For this purpose is very important to accurately characterise the polarization fraction of the thermal source, since it results in a relative gain error that causes intensity to polarization leakage. We propagate the related uncertainty to the systematics pipeline in order to assess the impact of relative gain errors on the BB spectrum.

Using observations of Saturn we perform a first absolute calibration of the different observations. The current model takes into account the variation of the intrinsic brightness of Saturn due to differences in the opening angle under which the rings are seen (see Weiland et al., 2011). Neglecting this effect caused up to 10% differences in the calibration products (see Peloton, 2015, Chapter 6, for more details).

Later in the analysis, the TT power spectrum is fitted to the one obtained from the *Planck* maps (or, formerly, WMAP best fit model). Since the correction for the time variation of the Saturn rings opening angle was introduced, this TT-estimated absolute gain correction gives a factor compatible with 1.

5.5.4 Polarization angle

Tau A (the Crab Nebula) is a polarised supernova remnant, characterised or used as a calibrator by many experiments (Farese et al., 2004; Weiland et al., 2011; Naess et al., 2014b, e.g.). Also in POLARBEAR we use Tau A for measuring the orientation of each of our detectors projected on the sky. Our calibration
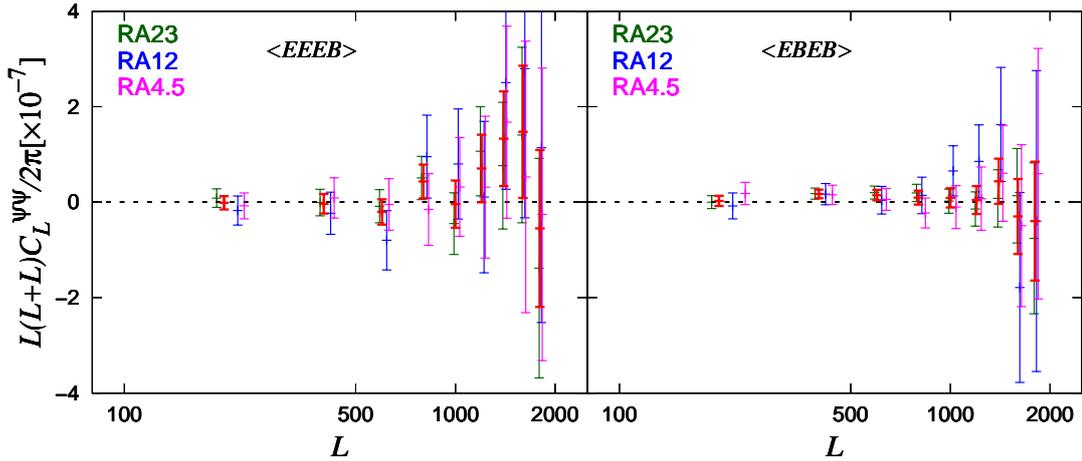
Figure 5.4: The power spectrum of the curl component of the deflection field. For both the $\langle EBEB \rangle$ and the $\langle EEEB \rangle$ estimator the spectrum computed for the three patches is displayed as well as their combination, they are all consistent with zero.

is based on the IRAM measurement (Aumont et al., 2010). With their 30 m telescope they produced a high resolution map ($30''$) at 90 GHz.

The method we use consists in producing simulated timestreams of each detector by scanning the IRAM map, taking into account each individual beam. Then we consider the difference timestream of the bolometers in a pair both for the data and the simulations. The simulated one has three free parameters: the polarization angle, the polarization efficiency and the relative gain between detectors. We have only one free polarization angle for each pair of bolometers because we assume their relative angle to be exactly 90°. Similarly to the beam analysis, simulated and real timestreams are filtered with a fifth order polynomial excluding from the evaluation of the amplitude of the polynomial all the samples within a $10'$ from the centre of Tau A. Samples within a $5'$ radius are then used for estimating the free parameters by fitting the simulated timestreams to the real one.

This determination of the polarization angle has a 1.2° r.m.s. variation across all the observation. This level of inaccuracy can produce relevant E to B leakage compared to our noise level. We therefore perform later a power spectrum-based calibration. Assuming CMB EB spectrum to be zero, a measured non-zero EB spectrum can be produced by a miscalibration of the polarization angle. Since our uncertainty is small, the model for the measured EB power described in Keating et al. (2013) simplifies to $C_\ell^{EB} \simeq \pm 2\Delta\psi(C_\ell^{EE} - C_\ell^{BB})$, where the plus or minus depends on the convention adopted for measuring the polarization angle. We fit $\Delta\psi$ in our measured EB spectrum and use that value to remove the consequent $4\Delta\psi^2 C_\ell^{EE}$ leaked EE power from our BB spectrum.
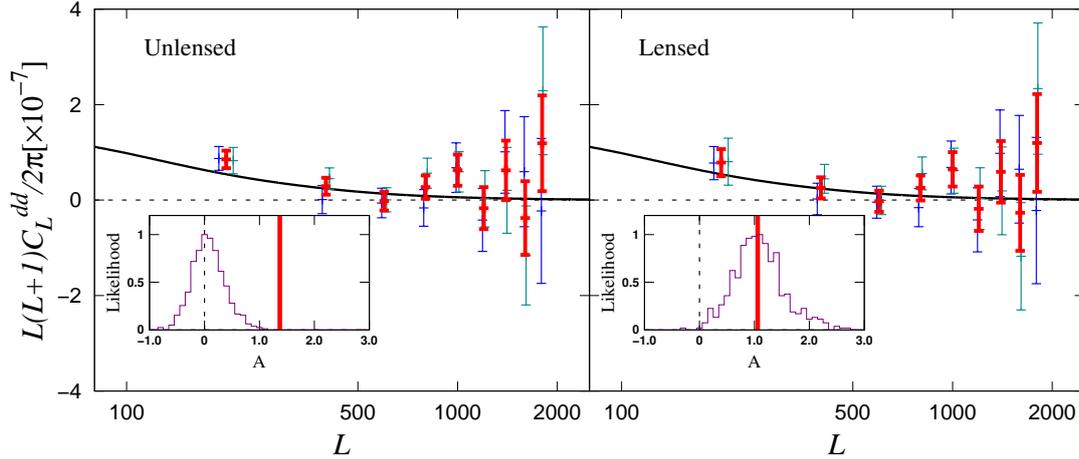
Figure 5.5: Power spectrum of deflection field assuming no lensing polarization signal (*left*) and assuming ΛCDM theoretical polarization lensing (*right*). The insets show the distribution of $\mathcal{A}$–the fitted amplitude of the ΛCDM theoretical C_L^{dd} shape–of 500 simulations. The left panel yields a 4.6σ statistical rejection of the null hypothesis.

5.6   Results from the first season

The data collected between June 2012 and June 2013 were used for the first set of results. Four publications were delivered using the sub-degree polarization measurements of POLARBEAR.

5.6.1   Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum with the POLARBEAR Experiment

From each of the three POLARBEAR patches the deflection field was estimated using the quadratic estimator Eq. (3.81). Two quadratic estimators were considered, $d_{EE}(\boldsymbol{L})$ and $d_{EB}(\boldsymbol{L})$. Since the focus of the POLARBEAR experiment are the polarization B modes, the power spectrum of the deflection field was estimated using only $\langle d_{EE}d_{EB}\rangle$ and $\langle d_{EB}d_{EB}\rangle$, respectively referred to as $\langle EBEB\rangle$ and $\langle EEEB\rangle$.

The curl component of the deflection field is found to be consistent with zero (Fig. 5.4), as expected from the linear theory of weak lensing (see Sec. 3.8). On the contrary, the spectrum of C_L^{dd} shows evidence for gravitational lensing: the null hypothesis of no signal is rejected at 4.6σ statistically (4.2σ combining the statistical and systematic errors in quadrature) and at 4.6σ statistically if only the $\langle EBEB\rangle$ estimator is considered.

The analysis is also performed assuming the predicted amount of polarization predicted by the ΛCDM. The first consequence is that the weighting given to the different $\boldsymbol{l}$ and $\boldsymbol{l'}$ in the quadratic estimator is different. The second consequence is that the error bar calculation has to include the sample variance of the signal. We fit the amplitude of the ΛCDM theoretical C_L^{dd} shape, the value measured is $\mathcal{A} = 1.06 \pm 0.47(\text{stat}) \pm {}^{+0.35}_{-0.31}(\text{syst})$, consistent with the ΛCDM value of 1, see Fig. 5.5.
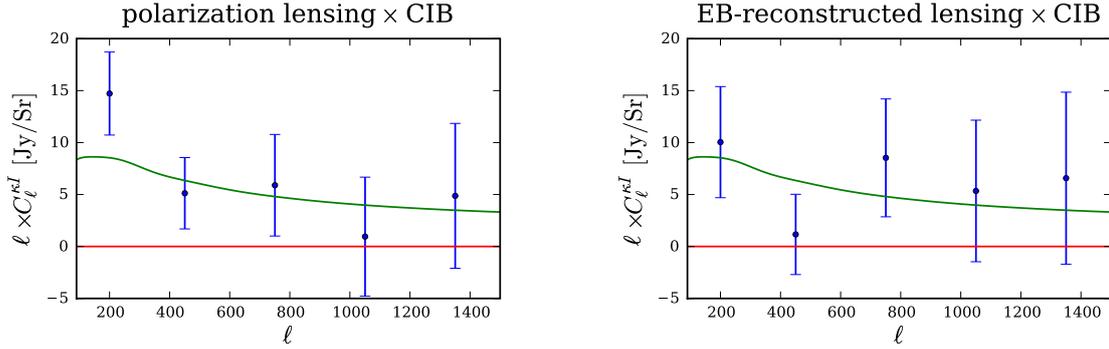
Figure 5.6: Spectrum of the correlation between the convergence field κ estimated from POLARBEAR polarization maps and the CIB map of the *Hershel* ATLAS fields. In the right panel only the EB quadratic estimator of κ is adopted. In the left panel EB and EE quadratic estimators are combined

This study was the first evidence of CMB lensing derived from polarization alone. More details on this study can be found in The POLARBEAR Collaboration (2014).

5.6.2 Evidence for gravitational lensing of the cosmic microwave background polarization from cross-correlation with the cosmic infrared background

As previously mentioned, the RA12 and RA23 patches overlap with ATLAS fields. The far-IR/submillimeter emission mapped by the *Hershel* satellite is mostly star emission absorbed and re-emitted by dust in the galaxies, therefore this cosmic infrared background (CIB) is a tracer of the matter density. As explained in Sec. 3.8, the convergence field κ is basically the Laplacian of the matter density integrated along the line of sight and therefore is expected to be correlated with the CIB. In POLARBEAR Collaboration (2014b) we correlate the CIB with the EB-estimated and EE-estimated κ field and find a 4.0σ correlation (2.6σ for the EB estimator alone). The power spectrum of the cross-correlation is reported in Fig. 5.6.

5.6.3 A measurement of the cosmic microwave background B-mode polarization power spectrum at sub-degree scales with POLARBEAR

The BB-power spectrum extracted from POLARBEAR polarization maps in the $[400, 2100]$ multipole range was published in POLARBEAR Collaboration (2014a). The spectrum is reported in Fig. 5.7, it yields a 97.2% confidence level rejection of the hypothesis of no B modes. Fitting the amplitude of the theoretical shape of the ΛCDM BB spectrum gives $A_{BB} = 1.12 \pm 0.61(\text{stat}) \pm^{+0.04}_{-0.12}(\text{syst}) \pm 0.07(\text{multi})$, where "stat" represents the statistical uncertainty, "syst" represents the possible systematic bias due to either the instrumental or the astrophysical foregrounds and "multi" represents a multiplicative uncertainty due to calibration. In addition, the paper provides details on the data processing adopted also for the previous two papers (outlined in the previous section) and also a thorough investigation of the possible impact of systematic effects (see Fig. 5.7).

5.6.4 POLARBEAR constraints on cosmic birefringence and primordial magnetic fields

Similarly to weak lensing, an anisotropic polarization rotation field α, arising for example from anisotropic cosmic birefringence, can be probed using quadratic estimators. In particular, a non zero $\alpha(\boldsymbol{l}+\boldsymbol{l}')$ sources a non zero $E(\boldsymbol{l})B(\boldsymbol{l}')$ correlation. Thus the power spectrum of the rotation field can be measured using a $\langle EBEB \rangle$ estimator, see Sec. 3.8.3 for more details. This power spectrum was measured in POLARBEAR Collaboration (2015) using the POLARBEAR first season data and is reported in Fig. 5.8. Its consistency with zero sets an upper limit the power spectrum of the magnetic fields that might source the cosmic birefringence. Assuming a scale-invariant spectrum, $B_{1\,\text{Mpc}}$–the strength of the magnetic field smoothed at a scale of $1\,\text{Mpc}$–is constrained to be lower than $93\,\text{nG}$ (95% confidence level, inclusive of systematic uncertainty).
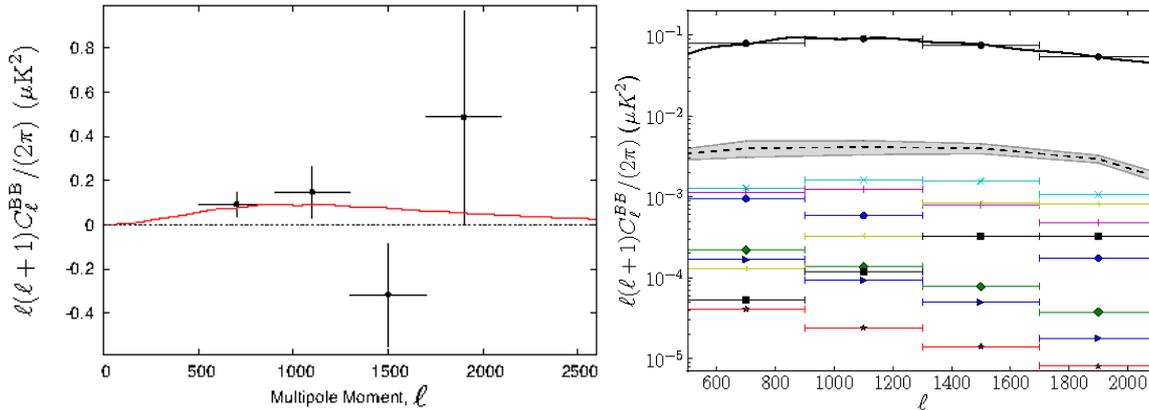
Figure 5.7: *Left: BB* power spectrum measured from the three POLARBEAR patches of the first season. For comparison, a theoretical WMAP-9 ΛCDM theoretical *BB* power spectrum is displayed in red. *Right:* Possible bias from instrumental systematics errors in the *BB* power spectrum. Colour points represent the individual uncertainty while the black dashed line is the cumulative bias. The estimation considers all the CMB patches. The grey-shaded region show the 1σ bounds on the cumulative bias limit, after the self-calibration procedure described in Sec. 5.5. The bias from Individual effects is evaluated by adding the systematics in simulated observations of an *E*-only simulated sky. The effects included in this analysis were the boresight and differential pointing uncertainty (light blue, cross mark), the residual uncertainty in instrument polarization angle after self-calibration (purple, plus mark), the differential beam size and ellipticity (yellow arrow and black square mark respectively), the electrical crosstalk (blue, arrow mark), the drift of the gains between two consecutive thermal source calibrator measurements (red star mark), and the HWP-independent and HWP-dependent terms of the relative gain model (green diamond and blue circle mark respectively).

Magnetic fields present before the recombination source vector and tensor perturbations in the metric. These perturbations leave an imprint in CMB *B*-mode anisotropies. At the small scales probed by POLARBEAR, the *B*-mode spectrum is particularly sensitive to the vector perturbations that might have been sourced by an inhomogeneous magnetic field before recombination (see. Fig. 5.8). In POLARBEAR Collaboration (2015) we also use the measurement of the *BB* power spectrum to constrain the amplitude of such primordial magnetic fields to $B_{1\,\mathrm{Mpc}} < 3.9\,\mathrm{nG}$ at 95% confidence level, assuming a flat prior on the magnetic field amplitude.

5.7   The future of the POLARBEAR project

Since May 2014 the focus of POLARBEAR moved larger angular scales. Since then the telescope started scanning the sky region observed by the BICEP2 experiment. The effective area is roughly 20×50 deg^2 and it is observable for nearly 14 hour per day. These large patch observations will be used mainly as a pathfinder for POLARBEARII (Suzuki et al., 2014a). These two phases of the POLARBEAR project not only have comparable scanning strategy—both are interested in scales considerably larger than the ones probed by POLARBEAR first and second season—but they also have important similarities in the instrumental set up. Since the beginning of the third observational campaign, a continuously rotating warm half-wave plate was mounted at the prime focus, between the secondary and the primary mirror. It rotates at 2 Hz, and consequently the incoming polarized radiation is modulated at 8 Hz.

A warm continuously rotating half wave plate is the current baseline of POLARBEARII. However, the most important innovation of POLARBEARII is the multi-frequency coverage, adding a 95 GHz band to the 150 GHz of POLARBEAR. The POLARBEARII focal plane is composed of 1897 pixel, each of them sensitive to the two orthogonal polarization states of the two frequency. The sinuous antennas (see Fig. 5.9a and O'Brient et al. (2013), for more details) enable to couple a broad range of frequencies to the microstrip lines. The two frequencies can then be selected by microstrip filters and directed to different bolometers. Thus each focal plane pixel has four bolometers, for a total of 7,588 detectors. Traditionally, detectors sensitive to different frequencies occupy separate focal plane pixels. This multichroic technology allows for a much more efficient usage of the focal plane area, enabling faster scanning speed. More
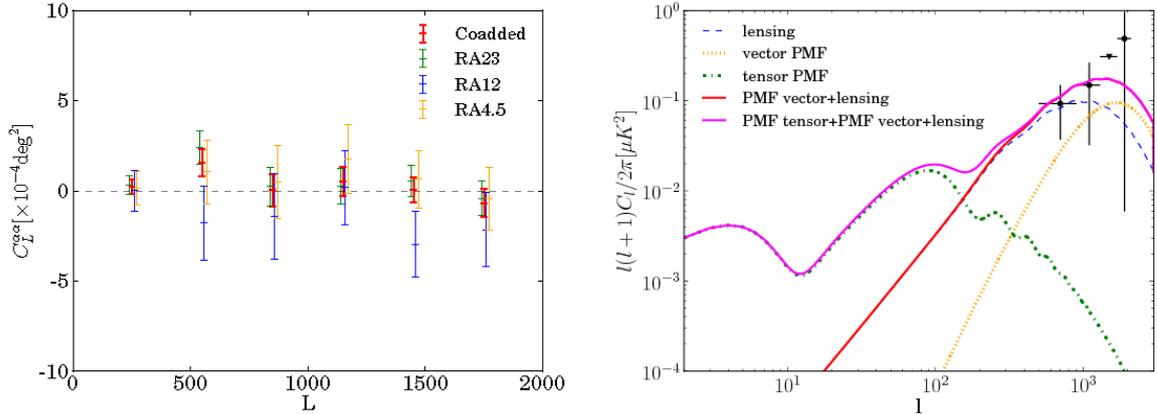
Figure 5.8: *Left:* the power spectrum of the rotation filed as measured from POLARBEAR using the first season data. *Right:* A representative B-mode polarization power spectrum sourced by a scale-invariant primordial magnetic field. Shown are the passive tensor mode (green), the compensated vector mode (orange), the gravitational lensing contribution (blue), and the combinations of the lensing and vector B modes (red) and all three components (magenta). See POLARBEAR Collaboration (2015) for more details.

details about the POLARBEARII focal plane can be found in Suzuki et al. (2014b). As of fall 2016, the POLARBEARII focal plane, optics and cryostat are under construction (Fig. 5.9b). A twin HTT is being assembled at the site and will soon be ready to host the new receiver. The complete POLARBEARII is expected to see first light in summer 2017.

In 2018, two POLARBEARII-like receivers will be deployed. One of them is an exact replication of POLARBEARII, the other adopts the same dichroic technology but is sensitive to 150 and 220 GHz. They will be mounted on the current HTT and on a third twin, currently under construction. This three-telescopes stage is called Simons Array (Arnold et al., 2014). The multi-frequency coverage will allow the rejection of both the dust and the synchrotron foreground signals. Including the statistical uncertainty endemic to this component separation process, Simons Array will constrain the tensor to scalar ratio and the total mass of the neutrinos down to $\sigma_{r=0.1} = 6 \times 10^{-3}$ and $\sigma_{\sum m_\nu} = 40 \, \text{meV}$. Fig. 5.9c shows the forecast sensitivity at the power spectrum level.

In the spring, a further stage has been funded. It is called Simons Observatory and will merge the POLARBEAR and ACT collaborations. The Simons Observatory will be an intermediate step towards Stage IV, a next-generation observatory in which dedicated telescopes from the south pole, the Chilean Atacama plateau and possibly northern hemisphere sites will observe a large fraction of sky with a total of 500,000 detectors. Stage IV is expected to be operational by 2020 and observe the sky for at least 5 years. More information can be found in the CMB S4 wiki[1] and in the (soon to be published) science book.

5.8 Summary

The POLARBEAR experiment is one of the leading CMB experimental projects currently in operation. So far it produced a host of results concerning the CMB polarization signal on small angular scales. The next campaigns (called POLARBEARII and Simons Array) will explore the *B*-modes on large scales. These efforts will culminate in the Simons Observatory, the comprehensive, multi-frequency, multi-telescope experiment aiming at setting unprecedented constraints on small and large angular scales, thus probing both primordial gravity waves and the masses of the neutrinos. At present, the final stages of the small-scale *B*-mode POLARBEAR analysis is taking place. I have actively taken part in this analysis since the beginning of my thesis and in the following chapters I describe the contributions I have made.

---

[1] https://cosmo.uchicago.edu/CMB-S4workshops/index.php/Main_Page

(a)

(b)

(c)
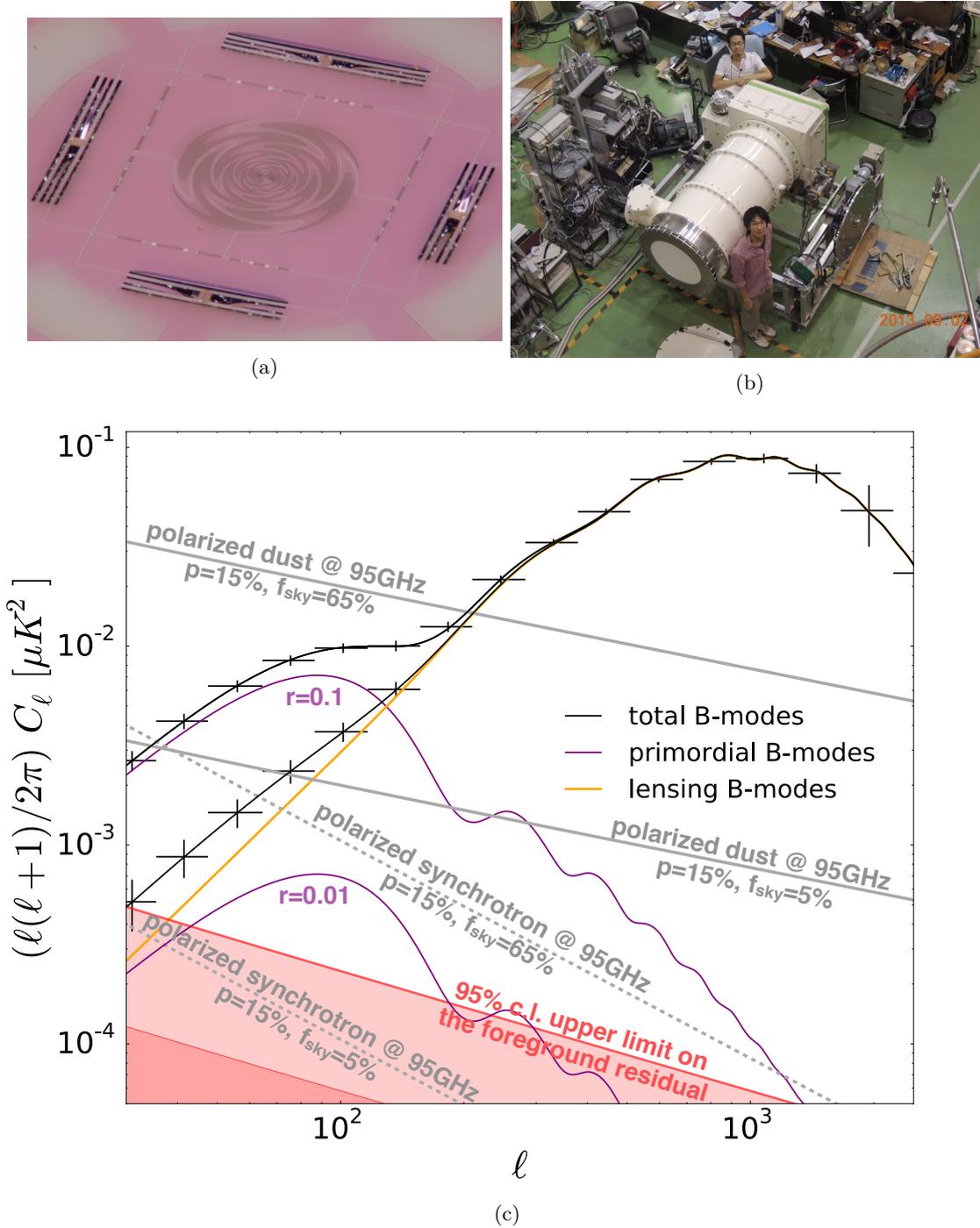
Figure 5.9: (a) A POLARBEARII focal plane pixel, the antenna has sinuous design, which makes them broadband. Four detectors are fitted in a single focal plane pixel, one for each polarization direction of the two frequency bands, centred at 95 and 150 GHz. (b) the POLARBEARII cryostat being assembled at KEK. (c) Forecast of the sensitivity of Simons Array (Errard et al., 2016).

Chapter 6

THE POLARBEAR UNBIASED MAP-MAKER

The project of an unbiased map-maker for the POLARBEAR experiment was started at APC in the summer of 2013 and now involve also our collaborators at SISSA and Sussex University. Two codes are currently available, an iterative and an explicit map-maker. The latter is the current baseline. I contributed to it by leading its design, implementation and validation, which will be illustrated in this chapter.

The implemented estimator is the unbiased estimator discussed in Sec. 4.1.4. Given the small sky area observed by POLARBEAR, the number of pixels observed was limited to about 5×10^4. On the contrary, as described in Sec. 6.2, the number of templates to be filtered out is roughly 10^9. Therefore the direct solution of Eq. (4.29) is preferable to the two-step solution involving Eq. (4.30) and Eq. (4.31). As already mentioned, the other fundamental property of the implementation is the explicit solution of the inverse problem in Eq. (4.29).

After a description of the formalism and its implementation, we assess its performances and compare it with both a different implementation (the popular iterative solver) and a different estimator (the biased estimator Eq. (4.32)).

6.1 Goals, requirements and constraints

In POLARBEAR we already had an iterative unbiased map-maker. The motivations for the implementation of the explicit solver were the following.

Simulation capability. The advantage of the iterative solution of the map-making equation is the speed of the algorithm. The cost indeed is equal to the cost of the application of the system matrix $\boldsymbol{A}^{\top} \boldsymbol{F_T} \boldsymbol{A}$ times the number of iterations. Moreover, the system matrix doesn't even need to be explicitly calculated, we can simply apply sequentially the linear operators of which it is composed. Postponing for a moment the quantitative derivation, a map-making run of the iterative solver cost $\sim 10^{14}$ operations while explicitly computing $\boldsymbol{A}^{\top} \boldsymbol{F_T} \boldsymbol{A}$ and inverting it costs $\sim 10^{15}$ operations. Thus, for a single run, the iterative solver is more convenient than the explicit solution. However, a fundamental requirement of any data analysis pipeline is the capability of running simulations, typically hundreds or thousands. As long as the observation and filtering are the same for all the simulations, the explicit implementation can recycle all the heavy computations and get an arbitrary number of simulations for a negligible extra cost. On the contrary, the iterative solution cannot take advantage of the fact that for each simulation the system being solved is the same.

Accurate reconstruction. Another disadvantage of iterative solvers is that they not always accurately converge (within a reasonable number of iterations) to the solution. As discussed in Sec. 6.6.2, we indeed encounter this problem in POLARBEAR.

Characterise the degeneracies. The iterative solver suffers from convergence problems especially if the system matrix is badly conditioned. As detailed in Sec. 4.2, this happens when there are some

degeneracies in the data model. The explicit inversion of the system matrix through the calculation of its eigenstructure provides quantitative insights on the possible degeneracies.

Strict implementation of the estimator.   The previous motivations are general differences between iterative and explicit solvers. In our case, we also had motivations specific to our implementation of the iterative estimator. One of the most important is the lack of orthogonalisation of the filters: instead of building the filtering operator in Eq. (4.13), two set of templates (the ground templates and the polynomial templates, discussed in Sec. 6.2) are sequentially filtered. Though internally orthogonal, the lack of mutual orthogonality (and even the degeneracies) between these two set of templates didn't ensure proper marginalisation over them. This approximation was forced by the serial implementation of the iterative mapmaker: the memory available to the single processor is insufficient to store the orthonormalisation kernel in the memory. This problem is thus related to the following requirement.

Intrinsically massively parallel implementation.   The serial implementation of the iterative solver also limits the amount of data it is able to grind at a time. Indeed, an unbiased map can be computed only from a maximum of 24 CESs. Moreover, obtaining the memory necessary to analyse this cases requires a severe under-population of the compute node (typically using only 2 processors out of 24), making an inefficient use of the computational resources.

Since the computational aspects are part of the motivations for the explicit implementation of the unbiased estimator, when devising the algorithm and the implementation we had to bear in mind the specifications of the machines on which the code was supposed to run. As part of the repository mp107— dedicated to the data analysis of post-*Planck* experiments—we had between 8Mcpuh and 11 65Mcpuh[1] available at the National Energy Research Scientific Computing Center (NERSC). The machine we used and their main specifications are

Name	Hopper	Edison	Cori Phase I
Machine	Cray XE6	Cray XC30	Cray XC based
Node # CPUs	24	24	32
Node memory [GB]	32	64	128
CPU clock rate [GHz]	2.1	2.4	2.3

The volume of POLARBEAR first and second season dataset for the RA23 patch is about 1.2 TB. As it will be explained later in Sec. 6.3, the map estimator requires other 3 TB of memory, in addition to the data volume loaded in the memory. Given the machines at our disposal, this required to implement a code that scales properly when run on thousands of processors.

6.2   The data model

In this section we illustrate in detail the POLARBEAR data model, providing a concrete and detailed example for the general considerations presented earlier in Sec. 4.1.4 and Sec. 4.2.4. Indeed, as any other ground-based CMB experiment, our raw TOD contain a contribution from the following two contaminating signals.

Atmospheric signal.   The observational site of POLARBEAR provides a very dry environment, which minimises the impact of atmospheric emission on the measurements of our detectors. The atmosphere signal gives nevertheless a relevant contribution to our time streams, especially at low frequencies. Fig. 6.1 shows the power spectral density (PSD) of the sum and the difference of two bolometers in the same focal plane pixel. The TOD of a single CES are noise dominated and therefore the figure essentially represents the PSD of our noise. Above a knee-frequency, the noise is basically white. The atmospheric emission is responsible for the $1/f$ tail of the noise. Since it is not polarized it affects mostly the sum of the detectors (i.e. the total intensity measurement). However, because of the bandpass mismatch, a small fraction of the $1/f$ intensity noise leaks into the difference of the detectors (i.e. the polarisation-only measurements). The figure highlights the frequency band where we expect our target CMB signal

---

[1]hours to be shared with all the other users of the repository
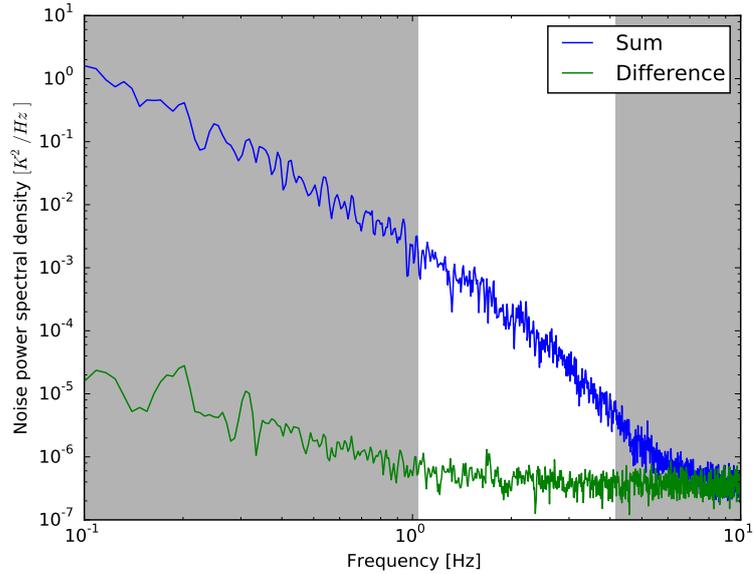
Figure 6.1: Power spectral density of the sum and the difference of a pair of detectors.

to be. For temperature, the whole science band is affected by the $1/f$ noise. For polarization, only the low-frequency part of the band is affected by $1/f$ noise. The extent of this $1/f$ contribution to the science frequencies varies depending on the detector, the weather, etc. A dedicated study of TODs recorded by different detectors in a variety of observing conditions empirically found that a first order polynomial filtering of each subscan suppresses the $1/f$ contribution to the polarization measurements most of the times. This procedure, of course, removes also the signal in those frequencies but has the advantage of making the map estimation sensitive only to the white part of the noise. Since total intensity measurements are more affected by atmospheric $1/f$ noise, the subscans of a detector sum are filtered up to a third order polynomial, which however does not remove completely the $1/f$ component. Our analysis and results focus on polarization but we mention that, consequently, our total intensity white noise simulations do not accurately reproduce the noise of the actual POLARBEAR data, analysed in the next chapter.

Ground pickup. As described in the previous chapter, the design of the POLARBEAR telescope include several comoving baffles for preventing the ground emission from leaking into the measurement. However, as any other ground-based experiment, we can not consider our TOD to be free from ground pickup. We therefore need to filter some signal stationary in ground coordinates. This is done by binning the azimuthal span of each detector and CES in 0.08 deg-wide intervals and removing for each bin the mean of all the samples taken in that range. In order to check that the size of the azimuthal bins enables to capture the structure due to the ground pickup we first used our data set to make a map in ground bound coordinates, then this map is scanned in a simulated observation and the resulting TOD are filtered and used to create a sky map that gives an estimation of the leaked ground signal into our CMB estimation. We checked that this residual is negligible compared to our statistical uncertainty.

Following the prescription illustrated in Sec. 4.1.4, the two set of filters just described and adopted in POLARBEAR Collaboration (2014a) are achieved by introducing appropriate templates in the data mode, which we describe in details in the rest of this section.

For convenience, we first consider data collected during a single CES by a single pair of detectors in the same focal plane pixel. This is the fundamental unit of our data model and we have a total of $\mathcal{N}_{pair} \times \mathcal{N}_{CES} \sim 1.5 \cdot 10^6$ of them. The two detectors in such a pair are sensitive to two orthogonal polarizations and are referred to as $\parallel$ and $\perp$. As discussed in Sec. 4.2.4, we model their TOD to include

signal, ground pickup, templates related to the correlated noise and an actual noise term:

$$
\begin{bmatrix} \boldsymbol{d}^{\|} \\ \boldsymbol{d}^{\perp} \end{bmatrix} = \boldsymbol{As} + \boldsymbol{Gg} + \boldsymbol{Bx} + \boldsymbol{w}
$$

$$
= \begin{bmatrix} \boldsymbol{A}^{\|} \\ \boldsymbol{A}^{\perp} \end{bmatrix} \boldsymbol{s} + \begin{bmatrix} \boldsymbol{G}^d & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}^d \end{bmatrix} \begin{bmatrix} \boldsymbol{g}^{\|} \\ \boldsymbol{g}^{\perp} \end{bmatrix} + \boldsymbol{Bx} + \begin{bmatrix} \boldsymbol{w}^{\|} \\ \boldsymbol{w}^{\perp} \end{bmatrix}, \tag{6.1}
$$

where we have arranged the data vector so that all the measurements of the first detector of the pair are gathered together and followed by the measurements taken by the other one.

The pointing matrices, $\boldsymbol{A}^{\|/\perp}$, are as given by (4.1), and thus have only three non-zero elements for each row, which correspond to three Stokes parameters of a given sky pixel, p, observed at the time assigned to the row. These are equal to 1, $\cos(2\varphi_t^{\|})$ and $\sin(2\varphi_t^{\|})$ for detector $\|$ and 1, $\cos(2\varphi_t^{\perp}) = -\cos(2\varphi_t^{\|})$ and $\sin(2\varphi_t^{\perp}) = -\sin(2\varphi_t^{\|})$ for detector $\perp$, and for the I, Q and U signal components, respectively. We are assuming that $\varphi_t^{\|} = \varphi_t^{\perp} + \pi/2$ is strictly true, a justified assumption in the POLARBEAR case. From now on we will write simply φ_t instead of $\varphi_t^{\perp}$. The block-diagonal structure of matrix $\boldsymbol{G}$ is due to the fact that we have introduced two independent ground templates, one for each detector of the pair. Since we always use the boresight azimuth to define the azimuthal bins, each of the blocks is the same for each detector and all focal plane pixels. Matrix $\boldsymbol{B}$ describes the time-domain filtering and thus can have more complex structure. In particular it needs to account for two types of contributions: the ones that are correlated between detectors and the ones that are independent. This can be achieved by assuming that matrix $\boldsymbol{B}$ has the following structure,

$$
\boldsymbol{B} \equiv \begin{bmatrix} \boldsymbol{B}_{corr} & \boldsymbol{B}_{uncorr} & \boldsymbol{0} \\ \boldsymbol{B}_{corr} & \boldsymbol{0} & \boldsymbol{B}_{uncorr} \end{bmatrix}, \tag{6.2}
$$

where we assumed that we use the same filtering of the uncorrelated part for each of the detectors. This corresponds to the following breakdown of vector $\boldsymbol{x}$,

$$
\boldsymbol{x} \equiv \begin{bmatrix} \boldsymbol{x}_{corr} \\ \boldsymbol{x}_{uncorr}^{\|} \\ \boldsymbol{x}_{uncorr}^{\perp} \end{bmatrix}. \tag{6.3}
$$

Here, $\boldsymbol{x}_{corr}$ collects the amplitudes of all the time-domain modes common to both the detectors, while $\boldsymbol{x}_{uncorr}^{\|/\perp}$ those specific to only one of them.

Owing to the orthogonality of the two polarization directions for the two detectors in a pair, we can represent their data with summed and differenced data streams, $\boldsymbol{d}^+, \boldsymbol{d}^-$, which contain the information about total intensity and polarized sky signals, respectively. These are defined as,

$$
\boldsymbol{d}^+ \equiv \frac{1}{2}(\boldsymbol{d}^{\|} + \boldsymbol{d}^{\perp}) \tag{6.4}
$$

$$
\boldsymbol{d}^- \equiv \frac{1}{2}(\boldsymbol{d}^{\|} - \boldsymbol{d}^{\perp}). \tag{6.5}
$$

Using (6.1) and introducing quantities specific to each of the new data streams, defined as,

$$
\boldsymbol{A}^{\pm} \equiv \frac{1}{2}(\boldsymbol{A}^{\|} \pm \boldsymbol{A}^{\perp}) \tag{6.6}
$$

$$
\boldsymbol{B}^+ \equiv \begin{bmatrix} \boldsymbol{B}_{corr}, & \boldsymbol{B}_{uncorr} \end{bmatrix}, \tag{6.7}
$$

$$
\boldsymbol{B}^- \equiv \boldsymbol{B}_{uncorr} \tag{6.8}
$$

$$
\boldsymbol{x}^+ \equiv \begin{bmatrix} \boldsymbol{x}_{corr} \\ \frac{1}{2}(\boldsymbol{x}_{uncorr}^{\|} + \boldsymbol{x}_{uncorr}^{\perp}) \end{bmatrix}, \tag{6.9}
$$

$$
\boldsymbol{x}^- \equiv \frac{1}{2}(\boldsymbol{x}_{uncorr}^{\|} - \boldsymbol{x}_{uncorr}^{\perp}), \tag{6.10}
$$

$$
\boldsymbol{g}^{\pm} \equiv \frac{1}{2}(\boldsymbol{g}^{\|} \pm \boldsymbol{g}^{\perp}), \tag{6.11}
$$

$$
\boldsymbol{w}^{\pm} \equiv \frac{1}{2}(\boldsymbol{w}^{\|} \pm \boldsymbol{w}^{\perp}), \tag{6.12}
$$

we can express these new data in a concise way as,

$$
\boldsymbol{d}^{\pm} = \boldsymbol{A}^{\pm}\boldsymbol{s}_{T/QU} + \boldsymbol{G}^d\boldsymbol{g}^{\pm} + \boldsymbol{B}^{\pm}\boldsymbol{x}^{\pm} + \boldsymbol{w}^{\pm}. \tag{6.13}
$$

Here $\boldsymbol{s}_T$ and $\boldsymbol{s}_{QU}$ denote sky signal vectors made of the total intensity and interleaved, pixel-by-pixel, Q and U Stokes parameters, respectively. These expressions emphasize that as intended each of the new data streams contains information either about the total intensity, $\boldsymbol{d}^+$, or polarization, $\boldsymbol{d}^-$. Moreover, as all the amplitudes appearing on the right hand side of these equations are specific for each data set, each of these two data sets can be analysed completely separately and, under the assumptions specified earlier, without any loss of accuracy. Specifically, the maps of the total intensity on the one hand, and the Q and U Stokes parameters on the other can be estimated independently. This is the approach we follow in this work.

We point out that a perfect separation of the total intensity and polarization information is strictly speaking only possible if the two detectors of each pixel pair are perfectly calibrated and have identical beams. Otherwise, some residual total intensity contribution may be present in $\boldsymbol{d}^-$ and, less harmfully, some polarization in $\boldsymbol{d}^+$. The beams of two detectors are more likely to be similar if they belong the same focal plane pixel. In any case, if needed, leakage from the total intensity to the differenced data, $\boldsymbol{d}^-$, can be modelled as a total intensity-like template and used in the map-making process, e.g. as proposed in Wallis et al. (2015). Though such tests were indeed performed as part of the analysis of the actual POLARBEAR data set (POLARBEAR Collaboration, 2014a), we do not consider them in the present work. Leaving aside this kind of systematic effect, it is mathematically equivalent whether we use one or the other data representation, as long as the filtered temporal templates, defined by $\boldsymbol{B}_{corr}$ and $\boldsymbol{B}_{uncorr}$, are used consistently.

In Eq. (6.13) the pointing matrices, $\boldsymbol{A}^+$ and $\boldsymbol{A}^-$, are given as in Eqs. (4.2) and (4.3) and therefore are composed of zeros and ones for the summed data and have two non-zeros per row given by $\cos(2\varphi_t^{\|})$ and $\sin(2\varphi_t^{\|})$ for the differenced data. The ground template operator, $\boldsymbol{G}^d$, is the same for the summed and differenced data. As an independent ground template is used for each detector pair and for each CES, the $\boldsymbol{G}^d$ matrix has as many columns as the number of bins used to discretize the ground signal and as many rows as the number of samples in a given CES. We bin the observed azimuths in intervals of the width of 0.08 deg and thus have $\mathcal{N}_{\boldsymbol{G}}^{CES,p} \sim 100$ bins per template. At any given time t the corresponding row of $\boldsymbol{G}^d$ has only one non-zero entry (equal to 1) in a column corresponding to the ground bin observed at this time.

The $\boldsymbol{B}^\pm$ matrices define the time-domain filtering applied to both data streams in order to suppress long-term correlations. In the POLARBEAR case the filtering is done subscan-by-subscan (POLARBEAR Collaboration, 2014a). Consequently, the $\boldsymbol{B}^\pm$ are block diagonal with one block per subscan, and each block displaying the same structure as in Eqs. (6.2) and (6.7). We denote such an elemental block with a subscript, s, to emphasize that we are referring to a single subscan. Each of these blocks removes from a given subscan time-domain trends given by time domain templates defined as polynomials up to some order, selected to ensure that the noise after filtering is nearly white. In our analysis, $\boldsymbol{B}_{corr,\,s}$ contains four templates (the Legendre polynomials up to the 3rd order, appropriately rescaled to become orthonormal over the time interval given by the subscan) and $\boldsymbol{B}_{uncorr,s}$ contains only the constant and linear templates. Consequently, the columns of $\boldsymbol{B}_{uncorr,\,s}$ are linearly dependent on those of $\boldsymbol{B}_{corr,\,s}$ and we restrict $\boldsymbol{B}^+$ to the latter ones, i.e., $\boldsymbol{B}^+ = \boldsymbol{B}_{corr,\,s}$ without any loss of generality, as discussed in Sec. 4.2.2. Moreover, for simplicity, for each subscan we use the orthogonalized version of the Legendre polynomials so matrices $\boldsymbol{B}^\pm$ are by construction column-orthonormal from the outset.

$\boldsymbol{w}^+$ and $\boldsymbol{w}^-$ are the noise terms, describing the noise in the data after the ground template and temporal trends removal. These noise terms are expected to be "prewhitened" with respect to the actual noise in the data. We model these vectors as white with inverse variances given by ω^+ and ω^- respectively. We allow for different weights ω for each CES and detector pair. These weights are evaluated from the actual data as the inverse of the average of the power spectral density of the timestreams sum and difference, taken between 1.04 Hz and 3.13 Hz.

It is now straightforward to generalize these considerations to multiple CESs and multiple detector pairs. In both cases we stack all the TOD for every detector pair and every CES together and, for concreteness, we do so for the summed and differenced data separately. The form of Eq. (6.13) for the concatenated summed and differenced data remains the same but the data objects and operators on its right-hand side need to be appropriately redefined. In particular, as we define a different ground template for each detector pair and each CES the global $\mathbf{G}_{all}^\pm$ matrix, will be block diagonal, with each block given by the detector-pair and CES specific matrix, $\mathbf{G}^d$. The vector of the ground template amplitudes, $\mathbf{g}_{all}^\pm$ will accordingly be made of the detector-pair and CES specific vectors, $\mathbf{g}^\pm$. Similar generalizations also apply to the temporal drifts term. However, in this case, one may need, or want, to account for effects that would be correlated between different detector pairs. Such effects could, for instance, be a result of contributions to the summed data due to atmospheric fluctuations. Consequently,

the ultimate filtering operator, $\mathbf{B}^{\pm}_{\text{all}}$, may not be strictly block-diagonal but have rather a form resembling that of (6.2). In the POLARBEAR data model, however, we do not include this possibility but introduce a separate template for each detector, each subscan and each polynomial order. This adds some flexibility that may permit better accounting for systematic effects, but it may not be always advantageous as far as statistical uncertainties are concerned due to the significant number of extra independent degrees of freedom this choice implies. The number of polynomial templates per CES and detector pair is $\mathcal{N}_{\mathbf{B}}^{CES,p} = \mathcal{N}_{sub} \times \mathcal{N}_{poly} \sim 150 \times \mathcal{N}_{poly}$, where $\mathcal{N}_{poly}$ is 4 or 2 if we are considering the sum or the difference of the detector pair. Consequently the total number of templates per CES and detector pair sum (resp. difference) is $\mathcal{N}_{\mathbf{T}}^{CES,p} = \mathcal{N}_{\mathbf{G}}^{CES,p} + \mathcal{N}_{\mathbf{B}}^{CES,p} \sim 700$ (resp. 400).

6.3   The implementation

We use the timestreams in the form of Eq. (6.13). Consequently the temperature and polarization problems can be solved independently. In the following we drop the $\pm$ superscript and specify if we are referring to the temperature of polarization case only when necessary.

   The implementation of the explicit unbiased map-maker can be divided in four main tasks.

1. The computation of the $\boldsymbol{A}^{\top} \boldsymbol{F_T} \boldsymbol{A}$ matrix.

2. The computation of the $\boldsymbol{A}^{\top} \boldsymbol{F} \boldsymbol{d}$ vector.

3. The inversion the $\boldsymbol{A}^{\top} \boldsymbol{F_T} \boldsymbol{A}$ matrix.

4. The application of $(\boldsymbol{A}^{\top} \boldsymbol{F_T} \boldsymbol{A})^{-1}$ on $\boldsymbol{A}^{\top} \boldsymbol{F} \boldsymbol{d}$.

Steps 1, 2, 3 are discussed in details in what follows. Step 4 does not deserve any explanation: it is just the application of a dense matrix on a vector. The computational cost is completely dominated by steps 1 and 3. Steps 1 and 2 both involve the filtering operator and consequently the orthonormalisation kernel of the templates, which is precomputed and used in both steps. We provide details about its precomputation in a dedicated section.

6.3.1   Filtering Operator $\boldsymbol{F_T}$

We describe here the implementation of the two different usages of the filtering operator: its application on a timestream and the evaluation of one of its entries. Both require the orthonormalisation kernel $\boldsymbol{K} \equiv (\boldsymbol{T}^{\top} \boldsymbol{T})^{-1}$, of which we describe the precomputation. The other relevant objects to be stored in the memory are the matrices $\boldsymbol{G}$ and $\boldsymbol{B}$. Since for any given row $\boldsymbol{G}$ is equal to one for the (unique) observed ground bin and zero elsewhere, we store it as a timestream of integers containing only the index of the non-zero elements. Also $\boldsymbol{B}$ is very sparse and is not stored explicitly, we rather keep in the memory a single copy of precomputed Legendre polynomial for all the required order and length. Since most the subscans have generally similar length, the memory required by $\boldsymbol{B}$ is limited. As a last important remark, the filtering operator inherits the block diagonal structure from the templates, with one block for every detector pair of every CES. Moreover, the corresponding blocks of the weight matrix $\boldsymbol{M}$ are equal to $\omega^{\pm}\mathbf{1}$ and, therefore, in the computation we can consider $\boldsymbol{M} = \mathbf{1}$ and multiply each of diagonal block by the corresponding ω. For the sake of clarity, all the following explanation refers to a single diagonal block. In order to extrapolate the global cost in term of memory and computational time, it is sufficient to multiply it by $\mathcal{N}_{CES} \times \mathcal{N}_{pair}$.

Orthonormalisation kernel $\boldsymbol{K}$

We first compute the inverse of the kernel and then the kernel itself. Since

$$(\boldsymbol{K}^{\pm})^{-1} = [\boldsymbol{G}, \boldsymbol{B}^{\pm}]^{\top}[\boldsymbol{G}, \boldsymbol{B}^{\pm}] \tag{6.14}$$

$$= \begin{bmatrix} \boldsymbol{G}^{\top}\boldsymbol{G} & \boldsymbol{G}^{\top}\boldsymbol{B}^{\pm} \\ (\boldsymbol{B}^{\pm})^{\top}\boldsymbol{G}^{\top} & (\boldsymbol{B}^{\pm})^{\top}(\boldsymbol{B}^{\pm}) \end{bmatrix} \tag{6.15}$$

the computation involves scalar products of timestream vector. We therefore loop over time, computing the product of templates and update the entry of $\boldsymbol{K}^{-1}$. These are, in principle, $\mathcal{N}_{t}^{CES} \times (\mathcal{N}_{\boldsymbol{G}} + \mathcal{N}_{sub} \times \mathcal{N}_{poly})^2$ operations. However, taking advantage from the sparse nature of $\boldsymbol{G}$ and $\boldsymbol{B}$, they are reduced to $\mathcal{N}_{t}^{CES} \times \mathcal{N}_{poly}^2$. Then, we compute the eigenstructure of $\boldsymbol{K}^{-1} = \boldsymbol{V}\text{diag}(\boldsymbol{e})\boldsymbol{V}^{\top}$ using the *dsyev* routine of

LAPACK (Anderson et al., 1999). We compute the regularised inverse by setting to zero the eigenvalues smaller than 10^{-6} times the larger eigenvalue and inverting all the others, and recomposing again the kernel by multiplying $\boldsymbol{V}$ and $\boldsymbol{V}^\top$ from left and right, respectively. The regularisation is particularly required because by construction we are in presence of the degeneracies described in Sec. 4.2.2. The scaling of the eigendecomposition is $(\mathcal{N}_T)^3$.

Summarising the computational cost of $\boldsymbol{K}^{-1}$ is $\mathcal{N}_t^{CES} \times \mathcal{N}_{poly}^2 \lesssim 5 \times 10^5$ operations, negligible compared to the cost of the inversion $(\mathcal{N}_T)^3$ which is roughly equal to 10^8 operations for temperature and 10^7 operations for polarization. The global cost is therefore approximately 10^{14} operations for temperature and 10^{13} for polarization. As far as the memory consumption is concerned, the $(\mathcal{N}_T)^2$ double precision values of the kernel are permanently kept in the memory, allocating 1.5 MB for polarization and 4 MB for temperature, for a total of 2 TB for polarization and 5 TB for temperature.

Application of the filtering operator

In order to apply the filtering operator on a timestream vector $\boldsymbol{d}$ we sequentially perform the matrix multiplication that composes the filtering operator

$$\boldsymbol{F_T}\boldsymbol{d} = \omega^\pm \{ \boldsymbol{d} - [\boldsymbol{G}, \boldsymbol{B}^\pm] \overbrace{\boldsymbol{K} \underbrace{[\boldsymbol{G}, \boldsymbol{B}^\pm]^\top \boldsymbol{d}}_{\text{(I)}}}^{\text{(II)}} \} \tag{6.16}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{(III)}}$$

Step (I) computes the scalar product between the templates and the data. It consists of one iteration over the timestream and, taking advantage of the sparsity of the templates, $1 + \mathcal{N}_{poly}$ multiplications for each of the $\mathcal{N}_t^{CES}$ samples. Step (II) applies the kernel, $\mathcal{N}_T^2$ operations, and the resulting vector is projected back to the time-domain in step (III). Also this last step—which moreover subtracts the result from the original data vector and apply the weighting—consists in few times $\mathcal{N}_t^{CES}$ operations. Summarising, the whole filtering operation scales like $\mathcal{N}_t^{CES}$ and is therefore negligible compared to the precomputation of K. The memory work-space required is also irrelevant.

Evaluation of a matrix element

The only non-trivial part in the evaluation of $(\boldsymbol{F_T})_{tt'}$ is to evaluate $\{[\boldsymbol{G}, \boldsymbol{B}^\pm] \boldsymbol{K} [\boldsymbol{G}, \boldsymbol{B}^\pm]^\top\}_{tt'}$. Again, we take advantage from that fact that, for any time t, the corresponding row of $[\boldsymbol{G}, \boldsymbol{B}^\pm]$ has just $1 + \mathcal{N}_{poly}$ non-zero elements. Therefore, the cost of the evaluation of a matrix element is $(1 + \mathcal{N}_{poly})^2$.

Considering the filtering matrix of the whole dataset, $(\boldsymbol{F_T})_{tt'}$ can be different from zero (and in general it is) only if t and t' refer to the samples of the same CES and pair of detectors. Therefore the non zero elements are $\mathcal{N}_{CES} \times \mathcal{N}_{pair} \times (\mathcal{N}_t^{CES})^2 \sim 10^{15}$.

6.3.2 Computation of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d}$

The non-trivial part is the application of the filter operator on the data vector. The time-domain filtered vector is never stored explicitly, instead, as its entry d'_t is calculated, we compute directly its contribution to $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d}$. It consists in adding d_t, $\cos(\phi_t)d_t$ and $\sin(\phi_t)d_t$ respectively to the I, Q and U Stokes parameters of the pixel being observed at time t.

6.3.3 Computation of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$

The basic idea of the algorithm is to iterate over all the non-zero elements of $\boldsymbol{F_T}$, a $\mathcal{N}_t \times \mathcal{N}_t$ matrix, and compute its contribution to the $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$, a $\mathcal{N}_p \times \mathcal{N}_p$ matrix for temperature and a $2\mathcal{N}_p \times 2\mathcal{N}_p$ matrix for polarization.

Consider a time sample t and call p the observed sky pixel. Since the only non-zero entries of the t-th row of $\boldsymbol{A}$ are the columns corresponding to pixel p, the t-th column of $\boldsymbol{F_T}$ contributes only to the p-th column of $\boldsymbol{F_T}\boldsymbol{A}$. Analogously the t'-th row of $\boldsymbol{F_T}$ contributes only to the p'-th row of $\boldsymbol{A}^\top \boldsymbol{F_T}$. Therefore, in order to build the $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ matrix we loop over the elements of $\boldsymbol{F_T}$: for the (t', t) entry

we compute its contribution to the (p', p) block of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$

$$\text{Temperature: } (\boldsymbol{F_T})_{tt'} \tag{6.17}$$

$$\text{Polarization: } (\boldsymbol{F_T})_{tt'} \cdot \left[\begin{array}{cc} \cos(\varphi_t)\cos(\varphi_{t'}) & \cos(\varphi_t)\sin(\varphi_{t'}) \\ \sin(\varphi_t)\cos(\varphi_{t'}) & \sin(\varphi_t)\sin(\varphi_{t'}) \end{array} \right] \tag{6.18}$$

Because of the symmetry of $\boldsymbol{F_T}$, we could add the same contribution to the (p, p') block of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$, without $(\boldsymbol{F_T})_{t't}$. Equivalently, we take advantage of this symmetry by evaluating the contribution to $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ only when $p \geq p'$. This was more convenient especially because the routine that we use for the computation of the eigenstructure of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ requires only the upper triangle of the matrix.

The computational cost of the calculation of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ scales as the number of non zero entries of $\boldsymbol{F_T}$ times the cost of its evaluation for given (t', t): $\mathcal{N}_{CES} \times \mathcal{N}_{pair} \times (\mathcal{N}_t^{\text{CES}})^2 \times (1 + \mathcal{N}_{poly})^2 \sim 10^{16}$ operations.

6.3.4   Inversion of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$

We then perform the eigendecomposition of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ representing it as,

$$\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A} = \boldsymbol{V} \text{diag}(\boldsymbol{e}) \boldsymbol{V}^\top. \tag{6.19}$$

This is done with help of a ScaLAPACK routine, `pdsyevr` (Blackford et al., 1997). The numerical cost is $\mathcal{O}(\mathcal{N}_p^3) \sim 10^{15}$ operations. This scaling relation is the main obstacle in the application of the explicit implementation to maps with a larger number of pixels. By construction the eigenvalues, $\boldsymbol{e}$, are all non-negative numbers though numerically some small eigenvalues may turn out to be negative. The inversion of this matrix is then performed by inverting its eigenvalues. Since the condition number of the matrix is typically very large, the inversion needs to be regularized by employing a (pseudo)inverse defined as,

$$(\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A})^{-1} = \boldsymbol{V} \text{diag}(\tilde{\boldsymbol{e}}) \boldsymbol{V}^\top, \tag{6.20}$$

where,

$$\tilde{\boldsymbol{e}}_i \equiv \begin{cases} \boldsymbol{e}_i^{-1}, & \text{if } \boldsymbol{e}_i > 10^{-6} \max_j \boldsymbol{e}_j; \\ 0, & \text{otherwise.} \end{cases} \tag{6.21}$$

6.3.5   Parallelism

There are basically two types of computations. The ones that compute objects from the time-domain information and linear algebra operations on dense pixel-domain objects.

In the second category the algebraic objects are distributed across the system according to the block-cyclic distribution of ScaLAPACK[2]. The map-domain square matrices ($\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$, its inverse and $\boldsymbol{V}$) are divided into square (or nearly square) blocks. We work with a square grid of processors. The size of the blocks is decided on the basis of the processor grid in such a way that each processor is assigned to one and only one block. We keep the same block structure for map-domain vectors, therefore involving only the first column of the processor grid.

In the first category computation is always done CES by CES, detector pair by detector pair. Each processor is responsible for a set of detector pairs belonging to several CESs. We could not assign an entire CES to a processor because the number of processors is comparable to the number of CESs. Different CES can have a considerably different number of detectors: having one (or few) CES per detector would result in a considerably uneven workload. A CES is typically scattered across 24 processors.

In the computation of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d}$, each processor computes the contribution of the detectors pairs it owns by storing a local copy of the full $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d}$. These processor specific copies are then gathered, summed and scattered on the processors' grid.

This method can not be used for the computation of the $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ matrix. The maximum number of pixels used in production run was $\sim 7.5 \times 10^4$, for polarization this means that storing $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ in the memory takes 180 GB, largely exceeding the average memory per processors and even the memory of each node. Therefore $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ is computed block by block: each processor computes its contribution to the block, all the contribution are summed and the result is sent to the processor that owns the final block.

---

[2] http://netlib.org/scalapack/slug/node75.html

6.4 Inputs

We summarise here all the information that have to be collected and the operations that have to be performed before solving for the algebra of the unbiased map estimator.

The timestreams have to be calibrated and the time domain cuts have to be performed, this includes the definition of the subscans. The resulting vector makes up the $\boldsymbol{d}$ vector.

In order to build the (compressed) pointing matrix, two more timestream-like vectors are required. For each time sample, they store the polarization angle and the index of the observed sky pixel.

In order to prepare for the ground pickup filtering, the boresight azimuth of each CES is discretised and a different index is given to each of the azimuthal bins. A fourth timestream-like array contains the index of the azimuthal bin being observed at each time sample. In our implementation, we also make sure that only the ground bins actually observed are included in the indexing. The regularisation of the kernel $\boldsymbol{K}$ before its inversion makes this procedure unnecessary. However, removing unobserved ground pixels from the indexing considerably reduces the size of the block of the orthonormalisation kernel, which is the bottleneck of the memory consumption.

In order to define the matrix $\boldsymbol{M}$, the weights ω have to be precomputed. These weights are evaluated from the actual data as the inverse of the average of the power spectral density of the real data sum and difference, taken between $1.04\,\mathrm{Hz}$ and $3.13\,\mathrm{Hz}$.

Finally, we have to perform the pixel selection. Traditionally this is done by setting a maximum threshold on the condition number of the blocks of $\boldsymbol{A}^{\top}\boldsymbol{M}\boldsymbol{A}$. We apply a first cut based on this criterion, discarding pixels with a condition number higher than 10^6. However, the reconstruction of some of the sky pixel can still be unsuccessful since it can only be ensured by the study of $\boldsymbol{A}^{\top}\boldsymbol{F_T}\boldsymbol{A}$, which is still unknown at this stage. We indeed saw that applying the selection purely on the basis of the $\boldsymbol{A}^{\top}\boldsymbol{M}\boldsymbol{A}$ matrix produced degenerate involving few pixels at the very boundaries of the patch. We empirically saw that imposing a lower limit on the number of different CESs observing a pixel was effective in solving these pathologies. Imposing a threshold of 5 CES was usually enough, though we typically use a threshold of 10. We occasionally used very high thresholds, ~ 120 CESs, but in these cases the choice was typically driven by the need to limit the total number of pixels, not by the pathologies at the boundaries. Once the pixel selection is done, it is propagated back to the time domain, excluding from the TOD the samples recorded in correspondence with the rejected sky pixels.

6.5 Validation of the implementation

We have performed multiple tests in order to validate and verify the map-making code. For validation, we test whether the filtering operator, $\boldsymbol{F_T}$, removes all the unwanted modes as desired; for verification we perform a number of full, end-to-end runs of the code, testing that the various outputs have the expected statistical properties.

6.5.1 Effectiveness of the filtering

Our validation test checks that the filtering operator $\boldsymbol{F_T}$ satisfies the relation $\boldsymbol{F_T}\boldsymbol{T}\boldsymbol{y} = 0$, for some vector of template amplitudes $\boldsymbol{y}$. Since we are interested in map domain residuals, we actually test whether

$$\boldsymbol{s}^{\mathrm{res}} = (\boldsymbol{A}^{\top}\boldsymbol{M}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{F_T}\boldsymbol{T}\boldsymbol{y} = 0. \tag{6.22}$$

We separately test the ground pick up filtering, $\boldsymbol{T} = \boldsymbol{G}$, and the polynomial filters, $\boldsymbol{T} = \boldsymbol{P}$. In the former case, we produce a simulated ground signal timesteam as follows. For every CES, there is an index i for each ground template bin. i ranges between 0 and ~ 100. For each CES we set the amplitude of the ith ground template to $y_i^G = 1 + 0.01 \times i$, while the entries of $\boldsymbol{y}$ corresponding to the polynomial templates are set to zero. We obtain a simulated data vector $\boldsymbol{d}^{\mathrm{ground}} = \boldsymbol{T}\boldsymbol{y}$, where $d_t^{\mathrm{ground}} = y_{i_t}^G$. This construction has been devised in order to ensure that the simulated timestream is a linear combination of all the ground templates, with elements that are of order unity and are always positive. This last condition mimics the worst-case scenario of a ground signal that is coherent in time. In a very similar fashion, we test the filtering operator on the polynomial filters by setting $d_t^{\mathrm{poly}} = 1 + 0.01 \times i$, where now i is the index of the swipe in azimuth within the CES (remember that we have a set of polynomial templates for each constant direction azimuthal glide). We find that the map domain residuals, $\boldsymbol{s}^{\mathrm{res}}$, never exceed 10^{-6} for temperature and 10^{-8} for polarization. These levels are expected given the precision of our filter orthogonalization procedure, which tends to be merely approximate for very short subscans. They are however negligible for any practical purposes.
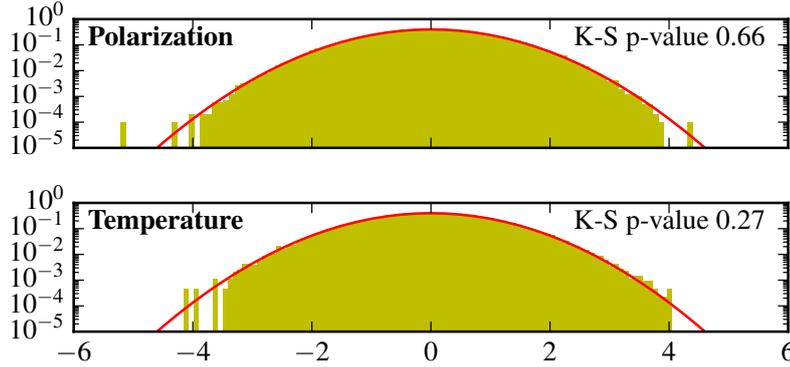
Figure 6.2: Histogram of the whitened unbiased map: $\boldsymbol{V} \text{diag}(\sqrt{e})\boldsymbol{V}^\top \hat{\boldsymbol{s}}$, where $\hat{\boldsymbol{s}}$ is estimated starting from a time-domain white noise simulation (see Sec. 6.5 for the details). The excellent compatibility of the distribution with the normal distribution shows that the $(\boldsymbol{A}^\top \boldsymbol{F_T A})^{-1} = \boldsymbol{V} \text{diag}(\boldsymbol{e}^{-1})\boldsymbol{V}^\top$ matrix (explicitly computed) correctly reproduces the covariance properties of the unbiased map, including the correlations due to the time-domain filtering.

6.5.2   Accuracy of the covariance of the estimator

As part of our end-to-end verification tests, we study the statistical properties of the noise-only maps produced by our map-making code. In this case, we produce the simulated noise-only stream in time-domain, with properties described by the diagonal weight matrix, $\boldsymbol{M}$, and processed it via our code. The output map was then prewhitened using the square root of the theoretically expected covariance, $\boldsymbol{A}^\top \boldsymbol{F_T A}$. The result was then histogrammed, Fig. (6.2), and compared to a Gaussian with unit variance. The agreement was found to be very good, e.g., the Kolmogorov-Smirnov test found p-values of 0.66 for polarization and 0.27 for temperature.

6.5.3   Accuracy of the estimator

Other examples of end-to-end tests involve a direct comparison of the known input map with the reconstructed map. As discussed in Sec. 6.6.2, the overall agreement is found to be excellent, compatible with the finite numerical precision.

6.6   A simulated noise-free reconstruction

In this section, we present the typical results produced by our unbiased map-making code on a simulated run. The realism of this benchmark comes from the use of the real pointing information and ω factors of POLARBEAR.

We create a simulated CMB sky using the `synfast` tool of the HEALPix package (Górski et al., 2005) and assuming the *Planck* best-fit parameters Planck Collaboration (2015j)[3]. The simulated sky and our pixel selection are reported in Fig. 6.3. We produce a signal-only timestream by using pointing and polarization angle information from the first and second seasons RA23 observations. These do not necessarily coincide with the dataset used for the cosmological analysis since we are only interested in a realistic scanning strategy.

As previously explained a key product of our map-maker is the eigenstructure of the $\boldsymbol{A}^\top \boldsymbol{F_T A}$ matrix. In Sec. 6.6.1, we describe an analysis of its eigenvalues and eigenvectors. Then we focus on the reconstructed map, the final product of our map-maker, and compare it the one produced by our iterative unbiased map-maker and a dedicated implementation of the unbiased estimator.
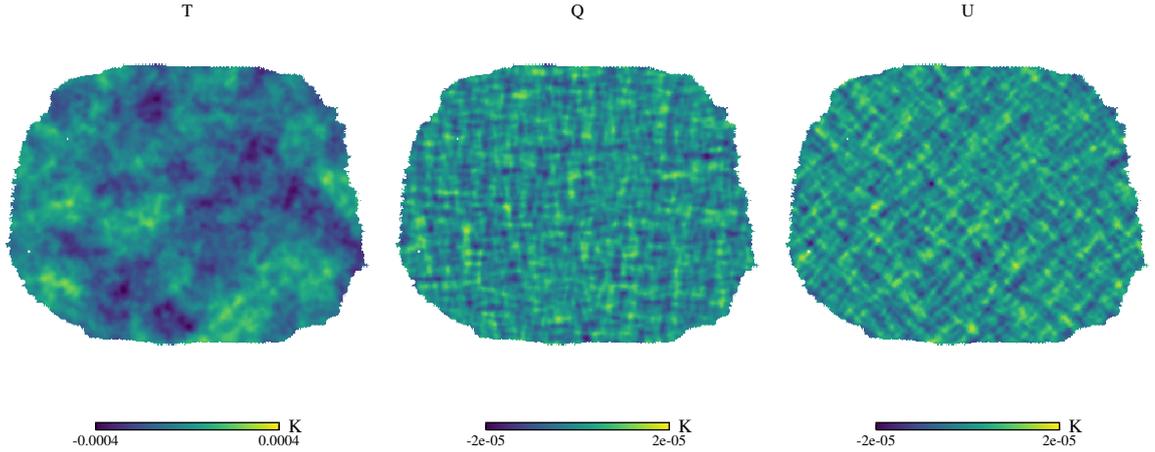
Figure 6.3: Maps of the input sky used here for the reconstruction comparison of the different map estimators.

6.6.1 Eigenstructure of $\boldsymbol{A}^\top \boldsymbol{F_T A}$.

The eigenstructure of $\boldsymbol{A}^\top \boldsymbol{F_T A}$ not only determines which modes are missing from the final unbiased sky estimate, it also provides information about the modes, which, while not singular, are not well constrained by the data. This is because the matrix is closely connected to the noise covariance in the pixel domain, $\boldsymbol{N_p}$, as shown by Eqs. (4.25) and (4.26).

In the case with no filtering at all the matrix, $\boldsymbol{A}^\top \boldsymbol{F_T A}$, reduces to the well-known $\boldsymbol{A}^\top \boldsymbol{M A}$, which, for the diagonal weights assumed here, is block diagonal with 2-by-2 blocks describing the (weighted) coupling between Q and U Stokes parameters in each pixel. The off-diagonal elements of these blocks will typically be negligible for pixels observed with a sufficiently homogeneous distribution of polarization angles, while the diagonal elements will be approximately equal to the eigenvalues of the 2-by-2 block. These are essentially given by the number of observations per pixel and their corresponding eigenvectors are spatial modes equal to zero everywhere but in the given pixel. Departure from such behaviour would then indicate the presence of strong off-diagonal coupling in some of the pixels.

The spectrum of $\boldsymbol{A}^\top \boldsymbol{M A}$ can, therefore, be used as a good reference for assessing the impact of the filtering on the map-domain noise spectrum. We compare the eigenstructure of the matrices, $\boldsymbol{A}^\top \boldsymbol{F_T A}$ and $\boldsymbol{A}^\top \boldsymbol{M A}$, in Fig. 6.4. In both cases, $\boldsymbol{A}^\top \boldsymbol{F_T A}$ and $\boldsymbol{A}^\top \boldsymbol{M A}$ have very similar spectra with the exception of tens of poorly constrained modes, defined as those with eigenvalues smaller by at least five orders of magnitude than the maximum eigenvalue. The corresponding eigenvectors, Fig. 6.6, are long modes and, in many cases, they exhibit a striped structure close to the boundaries. We interpret these modes as the result of the ground signal filtering discussed in Sec. 4.2.4 as their number roughly corresponds to the number of bins in the ground-bin. Indeed, following the discussion of Sec. 4.2.4, we expect that with every ground-template bin there should be an associated ill-constrained mode, corresponding to an offset of the constant declination strip swept by the azimuth range of the bin during the time of a single constant elevation scan. The fact that the recovered eigenvalues are not numerically zero demonstrates that these degeneracies are weakly broken as expected given that the presence of sky pixels observed with the telescope orientation corresponding to two different ground-template bins. Consequently, this leads to only one truly degenerate mode per each Stokes parameter map, see Fig. 6.5.

We also see that the two most singular eigenvalues of the polarized case are significantly larger than the most singular eigenvalue in the case of temperature. The difference is at least in part due to the numerical precision of the Computations, however it is also consistent with our earlier expectation that the polariser angle change across the constant elevation sweep can break the degeneracy between the sky signal offset for each strip and the amplitude of the ground template in the corresponding bin. In the case under study, given the limited primary mirror chop, the effect is very weak.

While regularizing the inversion of the matrix, $\boldsymbol{A}^\top \boldsymbol{F_T A}$, we remove these modes from the solution together with all the modes which are smaller than 10^{-6} of the maximal eigenvalue. This still leaves a significant number of ill-constrained modes in the estimated maps. Although they do not give rise to any

---

[3]We set the tensor-to-scalar ratio, r to zero, for definiteness, as the value of r is not relevant in the case considered here, given the focus of the first POLARBEAR campaigns on sub-degree angular scales.
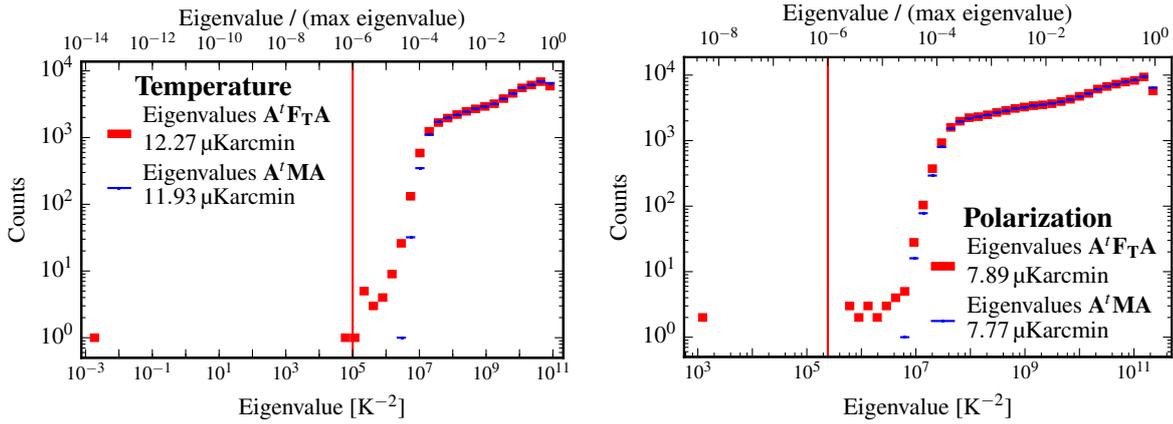
Figure 6.4: Eigenvalues of the blocks of the $\boldsymbol{A}^\top \boldsymbol{F_T A}$ and $\boldsymbol{A}^\top \boldsymbol{M A}$ matrices. The spectra of the two matrices are very similar in both temperature (left) and polarization (right). The qualitative difference is the presence of few degenerate modes and tens of poorly constrained modes. Temperature has a formally degenerate mode corresponding to the offset of the map. Polarization too is expected to have two nearly degenerate modes, one for each Stokes parameter. Though they are treated as singular because of numerical reasons, their degeneracy is partially broken, as explained in Sec. 4.2.4. In addition to these qualitative differences, filtering also produce a general decrease in sensitivity (see the legend), evaluated as the average over the whole observed area.

discernible artefacts in the case of the noiseless results as shown Fig. 6.8, when noise is included in the time-domain data, these modes may dominate the maps visual appearance. However, despite being noisy these modes are correctly estimated, and are neither artefacts of the estimator or its implementation, nor remnants of any incompletely-filtered parasitic signal, such as the atmosphere. Rather they reflect the actual uncertainty that the observation and filtering incur. These modes are typically missing in the unbiased maps derived with the iterative solver as well as in the biased maps. This is because these modes are either the most difficult to converge (in the case of the iterative solver) or are explicitly filtered out (in the biased map-maker). Consequently, although these maps may occasionally – and somewhat deceptively – look better, they will nonetheless be missing information which is correctly included in the unbiased map computed using the explicit solver.

The main sequence of eigenvectors are "pixel-like" modes, Fig. 6.7, in the sense that in each of these modes the most relevant structure involves a very limited number of pixels. As one intuitively expects, these pixels move from the boundary towards the centre of the patch as the eigenvalue of the mode grows (i.e. as the mode is better constrained). If there were no filtering, $\boldsymbol{A}^\top \boldsymbol{F_T A}$ would be block diagonal (it would be equal to $\boldsymbol{A}^\top \boldsymbol{M A}$) and therefore each eigenvector would correspond to exactly one pixel. Setting a threshold on the magnitude of the eigenvalues would be then equivalent to performing a selection of the best observed (here innermost) pixels. In the presence of filtering, the main effect of setting a threshold is still selecting the innermost pixels. However, because of the correlations that the filtering introduces, some signal is also removed from all over the map, affecting areas relatively far from the boundary pixels removed. This effect is visible at map-level in the lower panel of Fig. 6.7 and it is investigated at the power-spectrum level in Section 7.3 : removing modes that mainly involve pixels outside of the power-spectrum mask has an important impact on the power spectrum uncertainty and bias.

6.6.2   Reconstructed map

We assess the maps from the point of view of the fidelity with which they reconstruct the actual sky signal. Then, we also compare the maps reconstructed using both the biased and iterative unbiased estimators.

In Fig. 6.8 we show the reconstructed map as well as the residual with respect of the input map Fig. 6.3. The bottom panel of Fig. 6.8 shows the residual once the singular modes are removed from the input map and prove that the residual in the reconstructed map is only due to the presence of singularities in the system matrix, $\boldsymbol{A}^\top \boldsymbol{F_T A}$. As expected from our earlier discussion, Sec. 4.2, and confirmed by our results in Sec. 6.6.1, there are two such singular modes for polarization and one for temperature.

Eigenvalue 0 of 52115: $1.49 \times 10^{-3}\,\mathrm{K}^{-2}$ (1.49×10^{-14}) Eigenvalue 0 of 104230: $1.12 \times 10^3\,\mathrm{K}^{-2}$ (4.62×10^{-9})

Figure 6.5: *The degenerate modes* (the parenthesis in the title reports the corresponding eigenvalue divided by the largest eigenvalue). The degenerate mode of the temperature (left) map reconstruction is the global offset. The Q and U maps on the left are one of the two singular modes of polarization, the other mode is very similar: Q and U maps are swapped and the sign of one of them is flipped. The degenerate modes of the polarization are more complex because of the modulation of the polarization angle during the observation (see Sec. 4.2.4). For the same reason the degeneracy of these modes is partially broken (though they have to be treated as singular in order to preserve the numerical stability of the inversion of the $\boldsymbol{A}^{\top}\boldsymbol{F_T A}$ matrix).

The inversion regularization procedure removes these two modes from the estimated map, even if they contain actual sky signal. These singular modes result from the interplay between the scanning strategy and the filtering. They make the filtered data insensitive to these modes, which are unavoidably lost. Therefore, the map estimated from the unbiased estimator using the explicit solver may not be strictly speaking unbiased but provides an unbiased representation of all the modes that can be constrained from the available data given the chosen filtering.

Unbiased maps via the iterative solver

We compare these results with the those produced by the iterative implementation already available for the POLARBEAR data analysis.

In the case of the iterative solver, the system matrix, $\boldsymbol{A}^{\top}\boldsymbol{F_T A}$, is never explicitly constructed. Instead, we follow the general blueprint of maximum likelihood map-making code implementations (e.g. Cantalupo et al., 2010) and apply the factors defining the matrix from right to left. The filtering operator is applied as described above, again without being ever explicitly constructed.

We use a preconditioned conjugate gradient (PCG) technique (e.g., Golub and van Loan, 1996) with the preconditioner set to be $(\boldsymbol{A}^{\top}\boldsymbol{M A})^{-1}$, which is either diagonal, for the total intensity, or block-diagonal, for the polarization, and can therefore be straightforwardly computed, stored in the memory, and applied to a vector whenever needed. This is again the standard choice (e.g. Cantalupo et al., 2010). To quantify the convergence one typically uses the norm of the map level residuals,

$$r^{(i)} \equiv \frac{|\boldsymbol{A}^{\top}\boldsymbol{F_T A m}^{(i)} - \boldsymbol{A}^{\top}\boldsymbol{F_T d}|}{|\boldsymbol{A}^{\top}\boldsymbol{F_T d}|}, \tag{6.23}$$

where $\boldsymbol{m}^{(i)}$ stands for the solution after the ith iteration. Notice that r is a dimensionless quantity. A typical requirement for convergence would then be $r^{(i)} \leq 10^{-6}$.

Unlike the explicit implementation, the iterative solver is not applied to the full dataset. The constant elevation scans are divided into 267 nearly even groups and, using the iterative solver, one map $\hat{\boldsymbol{s}}_{\alpha}$ is obtained independently for each group α. Afterwards we coadd the maps as follows

$$\hat{\boldsymbol{s}} = \left[\sum_{\alpha}(\boldsymbol{A}^{\top}\boldsymbol{M A})\big|_{\alpha}\right]^{-1}\sum_{\alpha}(\boldsymbol{A}^{\top}\boldsymbol{M A})\big|_{\alpha}\,\hat{\boldsymbol{s}}\big|_{\alpha}, \tag{6.24}$$

Eigenvalue 16 of 52115: $1.14 \times 10^6\,\mathrm{K}^{-2}$ (1.14×10^{-5}) Eigenvalue 20 of 104230: $6.40 \times 10^6\,\mathrm{K}^{-2}$ (2.64×10^{-5})

Figure 6.6: *Examples of nearly degenerate modes* (the parenthesis in the title reports the corresponding eigenvalue divided by the largest eigenvalue). These modes are composed of some prominent feature at the boundaries and a (usually) weaker long mode. The structures at the boundaries correspond to sets of pixels that are heavily affected by filtering. For polarization the dominant effect is the ground removal. At the high and low declination ends of the observed area the redundancy of the observations is low and therefore the degeneracy breaking effects discussed in Sec. 4.2.4 are mild. For temperature, the high order of the polynomial filtering plays a significant role, adding prominent features at the boundaries at intermediate declinations and increasing the complexity of the long modes. Notice that the prominent features at the boundaries saturate the colour scale.

where as before $|_\alpha$ denotes a quantity computed using only the data belonging to subset α.

This divide-and-conquer approach is necessary because of the intrinsically serial implementation of the code: the estimation of the unbiased map for different α groups can be done in parallel but each of them has to take place on a single processor. This poses serious constraints on the volume of data that can be analysed simultaneously. The data and the map-making related objects have to be kept in the memory of a single processor, working with 24 CESs typically required to use only two of the processors available on the computational node. Moreover, in this case each iteration takes $\sim$5 minutes, setting a lower bound of $\sim$10 hours to the runtime of the program. This is typically a limitation on the modern supercomputers, which prefer short and massively parallel jobs. These features reflect the fact that the efficiency in the computation and the memory management was not the priorities of its design: iterative map-maker was devised to be well integrated within the whole POLARBEAR pipeline, which is intrinsically serial and written in python.

The product of this map-maker is shown in Fig. 6.9. The residual in this case is clearly more pronounced and complex (middle panel). One might expect that the result of the unbiased map estimator should be the same, whichever solver is applied. However, the result shown in the figure corresponds to an incompletely converged iterative solution. Indeed, we have found that the iterative solution residuals, (6.23), do not decay to zero, see the left panel of Fig. 6.10, but instead asymptote to between 1×10^{-4} and 1×10^{-3}, which is roughly ~ 3 orders of magnitude above our fiducial convergence criterion of 10^{-6}. This is the case even if we allow as many as a few thousand iterations. Such a behaviour is indeed expected in linear systems for which the system matrix is (numerically) nearly singular (Hanke, 1995; Szydlarski et al., 2014).

To understand the effects of this lack of convergence of the solver on the estimated maps, we have computed the (pseudo) power spectra of the estimated map after ith iteration, centre and right panels of Fig. 6.10. We see that although the very low ℓ part of the spectrum does indeed fail to converge, convergence is quickly reached in the intermediate and high ℓ-range. This again is consistent with the singular modes found in the explicit solver having only large angular scales. If the singular modes are known, we could readily remove them from the solution, and thus from the residuals, at each step of the iteration and restore the proper convergence. However, this typically would require as many computations as the explicit solver, undermining the most important advantage of the iterative one.

We can still use the PCG solver in such circumstances by using this practical workaround: instead of monitoring a single residual as given by (6.23), we track the behaviour of residuals at the scales of
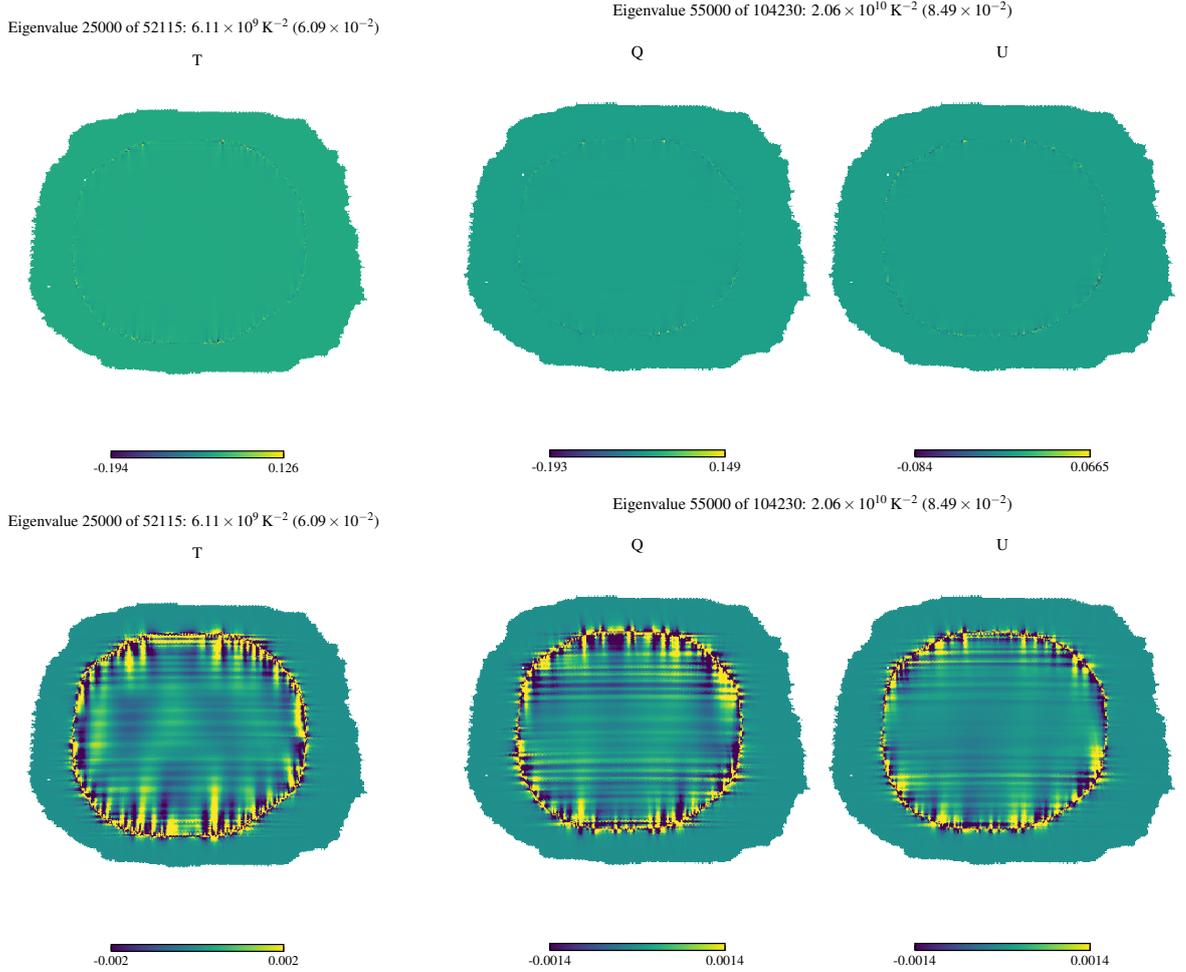
Figure 6.7: *Examples of pixel-like modes* (the parenthesis in the title reports the corresponding eigenvalue divided by the largest eigenvalue). These modes involve mainly a "ring" composed of a very limited number of pixels (see the first row). In the second row, the same eigenvectors are displayed with a colour scale squeezed by roughly two orders of magnitude, emphasizing structures inside and outside of the ring. Outside of the ring the structures quickly fade away. Also moving inwards the structures decrease their amplitude. However, compared to the outward structure, their typical length is much larger and, most important, they do not completely disappear: they have a relevant amplitude in the whole inner region. We emphasize that the "horizontal" structures represent the correlations induced by the ground template filtering.

interest for the power spectrum, as shown in Fig. 6.9. Nevertheless, we have to be aware of the fact that some of the power in the final solution may be compromised.

In the case of the map shown in the figure, convergence can indeed be reached in fewer than 100 iterations in the so-called science band defined in POLARBEAR Collaboration (2014a), which was the band of interest for the first round of the POLARBEAR papers.

Biased map-maker

An important term of comparison is the biased map estimator, Eq. (4.32). Because of its simplicity of implementation and low computational cost, this estimator is very popular in the analysis of current experiments. Also in POLARBEAR we implement this estimator, it is part of the fiducial pipeline adopted in the derivation of the results in POLARBEAR Collaboration (2014b); The POLARBEAR Collaboration (2014); POLARBEAR Collaboration (2014a, 2015). However, in the following comparison we will use a dedicated implementation of this estimator instead. The estimator we use strictly implements Eq. (4.32) given the data model in Sec. 6.2. The reason is that we want to compare the difference inherent to the unbiased and biased estimators, not on the extra small differences that the POLARBEAR implementation

has.

In our biased estimator Eq. (4.32) we adopt the very same $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{d}$ computation of our unbiased estimator and then compute and apply $(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}$. All these operations pose no issues given the block-diagonal structure of the matrix and the pixel selection procedure applied to the data beforehand, which ensures that each block of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ is invertible. The computational cost is then driven by the construction of the kernel $\boldsymbol{K}$ ($\sim 10^{13} - 10^{14}$ operations). As mentioned earlier, this kernel is block-diagonal and can be constructed and inverted efficiently on a single modern processing unit. Consequently, the entire estimator can be implemented and executed using serial or embarrassingly parallel programming models.

The biased map estimator, as expected, leads to the largest residuals (Fig. 6.11). These are particularly pronounced in the outskirts of the map, where the pixels crossing (and thus the filtering) may be highly anisotropic, but are still readily visible in the central part of the map, where the cross-linking and pixel sampling are better.

6.7   Summary

The new map-making formalism of Chapter 4 has been implemented in a flexible, massively parallel numerical code, which has been applied and validated on simulations mimicking all the important features of the POLARBEAR data set from the first and second observational campaigns. The applications of the code to real data is described in the next chapter.
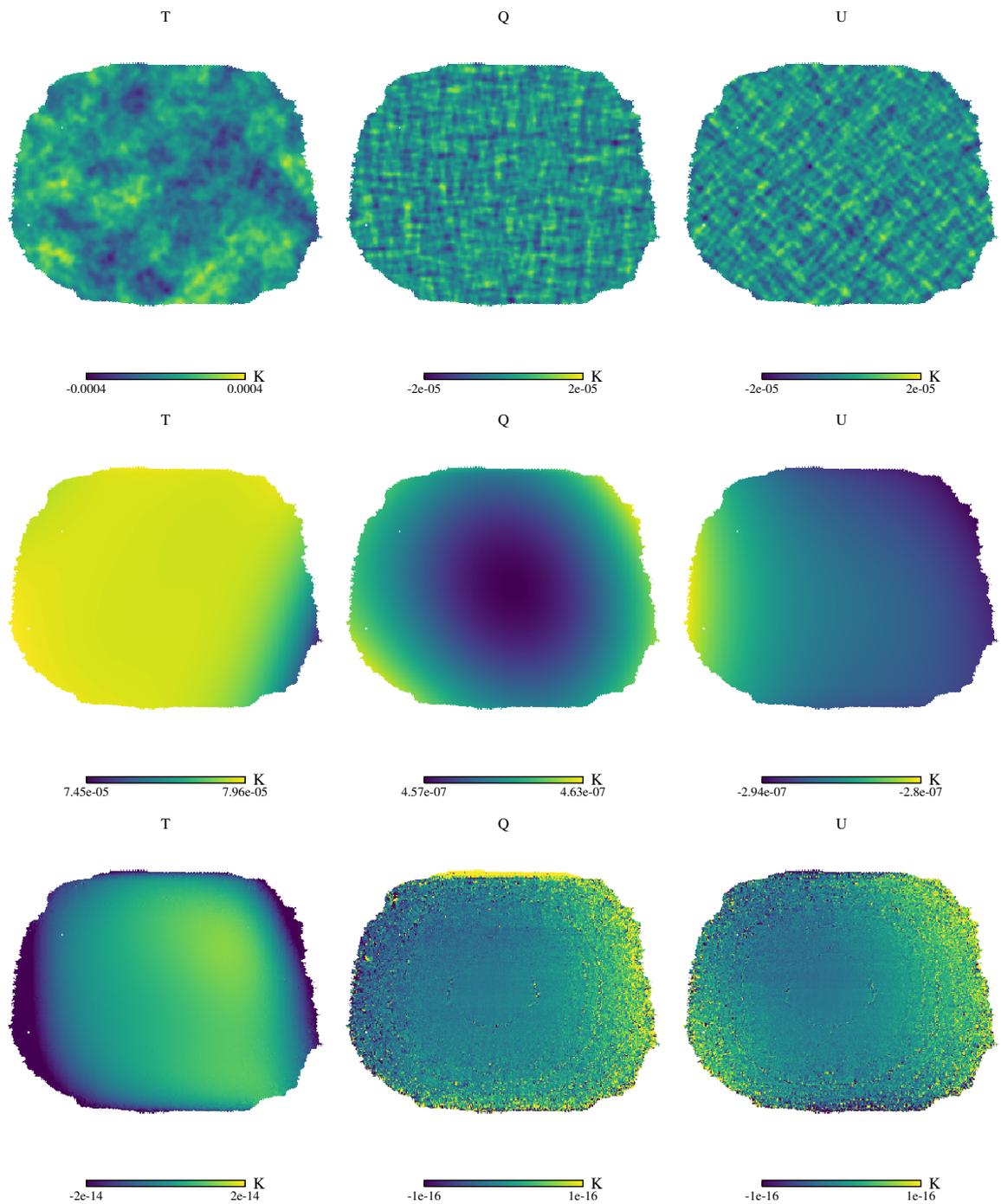
Figure 6.8: Maps derived with the explicit implementation of the unbiased map estimator. *Top row:* reconstructed maps. *Middle row:* difference between the reconstructed and the input maps. *Bottom row:* difference between the reconstructed and the input maps with the singular modes removed from the input maps.
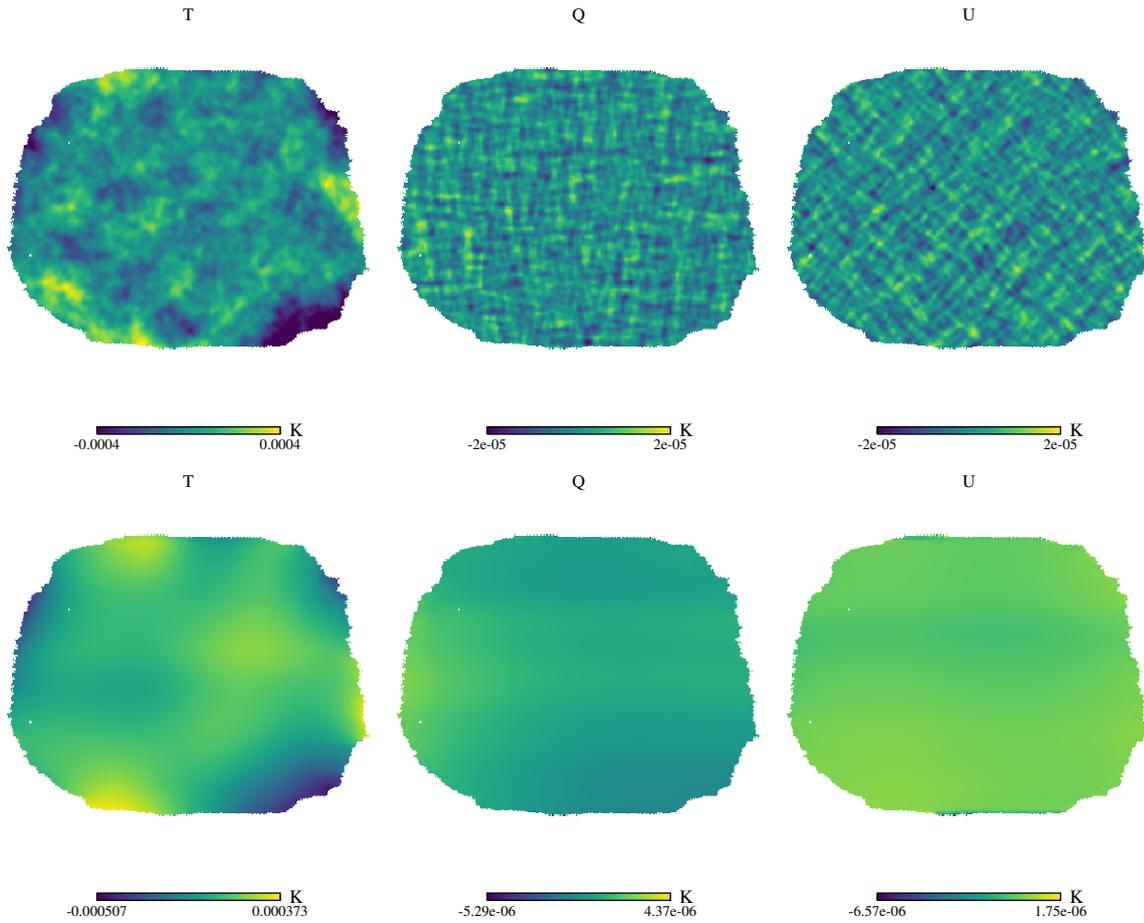
Figure 6.9: Results from the iterative implementation of the unbiased map estimator. *Top row:* reconstructed maps. *Middle row:* difference between the reconstructed and the input maps.
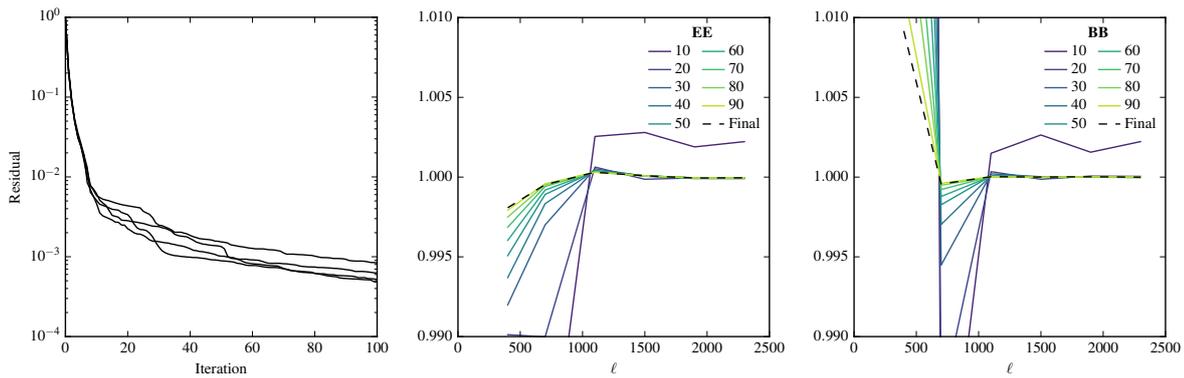
Figure 6.10: Visualization of different convergence criteria used in our iterative solver. The left panel shows that the standard residual, (6.23), saturates and does not converge to our fiducial 10^{-6} level in as many as 100 iterations. The middle and right panels show that this lack of convergence is due to the largest angular modes as the fractional difference between the power spectrum of the input map and the power spectrum of the ith map estimate in the multipole range $\ell \in [500, 2100]$ becomes quickly very small and reaches the level of better than 0.1% in fewer than $\sim 100 iterations$. This last observation has been used to set the convergence criterion used in the analysis of the first year POLARBEAR data set (POLARBEAR Collaboration, 2014a).
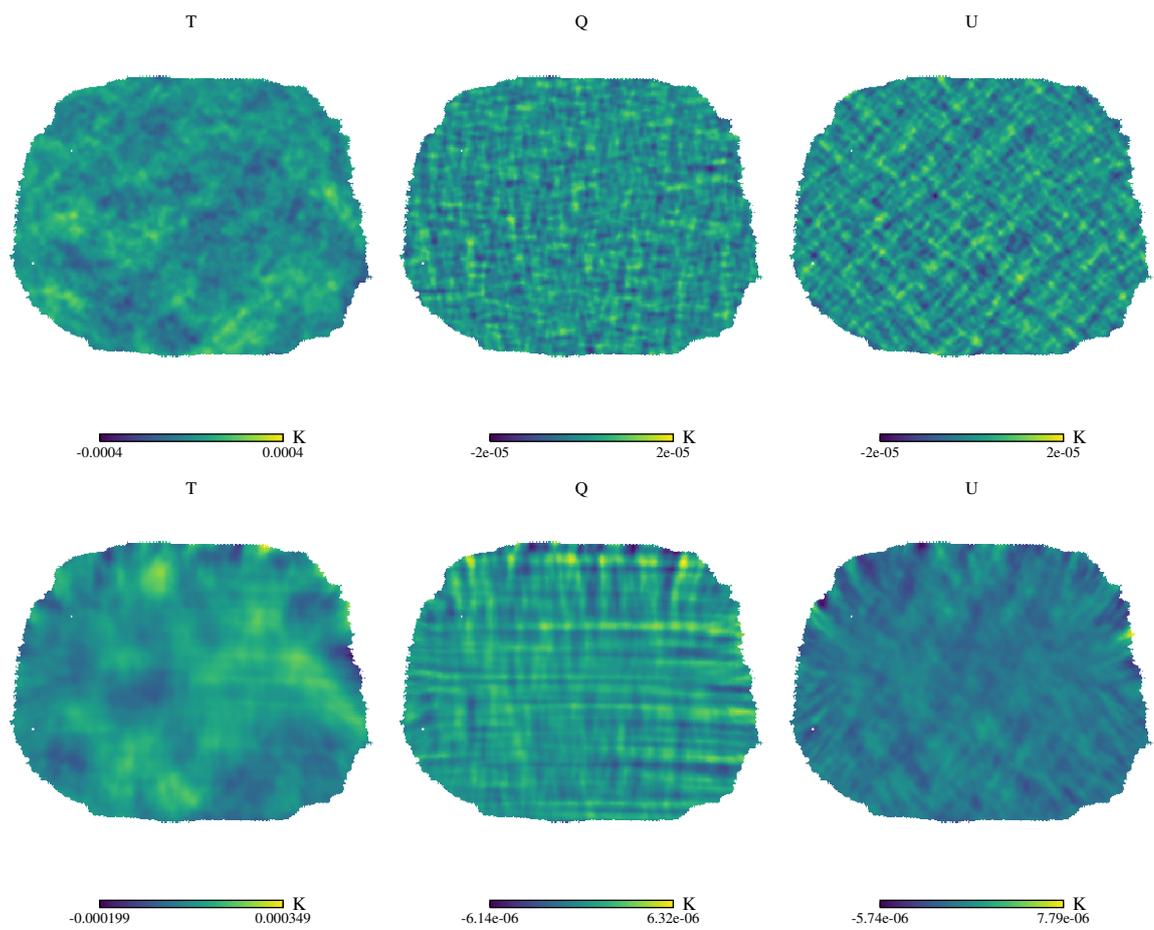
Figure 6.11: Map estimates derived using the biased map estimator, Eq. (4.32). *Upper row:* reconstructed maps. *Bottom row:* difference between the reconstructed and the input maps.

Chapter 7

Analysis of the Polarbear first and second season data

In this chapter I discuss the ongoing analysis of the Polarbear data from the first and second season of observation. The analysis should be concluded shortly, however at this time the results are not public.

The technique adopted in the analysis of the first season (Polarbear Collaboration, 2014b; The Polarbear Collaboration, 2014; Polarbear Collaboration, 2014a, 2015) employed a biased map estimator similar to Eq. (4.32) followed by a power spectrum estimator analogous to the MASTER procedure illustrated in Sec. 4.3.2. Because of its flexibility and low computational cost, in the current analysis the TOD selection and calibration are still performed with this pipeline. However, the actual data reduction (map-making and power spectrum estimation) is also performed using an independent pipeline devised by current and former members of the Polarbear APC team. This pipeline—subject of this chapter— combines the unbiased map-maker described in the previous chapter with X²PURE, a pure-pseudo-power spectrum estimator. My contribution to this stage of the project consisted in the development of the infrastructure around these two stand-alone codes. I also led the exploitation of the pipeline, validating its suitability for the analysis of the real data. As explained below, this required substantial amendments to the simple chaining of the two stand-alone codes.

The data set we are analysing is the one described in Sec. 5.4. The map-maker was described and validated in the previous chapter, I therefore start by briefly describing X²PURE and its validation in the Polarbear context. I then show two studies related to the power-spectrum estimation. The first one shows that the map domain correlations severely affect the power-spectrum uncertainty at large scales. In the second I discuss the need for a power-spectrum estimation based on the use of cross-spectra. The whole framework is validated by performing null studies on the real data. The focus of the described analysis is on the RA23 patch, the best observed of the Polarbear patches.

7.1 The Polarbear blindness policy

A distinctive feature of the Polarbear analysis workflow is that it conforms with the rules of the blindness policy. Blind analyses are popular in the particle physics community and were adopted in the CMB context by QUIET Collaboration (2011) and by Polarbear in the analysis of the first season. In blind analyses scientists are not allowed to look at the science product before the data sets has been defined, its quality assessed and the data analysis tools validated. The goal is to mitigate the observer bias: if the researcher has both access to the science products and the freedom to adjust some aspects of the analysis, he may subconsciously tune the free parameters to push the results towards his prior knowledge or some theoretical expectations.

The main science product of Polarbear is the BB power spectrum. The blind analysis is implemented with the help of the so-called *null tests*, which enable to assess the quality of the data set without ever looking at the BB spectrum of the data. A null test consists in deriving from the candidate science data set a quantity that has no expected contribution from the signal and check its compatibility with the expectation from the noise model.

In Sec. 8 we will describe in detail how this paradigm is implemented in Polarbear. Currently, the data set of the first and second season analysis has been defined, we are finalising the null test analysis
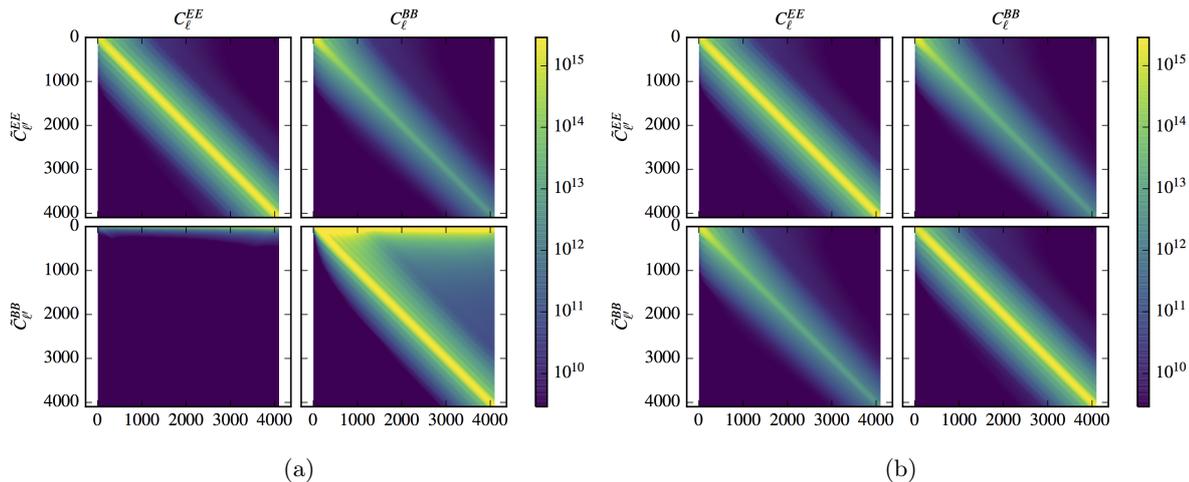
Figure 7.1: Example of polarization auto-spectra blocks of the mixing kernel $M_{\ell\ell'}$ for the hybrid (a) and the standard (b) formalisms.

and we are expected to unblind the BB spectrum soon. As a consequence, we did not yet look at any preliminary scientific results and they cannot be reported in the present work.

7.2   X^{2}PURE

The power spectrum estimation code that we adopt is X^{2}PURE (Grain et al., 2009). It implements the pure and hybrid formalism described in Sec. 4.3.1 and extends the approach of Smith (2006) to cross-spectra. This tool was already validated, tested and exploited in several works (e.g., Grain et al., 2012; Ferté et al., 2013, 2015; Planck Collaboration, 2016b; Krachmalnicoff et al., 2015). In this section we provide more details about our set-up and a brief demonstration of the code performances.

The weight function, W, that we adopt is the map of the temperature weights, i.e. the diagonal of the temperature block of $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$. The pure formalism requires the weight function to satisfy the Neumann-Dirichlet boundary condition and therefore we multiply W by an analytic apodisation function. The expression for the resulting window function at pixel i is

$$W_i^{(\mathrm{apod})} = \begin{cases} W_i \left(\frac{\delta_i}{\delta_c} - \frac{1}{2\pi} \sin(2\pi \frac{\delta_i}{\delta_c}) \right) & \delta_i < \delta_c \\ W_i & \delta_i \geq \delta_c \end{cases} \tag{7.1}$$

where δ_i is the angular distance of pixel i from the boundary and δ_c is the apodisation length. Given the small sky area observed, we found $\delta_c = 30\,\mathrm{arcmin}$ to be appropriate.

X^{2}PURE allows using both the pure and the hybrid formalism. In this work we make use of the latter. As a demonstration of the consequences of this formalism, in Fig. 7.1 we show an example of the polarization auto-spectra blocks of the mixing kernel Eq. (4.82) and, for comparison, we report the same matrix for the standard (i.e. non-pure) formalism.

As expected by the pure formalism, the $M_{\ell\ell'}^{BB,EE}$ matrix in Fig. 7.1a has very small (mostly negligible) values: the BB pseudo-power spectrum doesn't receive contribution from the EE power spectrum. On the contrary, in the standard case (Fig. 7.1b) the same block has notable values, especially at large scales. This is due to the non-orthogonality between the B and the E pseudo-basis. As discussed in Sec. 4.3.1, this effect is taken into account and corrected for, but the estimated spectrum will be affected by an uncertainty coming from the cosmic variance of the leaked E component. In Fig. 7.2b we show that the uncertainty on the BB power spectrum of the standard estimator is considerably higher than the hybrid case.

The $M_{\ell\ell'}^{BB,EE}$ block of the hybrid case is computed even if it should be analytically zero. The non-zero values typically involve the large scales and are due to the imperfect orthogonality between the B pseudo-basis and the E modes caused by pixelisation effects. As in the standard case, this effect increases the variance of the BB spectrum. However, for reasonably regular sky regions and properly apodised window functions the effect is by far subdominant.

(a)   (b)
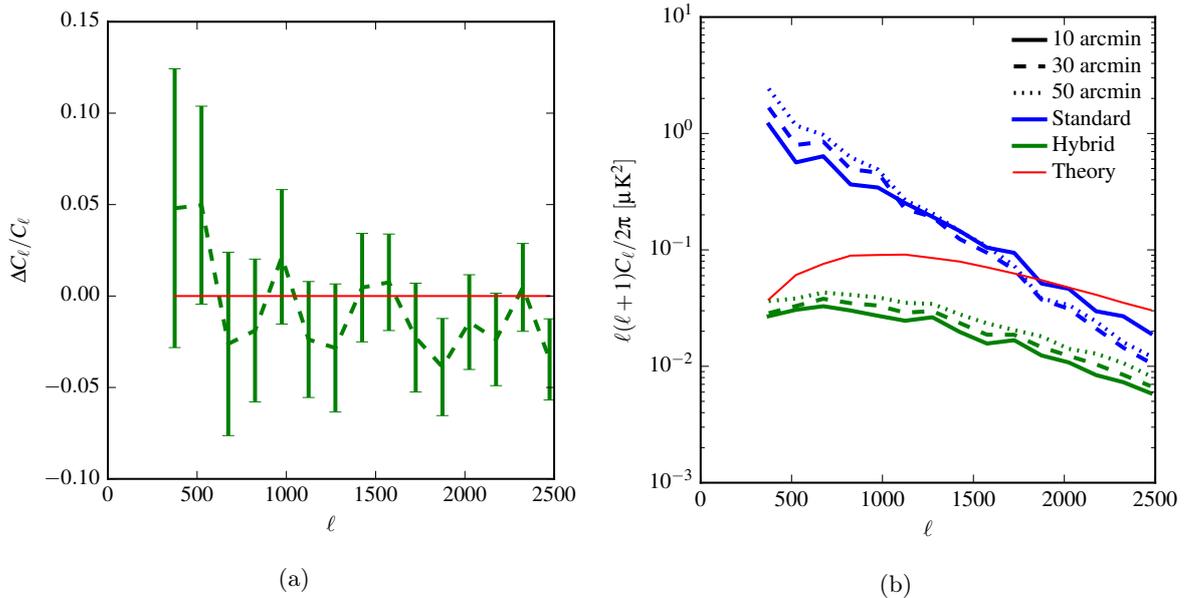
Figure 7.2: Example of the performances of X²PURE. (a) Fractional difference between the mean BB spectrum of 100 signal only simulations and the assumed theoretical power spectrum. (b) Standard deviation of the 100 signal only simulations, as an estimation of the cosmic variance, for different apodisation lengths. For comparison we report the same quantities for the standard pseudo-power spectrum estimator: they are severely affected by the cosmic variance due to E-to-B leakage. For reference, we report the theoretical BB power spectrum.

In the pure formalism, analogous statements can be made about the $M_{\ell\ell'}^{EE,BB}$ block. However, in the hybrid formalism the E pseudo-basis is not pure: it is the same pseudo-basis of the standard formalism and this block is indeed equal in the two cases.

7.2.1 Reconstruction of signal only simulations

In this section we show an example of the performances of X²PURE. We adopt a theoretical power spectrum defined by the *Planck* best-fit parameters (Planck Collaboration, 2015j) and use it for simulating 100 CMB realisations. We want to assess the performances of the tool in terms of fidelity of the reconstruction and E-to$B$ leakage, in comparison to the standard formalism. Trying to test realistic conditions of our usage, we adopt the mask and the window function actually employed for the analysis of the POLARBEAR RA23 patch.

In Fig. 7.2a we show that the mean BB reconstructed spectrum is compatible with the theoretical spectrum assumed in the simulations, demonstrating that the estimator is unbiased. Notice that the points present some correlations due to the fact that the size of the patch is too small for the binwidth of $\Delta\ell = 150$ that was adopted here, for demonstration purposes. Consequently, for the actual POLARBEAR analysis we use broader bins, with $\Delta\ell = 400$, and the estimated bin power are essentially uncorrelated. In Fig. 7.2b we show the standard deviation of the same spectra. Since every simulation is an independent sky realisation, the dispersion of the reconstructed spectra represents the cosmic variance of the single spectrum estimation. This cosmic variance includes both the BB intrinsic cosmic variance—due to the finite number of modes available—and the EE cosmic variance, which affects the BB estimation whenever the B pseudo-basis is not pure. This effect is visible in the figure: especially at large scales the standard estimator has higher variance than the hybrid formalism. This effect persists if larger sky areas are considered (see e.g. Grain et al., 2009; Ferté et al., 2013). In Fig. 7.2b the analysis is performed also varying the apodisation length of the window function (10, 30 and 50 arcmin). The performances of the estimator vary but, in all these cases, the difference is negligible compared to the total uncertainty level of POLARBEAR (as we discuss later). We therefore do not explore further possible optimisations related to the window function and its apodisation and adopt an apodisation length of 30 arcmin.

7.3   Computing the spectrum of biased maps

As it will be explained later, we have to be able to compute the spectrum of biased maps. In principle, we could implement the MASTER framework illustrated in Sec. 4.3.2. However, the iterative procedure for the computation of the F_ℓ functions and their integration in the power spectrum estimation require to modify X²PURE, which produces binned estimated power spectra from unbiased maps. In order to avoid this, we modify slightly the formalism. The key idea is simply to move the transfer functions at the level of the binned power spectrum instead of incorporating them in the computation of the binned mixing kernel.

We denote the power spectrum estimated by X²PURE as $\hat{C}_b^{\text{biased}}$. It is biased because we assume that we are applying the power spectrum estimator on a biased map. We then assume that the following relation holds

$$\begin{pmatrix} \langle \hat{C}_b^{TT,biased} \rangle \\ \langle \hat{C}_b^{EE,biased} \rangle \\ \langle \hat{C}_b^{BB,biased} \rangle \end{pmatrix} = \begin{pmatrix} F_b^{TT,TT} & & \\ & F_b^{EE,EE} & F_b^{EE,BB} \\ & F_b^{BB,EE} & F_b^{BB,BB} \end{pmatrix} \begin{pmatrix} C_b^{TT} \\ C_b^{EE} \\ C_b^{BB} \end{pmatrix} + \begin{pmatrix} \langle N_b^{TT} \rangle \\ \langle N_b^{EE} \rangle \\ \langle N_b^{BB} \rangle \end{pmatrix} \qquad (7.2)$$

Compared to the MASTER framework of Sec. 4.3.2, we are making an ansatz equivalent to Eq. (4.83).

The transfer functions have to be estimated using noise-free Monte Carlos. Similarly to the previous case, the temperature transfer function is estimated as $F_b^{TT,TT} = \langle \hat{C}_b^{TT,biased} \rangle_{\text{MC}}/C_b^{TT}$ where the $\langle \cdots \rangle_{\text{MC}}$ means average over a set of simulations with (known) theoretical power spectrum C^{TT}. For polarization, in order to evaluate $F_b^{X,Y}$, where X and Y can be either EE or BB, we perform simulations of CMB containing only the Y component and the transfer function is estimated as $F_b^{X,Y} = \langle \hat{C}_b^{X,biased} \rangle_{Y\text{-only MC}}/C_b^{Y}$. Also, the expected value of the noise bias $\langle N_b^X \rangle$ is evaluated as the average of noise-only MC. The unbiased estimator of the power spectrum can be finally obtained as

$$\begin{pmatrix} \hat{C}_b^{TT} \\ \hat{C}_b^{EE} \\ \hat{C}_b^{BB} \end{pmatrix} = \begin{pmatrix} F_b^{TT,TT} & & \\ & F_b^{EE,EE} & F_b^{EE,BB} \\ & F_b^{BB,EE} & F_b^{BB,BB} \end{pmatrix}^{-1} \left[\begin{pmatrix} \hat{C}_b^{TT,biased} \\ \hat{C}_b^{EE,biased} \\ \hat{C}_b^{BB,biased} \end{pmatrix} - \begin{pmatrix} \langle N_b^{TT} \rangle \\ \langle N_b^{EE} \rangle \\ \langle N_b^{BB} \rangle \end{pmatrix} \right] \qquad (7.3)$$

7.4   Filtering-induced correlations and the map-domain mode removal

In the following analysis we devise a power spectrum estimation procedure for the maps produced by our map-maker. For comparison, our analysis is conducted in parallel on the biased map estimator.

7.4.1   Set-up

Power spectrum estimation.   The $\hat{C}_\ell^{biased}$ spectra are evaluated using X²PURE, the bins in ℓ are the official POLARBEAR first season and second season bins, with $\Delta\ell = 400$ and centred at $\ell = 700, 1100, 1500, 1900$, plus a low ℓ and a high ℓ bins, which respectively have $\Delta\ell = 200$ centred at $\ell = 400$ and $\Delta\ell = 400$ centred at $\ell = 2300$. These two bins are considered just in the simulation studies of this section and will be excluded in the subsequent analyses of the real data.

As discussed before, the window function W that we use is the entire map of the temperature weights. However, the actual sky area retained for the power spectrum estimation is smaller than the full patch. As we explain later, we introduce a map-domain filtering that effectively selects the inner part of the map. The portion of the map considered by X²PURE is tailored on the most aggressive filtering. This pixels selection discards $\sim 50\%$ of the observed sky area but retains $\sim 95\%$ of the integrated weight.

Simulations.   Our analysis, as well as the estimation of the transfer functions, requires signal and noise simulated maps. We assume the same model and simulated data set as in Chapter 6.

As before, for the simulations of the CMB sky signal, we assume the power spectrum defined by the *Planck* best-fit parameters Planck Collaboration (2015j). We set the tensor-to-scalar ratio, r to zero, for definiteness, as the value of r is not relevant in the case considered here, given the focus of the first POLARBEAR campaigns on sub-degree angular scales. We synthesize the 'true sky' maps, denoted here as s^{sim}, using the synfast tool of the HEALPix package (Górski et al., 2005). To simulate the CMB-only sky maps as reconstructed with the explicit implementation of the unbiased map-maker we take the

simulated true sky maps, $\boldsymbol{s}^{sim}$ and remove from the simulated realization of the sky maps the eigenvectors corresponding to the eigenvalues set to zero in the $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ inversion regularization. We have validated that this is equivalent, algebraically and numerically, to first projecting the sky signal, $\boldsymbol{s}^{\mathrm{sim}}$, to the time-domain, $\boldsymbol{A}\,\boldsymbol{s}^{\mathrm{sim}}$, and then running the map-making procedure on the derived time-ordered data (see Sec. 6.6.2 and Fig. 6.8). For the biased map estimator we apply the operator, $(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$, directly to the generated, true sky maps, $\boldsymbol{s}^{\mathrm{sim}}$. Again this is algebraically and numerically equivalent to first producing the signal-only data streams, $\boldsymbol{A}\boldsymbol{s}^{\mathrm{sim}}$, and then applying the biased map-making operator, $(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{F}$, to them.

The noise-only maps reconstructed with the explicit implementation of the unbiased map estimator are computed directly in the pixel-domain as $\hat{\boldsymbol{n}} = \boldsymbol{V}\boldsymbol{z}$, where $\boldsymbol{z}$ is a pixel-domain vector of Gaussian random variables with variance given by the corresponding entry of $\tilde{\boldsymbol{e}}$, and $\boldsymbol{V}$ and $\tilde{\boldsymbol{e}}$ are defined in Eq. (6.20). This assumes that the weight matrix, $\boldsymbol{M}$, provides a correct description of the TOD noise, i.e., $\boldsymbol{M} = \boldsymbol{C_n}^{-1}$. Were this not the case, we would have to start from simulating a noise timestream, $\boldsymbol{n}$, with the desired noise properties and then process it with the map-making algorithm. For the biased map-making we follow the latter path even if the noise is uncorrelated. We therefore start by generating a timestream of uncorrelated Gaussian numbers of appropriate variance and projecting it into the pixel-domain using the corresponding map-making procedure.

Simulations containing both signal and noise are obtained by summing directly in the map-domain a signal-only and a noise-only simulation. Again, this is equivalent to summing noise and signal in the time domain and then applying the map estimator, it is a trivial consequence of the linearity of all map estimators considered.

7.4.2 Noise spectrum of unbiased maps

The maps generated by the different map-makers have different noise properties, in terms of both the noise amplitude and its correlations.

Fig. 7.3 shows the mean of the spectra produced by X²PURE from the noise-only simulations, measuring the noise bias. We then take the noise simulations in pairs and evaluate the uncertainty on the noise bias as the standard deviation of the cross-spectrum of the two noise maps within a pair.

In order to interpret these results, consider the unbiased map estimator (labelled with 10^{-6} in the figure). In Sec. 6.6.1 we illustrated the properties of the eigenstructure of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$—which is the covariance matrix of the map estimator—and have shown that the noise is correlated between pixels. In particular, in Fig. 6.7 we show that the typical uncorrelated mode has most of its power concentrated in few pixels and some minor contribution from a non-trivial structure that involves the whole inner part of the map. This latter feature is the most significant difference between the eigenstructure of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ and the one of $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$, whose eigenmodes involve only one pixel at a time. Consider a mode belonging to the noisy end of the spectrum of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$, most of the power carried by the mode is concentrated in some pixels at the boundaries of the patch. These pixels have huge statistical fluctuations but these are reasonably predicted by the $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$ matrix. The window function of X²PURE is based on the eigenstructure of $\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A}$ and in Sec. 4.3.1 we explained that the window function acts as a pixel-based weighting. Therefore, our power spectrum estimator properly handles the pixel-like component of the noisy mode but its correlated part is not downweighted by the power spectrum estimator. Therefore, even though it carries a minor fraction of the power of the noisy mode, it can be particularly pernicious because our estimator is suboptimal with respect to its statistical fluctuations. We identify this suboptimality as the cause of the noise increase at large scales compared to the usual ℓ^2 trend.

Most of the cosmological information contained in the noisy mode is nevertheless carried by the pixel-like component of the mode, which is downweighted by the power spectrum estimator. Therefore, the contribution of the mode to the power spectrum signal is marginal while it substantially increases the uncertainty, because of the correlated component, thus suggesting to discard the mode completely. The reasoning above is purely qualitative but suggests exploring the following mode removal procedure. We filter our unbiased map, progressively removing the noisiest eigenvectors of the $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ matrix: if the eigenvalue of a given eigenvector is less than α times the maximum eigenvalue, the mode is filtered out of the map. We consider several values of α: 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 0.05, 0.08, 10^{-1}. The 10^{-6} case corresponds to our unbiased estimator because the regularisation procedure of the $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ matrix effectively removes the modes that are below the 10^{-6} threshold. Notice that this procedure is similar to the rejection of the noisiest pixels in a map, the only difference is that we remove modes, because our noise is correlated.

Looking again at Fig. 7.3, as the noisiest modes are removed from the unbiased map (i.e. as α
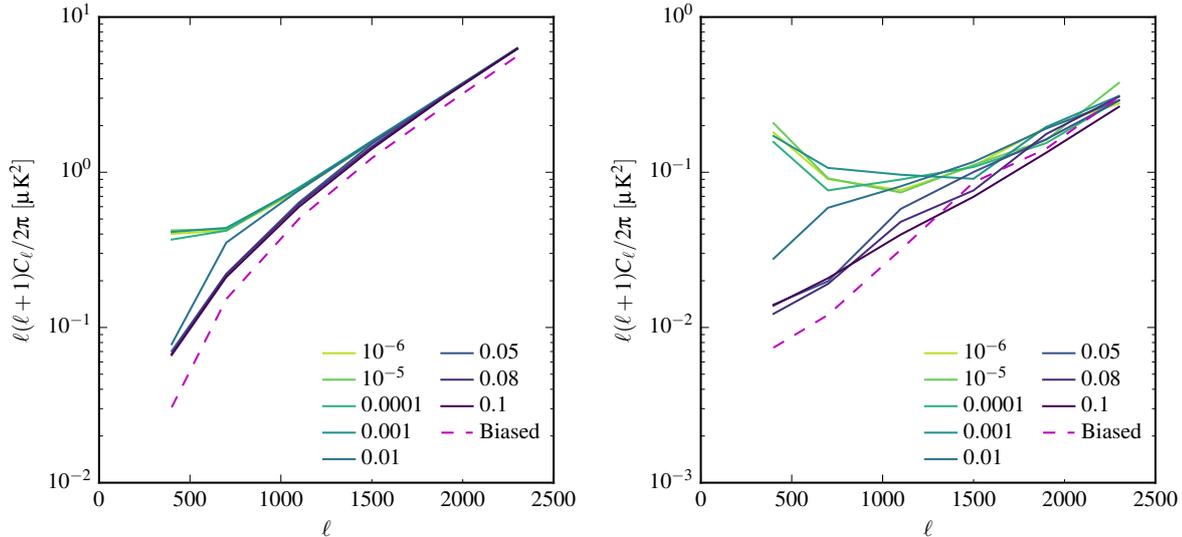
Figure 7.3: Mean (left) and standard deviation (right) of the B-mode power spectra (as computed by X^{2}PURE) of 100 noise only simulations for different input maps. "Biased" refers to the biased map estimator defined in Eq. (4.32). The other lines refer to different maps derived from the unbiased map estimator Eq. (4.29). The value in the legend is the α parameter that quantifies the amount of map domain filtering applied to the unbiased map. Higher values of α correspond to more aggressive filtering, see Sec. 7.4.2 for more details

increases) the power spectrum slowly converges to ℓ^2 behaviour. Notice that this low ℓ noise increase is not caused only by the "long modes" related to the ground template, it is actually dominated by poorly constrained "pixel-like modes": using the 10^{-4} threshold the long modes are removed but we still observe an important noise excess. This suggests that the cause is not the ground template marginalization but the polynomial filtering (or a combination of the two).

On the contrary, in the biased map the way noise is correlated does not cause any noise increase at large scales, both the mean and the standard deviation of the noise power spectra follow the usual ℓ^2 behaviour. Qualitatively speaking, we interpret this as a result of the presence of $(\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A})^{-1}$ in the unbiased estimator, compared to $(\boldsymbol{A}^\top \boldsymbol{M} \boldsymbol{A})^{-1}$: while the latter matrix only increases the power in pixels we little observations, the former also increases the power of the (correlated) modes that lost power because of the time-domain filtering.

We stress that in this section both the spectra derived from the biased map estimator and the ones derived from filtered unbiased maps are biased: we have to debias them as described in Sec. 4.3.2 before making quantitative statements about the uncertainty on their spectra.

7.4.3 Comparison of the performances

In Sec. 7.4.2, we consider the spectrum of noise-only simulations as computed by X^{2}PURE. The qualitative results support filtering the noisiest modes from maps produced with the unbiased map estimator. This causes a bias that percolates to the power spectra estimated by X^{2}PURE and, therefore, these spectra have to be debiased with the same procedure necessary for the maps issued by the biased map estimator.

As described in Sec. 7.3, in order to debias the spectra, signal only simulations are used to evaluate the transfer functions f^{XY} (Fig. 7.4) and noise only simulations for evaluating the noise bias Fig. 7.3.

We use these quantities to get an unbiased power spectrum estimator for each map estimator and apply it to 3 sets of 100 simulations:

- E modes only for evaluating the uncertainty due to E to B leakage;

- B modes only for evaluating the uncertainty B mode cosmic variance;

- noise only for evaluating the uncertainty due to the noise;
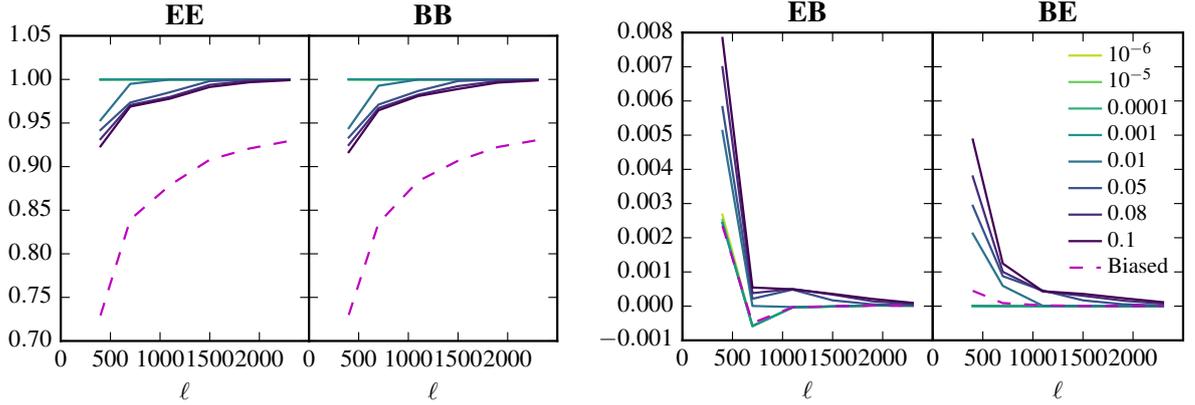
Figure 7.4: Transfer functions of the different map estimators. For the explanation of the legend, see Fig. 7.3

- E and B modes and noise for evaluating the total uncertainty on the BB spectrum (E to B leakage, BB cosmic variance and noise uncertainty).

When the simulations do not contain noise, the uncertainty is evaluated as the standard deviation of the simulations. When they contain noise, they are grouped in pairs and the uncertainty is evaluated as the standard deviation of the cross-spectrum of the two simulations of each pair.

E-to-B leakage and B-modes cosmic variance

Fig. 7.5a shows that, as expected from the pure formalism implemented in X^{2}PURE, the unbiased map estimator has basically no E-to-B leakage (the leakage is only due to pixelization effects).

However, the filtering operation that we apply on unbiased maps partially mixes E and B modes. When we filter a map containing only E modes, the result contains also a fraction of B modes. This effect is quantified by the off-diagonal blocks of the transfer functions in Eq. (7.2), reported in Fig. 7.4. The departure from zero is relevant for the biased map estimator only for the first bin and it is considerably more pronounced for $\alpha \geq 0.001$.

Eq. (7.3) account for this effect and yields an unbiased estimation of the spectrum. Nevertheless, this E-to-B leakage increases the uncertainty of the estimator. As shown in Fig. 7.5a, the leakage becomes relevant at large scales for any threshold higher than 0.001. The biased map estimator too has a significant amount of leakage at large scales but performs better than any threshold greater than 0.001.

As far as the BB cosmic variance is concerned, all the estimators considered perform equally well. In principle using a low threshold retains more sky area compared to the cases of with high threshold, thus they should have smaller sampling variance. However, in this analysis the sky area considered for the power spectrum estimator is always the same. Therefore the fact that the cosmic variance is roughly the same for all the values of α quantitatively confirms a qualitative statement made in Sec. 7.4.2: the correlated component of the pixel-like noisy modes does not carry any significant cosmological information while, as we show in the next section, adds substantial noise uncertainty to the power spectrum estimator.

Noise uncertainty

In Sec. 7.4.2 we have shown that, for different map estimators and different thresholds α, we get different dispersions of the raw spectra of noise simulations. However, the BB, BB transfer function in Fig. 7.4 shows that they also have different loss of BB power. Restoring this power boosts the spectrum of the noise too. The interplay between the two effects can be non-trivial.

However, Fig. 7.5c shows that the latter effect has minor impact: the dispersion of the unbiased spectrum of noise only simulations is still the higher the lower the threshold α and the lowest for the biased map estimator.

Because of the correlated nature of the noise in the estimated maps, the power spectrum estimator cannot properly down-weight the noisy modes. Their large fluctuations dominate the power at large scales, boosting the noise uncertainty. Filtering these modes out of the map alleviates this noise excess

at large scales and the lowest noise uncertainty is reached by the most aggressive filtering ($\alpha = 0.1$). The biased map estimator performs extremely well in this respect: despite the fact that its noise is correlated too, the uncertainty due to noise is comparable with the ℓ^2 trend expected by the uncorrelated noise case. We stress that the unbiased and biased map estimators preserve the same amount of information (we can convert one into the other anytime using an invertible linear operator). The disparity in their power spectrum noise uncertainty is purely due to the fact that we are using a suboptimal weighting for the power spectrum estimation.

Total uncertainty

Finally we consider the total uncertainty on the BB spectrum for the different map estimators. In Fig. 7.5d we show its spectrum while in Fig. 7.6 we express it as ratio of the lensing BB spectrum and total uncertainty.

   As far as the unbiased map estimator is concerned, given the specific noise level of these simulations, the noise plays a dominant role. Controlling its large-scale excess is more important than controlling the E to B leakage and, consequently, we find that the higher the threshold on the eigenvalues the lower the overall uncertainty is. We note that the most aggressive threshold removes more than 50% of the modes but retains more than 90% of the information (computed as the sum of the eigenvalues retained over the sum of all the eigenvalues)

   The biased map estimator has good noise level over the entire spectrum and, even if the estimator produces E to B leakage at large scales, the resulting uncertainty is below the noise level. As a result it performs substantially better than any other estimator studied here in the low ℓ part of the spectrum.

   We also investigate how the situation would change if the noise level was lower by extrapolating the total uncertainty assuming the observation time was x times longer. For each power spectrum bin, this total uncertainty is evaluated as

$$\sqrt{\sigma_S^2 + \left(\frac{\sigma_N}{x}\right)^2 + 4\frac{N}{x}S}, \tag{7.4}$$

where S and N are the mean power of the signal only (E and B) and noise only simulations respectively and the σs are their standard deviations. For the biased map estimator the uncertainty due to E to B leakage is smaller than the BB cosmic variance (see Fig. 7.5b). Therefore, the unbiased map estimator has superior leakage control but it is not the limiting factor in the case we are considering. Consequently, the factor x required for the unbiased map estimator to have better performance then the biased one is very large (about 10). We emphasize that this statement depends strongly on the specific case we are considering. The situation might be very different if the E to B leakage were to provide a more significant contribution to the overall uncertainty, as can happen when larger scales are probed. In such cases the unbiased map-making approach may be more readily favoured also from the power spectrum perspective.

Conclusions

The time domain filtering produces map-domain correlations. In particular, poorly constrained modes tend to have most of their power concentrated in boundary pixels but have nevertheless some power in the inner part of the map. The pseudo-spectrum-based methods can be significantly affected by the presence of such modes because they typically rely on the window functions which weight different sky pixels and are not capable of weighting such noisy modes. The result is a noise excess at low multipoles if they are left in the map, or enhanced E-to-B leakage if they are excised. We have developed a practical approach to compensate for such deficiency but found out that, at best, we can only match the performance of the quicker and simpler biased map-making. This approach consists in taking advantage of the eigendecomposition of $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ and deprojecting from the map all the eigenmodes smaller than 0.1 times the larger eigenvalue.

   We stress that these results are solely caused by the time-domain filtering adopted, and are not due to some incorrect data model. Indeed these results were derived only from simulations that assume the same data model adopted to devise the map estimator.

(a)

(b)

(c)

(d)
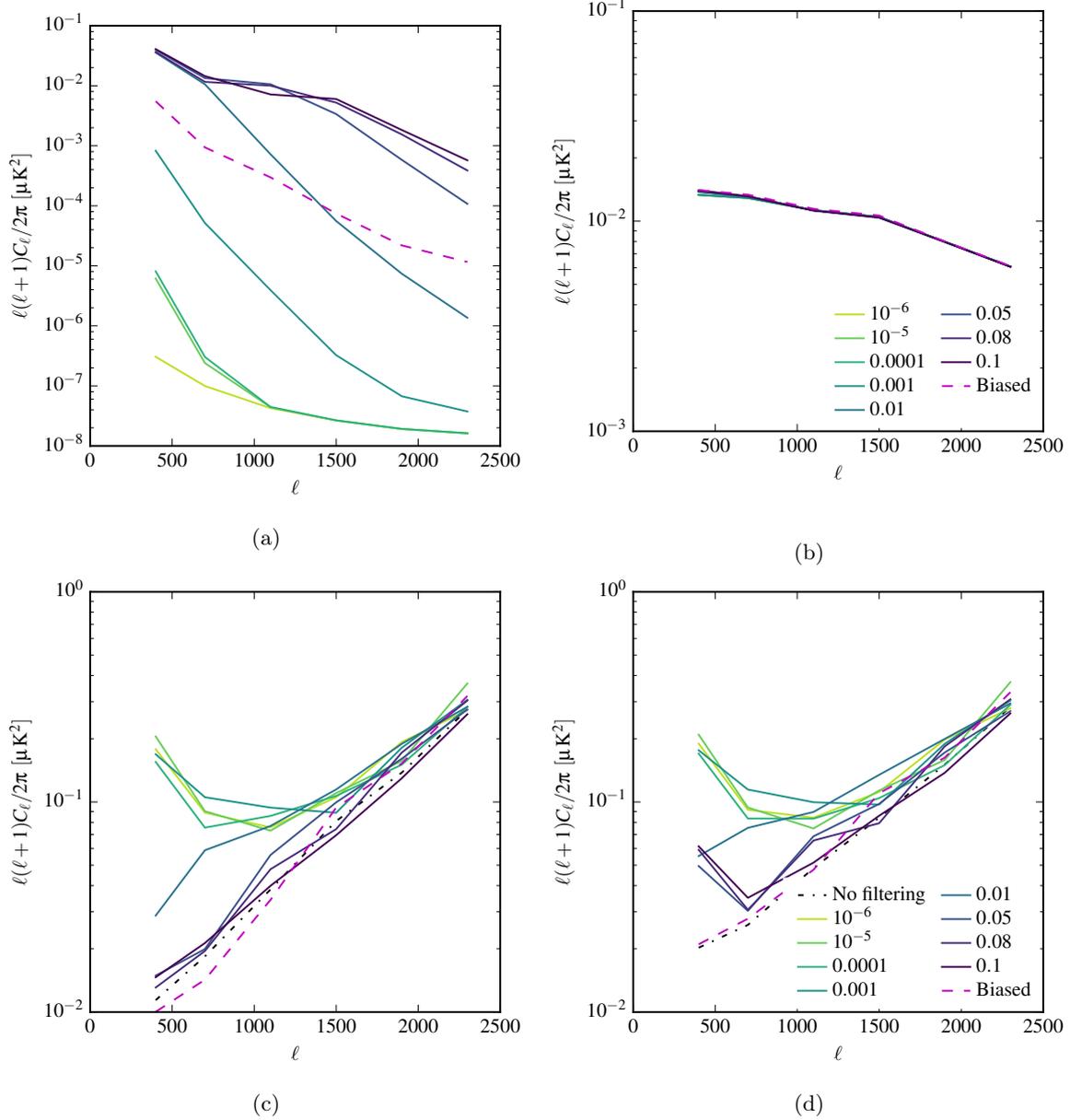
Figure 7.5: Uncertainty on the unbiased estimation of the B-mode power spectra based on 100 simulations containing E only (a), B only (b), noise only (c), E, B and noise (d). They respectively estimate the uncertainty due to E-to-B leakage, the B sample variance, the noise uncertainty and the total uncertainty on the BB power spectrum estimation. When containing noise, the spectra are estimated with help of cross-spectra of simulated pairs of maps. For the explanation of the legend, see Fig. 7.3. In the case of simulations containing noise we also display the "No filters" case, in which the simulated white noise TOD are not filtered: the simple map-making Eq. (4.5) is adopted.
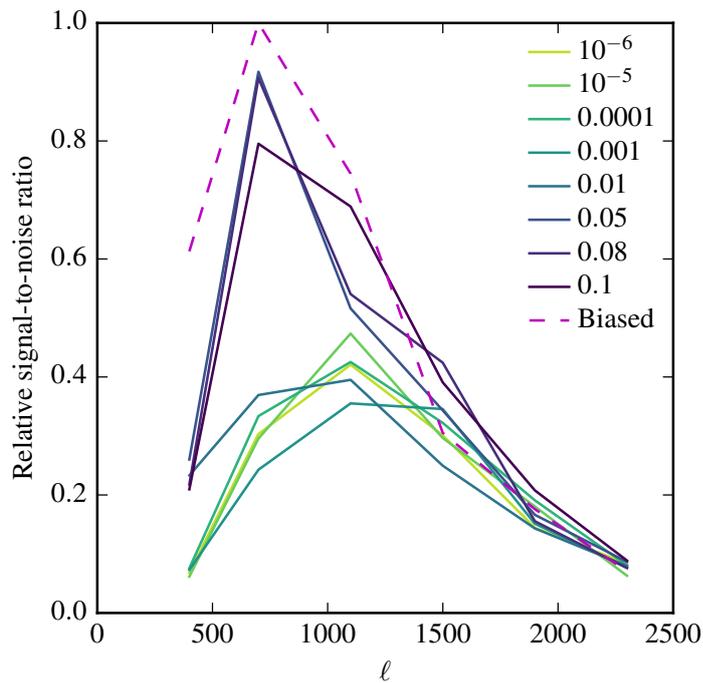
Figure 7.6: Ratio between the signal and the total uncertainty for B-mode power spectra derived with the different power spectrum estimation choices. The results are quoted relative to the highest ratio obtained for the biased map. The loss of precision of the spectra derived from the unbiased map is apparent and related to the strong correlated noise modes present in the map. The loss can be mostly recovered with help of a progressively more aggressive removal of the noisiest modes as expressed by the increasing value of parameter α as defined in the text and given in the legend.

7.5   The noise-bias problem and the cross-spectrum approach

In this section we use the POLARBEAR data to probe if our noise model is adequate for the POLARBEAR power spectrum estimation. We show that, while our noise model captures all the major features of the true noise in the data, it may not be sufficient to model the noise bias on the power spectrum level. We therefore resort to the cross-spectrum technique which bypasses this issue at some loss of statistical precision.

7.5.1   Probing the noise bias

The power spectrum estimation procedure described in Sec. 7.3 relies on the estimation through simulations of the expected value of the noise bias. Since POLARBEAR is still in the noise dominated regime (i.e. the expected lensing BB power spectrum is much smaller then the expected power spectrum of the noise), the accuracy of the noise bias subtraction is crucial for our BB power spectrum estimation.

In this section, we produce null maps from the data and from noise simulations, compute their spectra and show indications that the simulation-based noise-bias estimation might underestimate the actual noise bias of the data.

Data splitting

We divide our timestreams into 8 parts with approximately equivalent statistically weight. While splitting, we preserve the chronological order of the data (for example, all the data taken is split 3 were acquired before those of split 4). The actual metric adopted for assessing the equivalence of the slicing is the volume of the data set on disk. This choice is based on four motivations. First, we found it to be a good proxy (at 10% level) of the depth of the map produced by so-selected data. Second, the criterion is easily applicable a priory, without performing involving statistical studies on the data set. Third, it is not sensitive to the chronological density of observations and the medium-term yield variations, unlike more standard criteria like the number of days or the number of CESs per split. Last, this criterion ensures a good balance of the computational burden necessary to process the different splits, an operation that is performed in parallel.

Null maps

An unbiased map-making run is performed on each of the timestream splits, obtaining 8 unbiased maps $\boldsymbol{m}_i$. Then four null maps are derived from the 8 unbiased maps. We do this according to two different schemes. A first one, that we call "first v.s. second" subtracts from the maps of the first half the ones of the second half of the data: the four noise maps $\boldsymbol{n}$ are computed as $\boldsymbol{n}_0 = (\boldsymbol{m}_0 - \boldsymbol{m}_4)/2$, $\boldsymbol{n}_1 = (\boldsymbol{m}_1 - \boldsymbol{m}_5)/2$, etc. In a second scheme the null maps are produced from a map i and $i+1$, we refer to this scheme as "even v.s. odd". Notice that the null maps can be obtained only thanks to the unbiased map-making: if two maps were differently biased their difference would have some leftover signal.

Noise simulations

We produce 100 noise only simulations of each null map. We actually simulate noise only maps of each of the eight splits and combine them in the same fashion as the real data. As explained in Sec. 7.4, the noise-only simulations of each of the splits are produced directly in the map domain by taking advantage of the eigendecomposition of the expected covariance matrix—obtained from our unbiased map-making code.

Power spectrum estimation

Before computing the power spectrum of the null maps, we have to perform the map-domain filtering described and justified in Sec. 7.4. In principle, we should construct the covariance matrix of each null map, compute its eigenstructure and use it to filter the null map. In order to save on the (expensive) singular value decomposition of other four large matrices, we compute the $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ matrix (the inverse covariance matrix) for the full data set and use its eigenstructure to filter the four null maps with a threshold equal to 0.1. Thus, we perform only one extra singular value decomposition instead of four, also building the matrix is cheap, since we compute it by summing the (already computed) $\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A}$ matrices of each of the 8 splits. We decide to adopt this approximation not only for computational
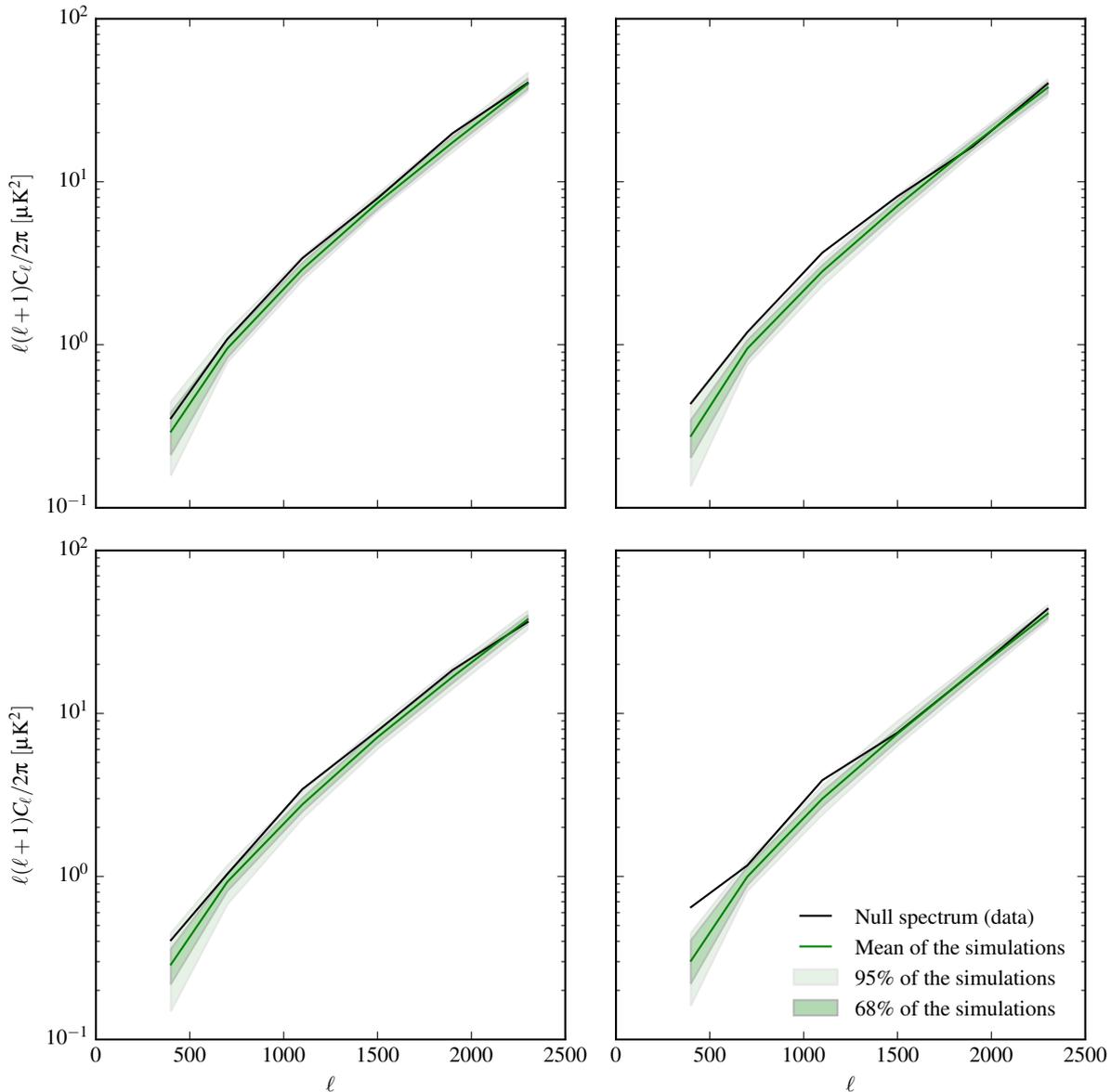
Figure 7.7: Comparison of the noise power in the four null maps derived with the "first v.s. second" scheme and the corresponding noise simulations.

convenience, but also for consistency with our final power spectrum estimation procedure (see Sec. 7.5.2 and Sec. 7.6).

We then use X^{2}PURE to compute the auto-spectrum of all the filtered maps.

Results

For each null map, we compare the auto-spectrum of filtered null maps from the data and from simulations. The mean of the latter corresponds to the evaluation of the noise bias and we expect it to be compatible with the auto-spectrum of the real-data null maps, within the scatter of the simulations. For each null map we report this comparison in Fig. 7.7 for the "first v.s. second" case and in Fig. 7.8 for the "even v.s. odd case". Considering each case individually, the discrepancy between data and simulations is not very significant. However, a global look at all the plots suggests that the data noise spectrum might be systematically higher than the one of the simulations.

Given the stringent requirements we have on the noise bias estimation, we resort to the cross-spectra technique as introduced in the next section.
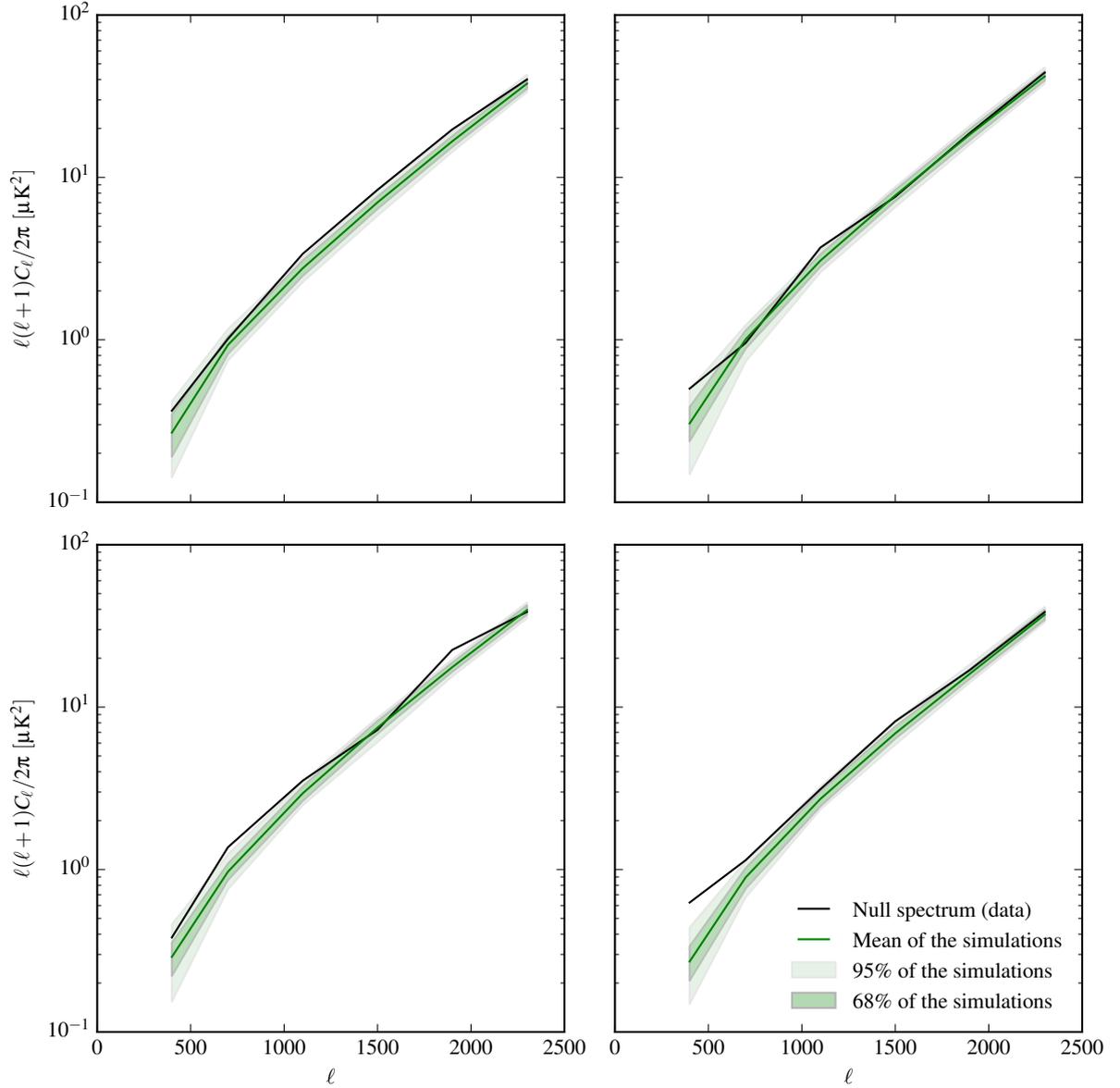
Figure 7.8: Comparison of the noise power in the four null maps derived with the "even v.s. odd" scheme and the corresponding noise simulations.

7.5.2   Power spectrum estimation with the help of cross-spectra

The typical solution to the noise bias problems is to use cross-spectra. Instead of computing one single map from the entire data set, we can divide the data set into subparts and compute a map for each of them. The cross-spectrum of different maps is free from noise bias. This is easy to show for two maps Y and X with same signal and uncorrelated noise n^X and n^Y: using the simplified cross-power spectrum estimator of Eq. (3.39)

$$\langle \hat{C}_\ell^{BB(XY)} \rangle = \sum_{m=-\ell}^{\ell} \frac{{}_B X_{\ell m}^* \, {}_B Y_{\ell m}}{2\ell + 1} \tag{7.5}$$

$$= \sum_{m=-\ell}^{\ell} \frac{\langle {}_B a_{\ell m}^* \, {}_B a_{\ell m} \rangle + \langle ({}_B n_{\ell m}^X)^* \, {}_B a_{\ell m} \rangle + \langle {}_B a_{\ell m}^* \, {}_B n_{\ell m}^Y \rangle + \langle ({}_B n_{\ell m}^X)^* \, {}_B n_{\ell m}^Y \rangle}{2\ell + 1}$$

$$= \sum_{m=-\ell}^{\ell} \frac{\langle {}_B a_{\ell m}^* \, {}_B a_{\ell m} \rangle}{2\ell + 1} = C_\ell^{BB}, \tag{7.6}$$

where a key role is played by the assumption of no correlation between noise in the two maps: $\langle (_B n_{\ell m}^X)^* \, _B n_{\ell m}^Y \rangle = 0$.

In the case of the pseudo-power spectrum estimators the formalism is more complicated. In particular, it has to allow for a different pseudo basis for each map because, in general, they have different sky coverage and observation depth. However, in the POLARBEAR first and second season observations the sky area is very small and we have a huge redundancy in our scanning strategy. Therefore, two reasonably large subsets of the data have similar sky coverage and observation depth. We therefore continue to use the window function of the full data set to define the pseudo basis of any map produced from a subset of the data. This has a considerable computational advantage because it allows computing only one mixing matrix instead of one for each pair of maps. Consequently, we continue using the pseudo-power spectrum formalism illustrated in Sec. 4.3.1, the only difference is that in the pseudo spectrum, Eq. (4.75), X and Y refer to different maps. For the general treatment see Tristram et al. (2005) for the standard formalism, Grain et al. (2009) for the pure formalism and Grain et al. (2012) for the hybrid formalism.

We want to divide the data set into disjoint subsets, compute a map for each of them and use the average of their cross-spectra as the estimator of the power spectrum $\hat{C}_\ell$. One further complication comes from the map-domain filtering procedure and the consequent map and power spectrum bias. In principle, each map has to be filtered using the eigenstructure estimated during the map-making. The formalism of Sec. 7.3 can be applied to each cross-spectrum if we estimate dedicated transfer functions for each pair of maps. Once all the unbiased cross-spectra are computed with the use of the transfer functions, they can be averaged. However, this procedure is quite costly because it requires a full set of simulations for each map in order to compute the transfer functions. Therefore we decide to apply the same filtering to every map, this allows to use the same transfer function for all the cross-spectra, similarly to how using a single window function allowed to use only one mixing matrix. As mentioned in the previous section, the common mode removal operator is based on the $\boldsymbol{A}^\top \boldsymbol{F_T A}$ matrix of the full data set and keep the usual 0.1 threshold on the eigenvalues. Using a common filtering for all the maps allows computing only one set of transfer functions in the auto-spectrum framework illustrated in Sec. 7.4.

The division of the data set in subsets is the same described in Sec. 7.5.1. The number of data splits that we use is driven by our computational resources, since for each data split we need a map-making run. As discussed in Sec. 6.3, the expensive parts of the algorithm are the computation of the $\boldsymbol{A}^\top \boldsymbol{F_T A}$ matrix and its inversion. The former scales linearly with the number of observations and, therefore, dividing the observations into subsets does not change the total cost. However, the cost of a single inversion depends only on the number of pixels, regardless of the volume of data involved in the computation of the matrix to be inverted. Therefore the global cost of this step of the map-making is proportional to the chosen number of subsets.

We decide to split the data set into 8 parts because it provides a good balance between the additional map-making computational cost and the extra uncertainty we get on our power spectrum estimation (compared to the auto-spectrum case). This latter can be estimated as follows. We are in the noise dominated regime, therefore $C_\ell \ll N_\ell$, where N_ℓ is the power-spectrum of the noise on the full-data set map. Assuming an even division of the data set in k split, the noise power spectrum in each split is kN_ℓ. The variance of the power spectrum estimator using the auto-spectrum of the full-data set map is $2N_\ell^2/(2\ell + 1)$, while if we use the cross-spectra of k splits the variance is $2N_\ell^2 k/[(2\ell + 1)(k - 1)]$ (see Polenta et al., 2004, Eq. (10) and (17)). Therefore, using k cross-spectra the variance increases by a factor $k/(k - 1)$. For $k = 8$, the increase in the error bars is $\sqrt{k/(k-1)} \sim 7\%$.

The uncertainty on the power spectrum estimation is evaluated using signal+noise simulations. Similarly to Sec. 7.4, a set of 100 simulations analogous to the data is created and processed in the same fashion. For each signal simulated map, eight maps are created by adding on the top of the signal map a noise simulation produced using the eigenstructure specific to each data split. These 8 maps are filtered and their cross-spectra are computed and averaged. These 100 biased cross-spectra are then debiased using the transfer functions. The error-bars are estimated as the standard deviation of this set of 100 unbiased spectra.

As a last remark, the average that compresses the cross-spectra into a single power spectrum estimate is actually weighted. We assign to a map i a weight w_i equal to the trace of its $\boldsymbol{A}^\top \boldsymbol{M A}$ matrix. The cross-spectrum $\hat{C}_b^{ij}$—produced by X^{2}PURE—between maps i and j receives a weight equal to $\sqrt{w_i w_j}$. Summarising, the power spectrum is estimated as follows

$$\hat{C}_b = F_b \frac{1}{\sum_{i \neq j} \sqrt{w_i w_j}} \sum_{i \neq j} \hat{C}_b^{ij} \sqrt{w_i w_j}. \tag{7.7}$$

7.6 Recap: the standard POLARBEAR analysis set-up

We now summarise the analysis set-up described and justified in the previous sections and chapters. Starting from raw data samples, the information is compressed into angular power spectra through the following steps.

1. Low level data processing (Sec. 5.5):

   - pointing reconstruction;
   - beam characterisation (estimation of B_ℓ);
   - timestream calibration;
   - polarization angle reconstruction.

2. Division of the data set into 8 parts (Sec. 7.5).

3. Map-making on each of the 8 parts (Chapter 6), which produces:

   - 8 reconstructed maps (to be used for power spectrum estimation);
   - 8 map covariances (to be used for efficiently producing noise simulations).

4. Map-making run on the full data set

   - reconstructed map (used just for displaying purposes);
   - map covariance (defines the mode removal);
   - map depth (the window function for the power spectrum estimation).

5. Map-domain mode removal: we retain only the modes that, in the full data set map, have a variance smaller then 10 times the one of the best observed mode (Sec. 7.4).

6. Trimming of the window function to the area retained by the mode removal (Sec. 7.4).

7. Biased power spectrum estimation: X^{2}PURE cross-spectra of the 8 filtered maps (Sec. 7.5.2).

8. Estimation of the transfer functions (Sec. 7.4)

   - E-only and B-only map simulations;
   - application of the mode removal;
   - auto-spectrum estimation using X^{2}PURE;
   - average of the spectra issued by the different types of simulations.

9. Transfer function inversion and unbiased power spectrum estimation (Sec. 7.4).

10. Error-bar estimation from signal+noise simulations processed with the same analysis procedure of the data (Sec. 7.5.2).

7.7 Foreground treatment in the POLARBEAR analysis

The current POLARBEAR data sets consist of measurements in a single frequency band centred at $\sim$150 GHz. The CMB power spectra are then directly derived from the single frequency map produced via the map-making procedure, implicitly assuming that no appreciable foreground signal is present. In this section I argue that this is indeed the case for the analysis of the current POLARBEAR patches.

The angular power spectrum of both dust and synchrotron can be approximated with a power law $\propto \ell^\alpha$ with a negative α (see, for example, Fig. 5.9c and Fig. 3.13). Consequently, foregrounds are not a major issue for POLARBEAR first and second seasons because we probe intermediate/small angular scales ($\ell > 500$) of very small patches, which were moreover selected for their low intensity of foreground emission. Currently we handle the foregrounds by using publicly available multi-frequency data sets (*Planck* and WMAP) to derive upper bounds on the contaminating power in our patches. The preliminary estimates find the amplitude of the foregrounds in our patches consistent with zero, where the limiting factor in these constraints come from the sensitivity of *Planck* and WMAP on the very small sky regions probed by POLARBEAR. This analysis is coordinated by our POLARBEAR collaborators in SISSA. In the analysis of

the first season (Polarbear Collaboration, 2014a), conservative upper bounds on dust and synchrotron were respectively derived from the *Planck sky model* and the measurements of the QUIET experiment.

As far as radio galaxies are concerned, resolved sourced were masked during the power spectrum estimation and an upper bound is set on the contribution of unresolved sources using simulations— based on number counts from de Zotti et al. (2005). The total upper bound on the foregrounds power spectrum $\ell(\ell+1)C_\ell/2\pi$ is 50, 48, 57 and 64 $\times 10^{-4} \mu K^2$ respectively at $\ell = 700$, 1100, 1500 and 1900. All these values are well below the expected lensing B mode signal (see e.g. Fig. 3.13) and the statistical uncertainty of our BB measurement (see Fig. 5.7).

Chapter 8

VALIDATION OF THE DATA SET

In the previous chapter I have described the POLARBEAR analysis set-up and the studies that led to it. As mentioned in Sec. 7.1, we cannot look at the products of the analysis before we validate the data set and the pipeline. This is implemented in practice with the help of the so-called null tests and involves testing signal-free combinations of the data. These are obtained by first dividing the data sets into two subsets with similar statistical power and producing from them maps of the same sky area. The signal-free combinations of their spectra are then tested for consistency with the data noise model.

The predefined data splitting criteria are the following.

- First v.s. second half of the data set

- First v.s. second season of observation

- Rising v.s. setting

- Far v.s. close to the moon

- Far v.s. close to the sun

- Sun above v.s. below the horizon

- High elevation v.s. low elevation

- Good v.s. bad weather

- High v.s. low gain

- Left v.s. right side of the focal plane

- Focal plane pixel types (the antennas of one group are rotated by 45° with respect to those of the other group)

- Left- v.s. right-going subscans

The splits are chosen to be sensitive to different systematics that were considered particularly pertinent to the POLARBEAR measurements, see POLARBEAR Collaboration (2014a) for more details.

For the pipeline described in this thesis, the full null test suite is not computationally feasible due to the hundreds of unbiased map-making runs that it requires. Especially, the pipeline is not suitable because the data set definition is an iterative procedure and the full suite has to be run many times. This procedure is indeed performed by the other POLARBEAR pipeline: thanks to the use of biased maps, it avoids the computationally intensive parts of map-making and, roughly speaking, the cost of the full null test suite is of the order of a single unbiased map-making run.

Once the data set is defined we validate it using our pipeline. In this chapter I describe this validation.

In the subsequent sections we describe other three types of tests, two of them are still power-spectrum-based while the other is map-based. These tests take advantage of our unbiased map-making and are therefore specific to our pipeline. In this sense this data validation is complementary to the data selection and validation process performed with the other pipeline.
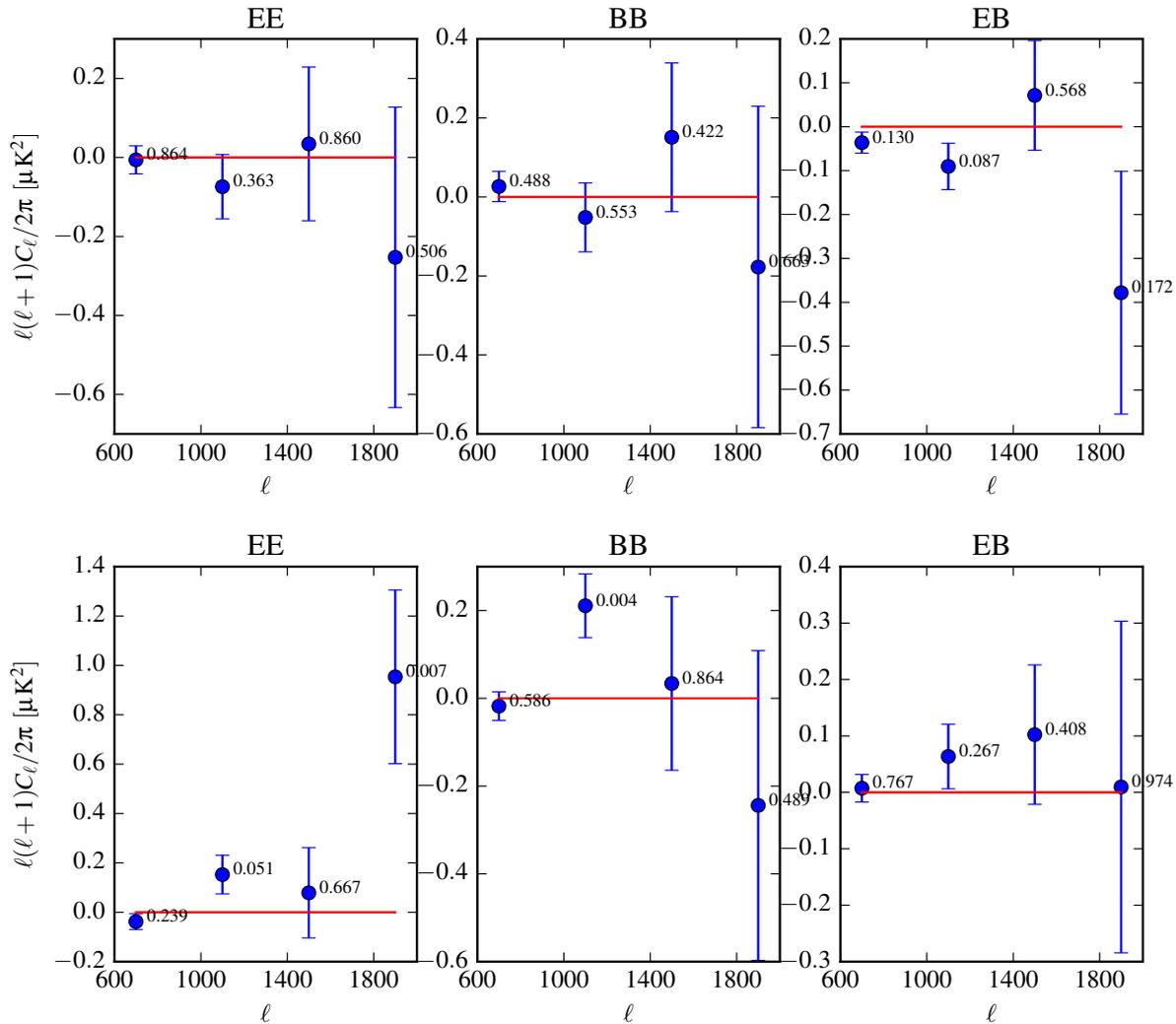
Figure 8.1: Null spectra of the "focal plane pixel type" (top) and the "high v.s. low gain" (bottom). The annotations next to each point correspond to the p-value of the point, assuming normal distribution and sigma given by the error bar.

8.1   Null power spectra analysis

We compute and discuss here null spectra from the cross-spectrum of null maps. This is made possible thanks to a capability of our pipeline to produce unbiased maps of the sky. Indeed, the null power analysis based on the biased sky maps can be only attempted on the power spectrum level, with help of a signal free combinations of cross-spectra of pairs of the maps. This approach has been used in the first POLARBEAR paper POLARBEAR Collaboration (2014a). Our map-based approach is more direct, requires fewer assumptions and its results are easier to interpret. We describe it here.

Given a null test type, we divide the data set of each of the two splits into 8 parts. We get a total of 16 splits and we run the map-making for each of them. We then compute 8 null maps by subtracting the first map of the two groups, and so on for the second, third, etc. We then apply the map domain mode removal based on the eigenstructure of the full data set. Finally, we compute the null spectrum as the weighted average of the cross-spectra of the 8 null maps. Then we use the eigenstructure of each of the 16 maps to produce 100 noise simulations that are processed in the same fashion as the real data. We then compute the standard deviation of the 100 simulated noise spectra and use it as error bar for the null spectrum computed from the data. In summary, we are running the same procedure of the real analysis on 8 null maps instead of 8 unbiased maps.

We perform two null tests of this kind: the "high v.s. low gain" and the "focal plane pixel type" null tests. The resulting spectra are reported in Fig. 8.1.
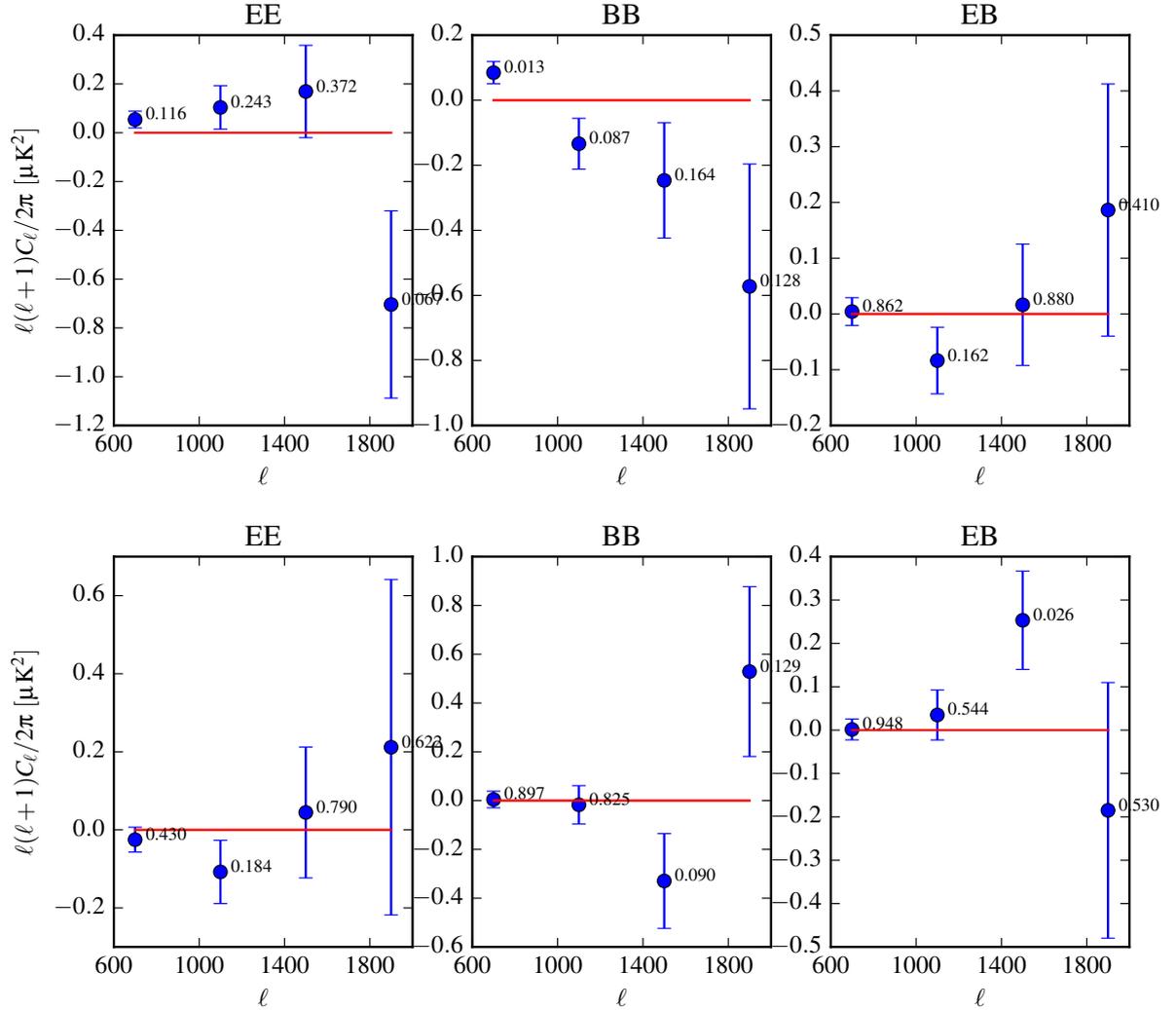
Figure 8.2: Null spectra extracted from the standard analysis set-up. The eight maps are combined into four null maps, which are exploited for additional null tests. The null maps calculation is done either according to the "fist v.s. second half" (top) or the "even v.s. odd" (bottom) schemes. The annotations next to each point correspond to the *p*-value of the point, assuming normal distribution and sigma given by the error bar.
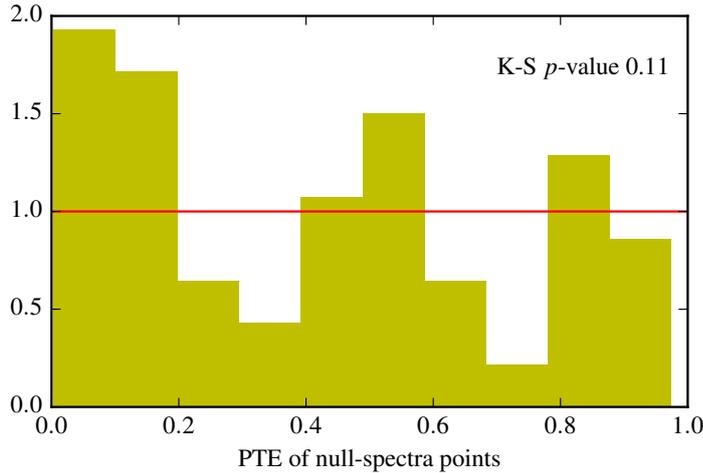
Figure 8.3: Normalised histogram of the p-values of all the points of the null spectra (4 Tests $\times$ 3 Spectra $\times$ 4 Bins = 48 Entries)

In order to exploit further the products of our analyses, we also use the run of the real analysis to perform two further types of null tests. From the 8 unbiased maps we obtain 4 null maps using the "first v.s. second half" and the "even v.s. odd scheme", described in Sec. 7.5.1. The rest of the null test procedure is completely analogous to the tests of the previous paragraph: the only difference is that we have 8 unbiased maps to start with, instead of 16. The null spectra are reported in Fig. 8.2.

We have to define how to quantify the compatibility of the null spectrum with the noise only simulations. As a first approximation we assume as null hypothesis that the spectra points have a Gaussian distribution centred at zero and with sigma given by the standard deviation of the noise simulations. The p-value (also referred to as probability to exceed, PTE) of a given point of the null spectrum $\hat{C}_b^{\mathrm{null}}$ is given by

$$p\text{-value} = 1 - \int_{-\hat{C}_b^{\mathrm{null}}}^{\hat{C}_b^{\mathrm{null}}} \frac{\mathrm{d}x}{\sqrt{2\pi}\sigma_b} \exp\left(\frac{x^2}{2\sigma_b^2}\right) \tag{8.1}$$

We report the p-value of each of the null spectrum points in the Figs. 8.1 and 8.2.

The null spectra are generally compatible with zero. Nevertheless, some individual null spectra points have quite low p-value. In order to assess if they represent a null test failure or a statistical fluke, we study the distribution of all the p-values and assess its compatibility with the flat distribution (Fig. 8.3). The Kolmogorov-Smirnov test gives a p-value of 11%, meaning that globally our data are statistically compatible with the null hypothesis.

8.2   Detailed study of the outliers

Each null test is in principle sensitive to different systematics and a given outlier might indicate the presence of some systematics. The results of the global analysis proposed in the previous section are reassuring but require some complementary investigation. If we suppose that an outlier is indeed some systematics and we combine it with a large number of other points that are insensitive to the systematic effect, the outlier might become statistically insignificant and incapable to affect the global p-value, which is driven by the much larger set of points. For this reason we probe the outliers also individually. In particular, we study the behaviour of the cross-spectra that, once averaged, give the offending null spectrum point.

Since the cross-spectra are computed from null maps, they are also expected to be individually compatible with zero and we can assess this with the set of 100 noise simulations of the same cross-spectrum. Therefore, instead of computing cross-spectra, combining them and checking the compatibility with zero signal, we compute the cross-spectra and directly study their distribution and individual compatibility with the zero-signal hypothesis.

As an example of this procedure, we show the investigation on the second bin $\ell = 1100$ of the BB spectrum of the "high v.s. low gain" null test, which has a particularly low p-value: 0.4%.

(a)

(b)
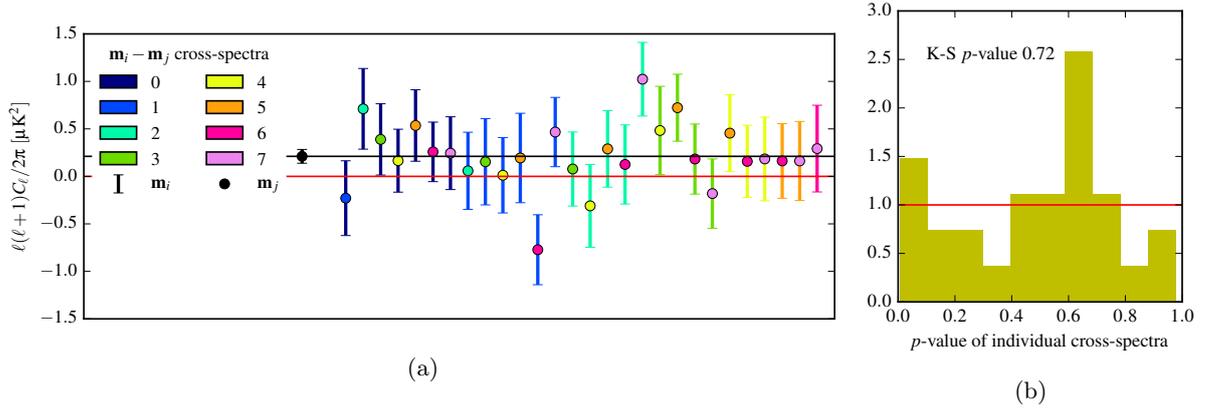
Figure 8.4: Detailed study of the second bin of the BB "high v.s. low gain" null spectrum, the leftmost point in the left panel. (a) All the cross-spectra C_ℓ^{ij} that are coadded into the null spectrum point being investigated (only $i < j$ are displayed). The error bars are the standard deviation of the simulated C_ℓ^{ij} cross-spectra. No particular index is responsible for the behaviour of the combined point. (a) The right panel is a histogram of the p-values of the individual cross-spectra, it is compatible with the flat distribution.

In Fig. 8.4a we report all the cross-spectra with their individual error bars. First of all, the outlying combined spectrum does not seem to be dragged by the cross-spectra involving a specific split, which would suggest a problem in a portion and trigger investigation at the TOD level. Secondly, the distribution of the p-values of the individual cross-spectra is compatible with the flat distribution, with an excellent p-value of 72% from the Kolmogorov-Smirnov test. This test ensures that the error bars are not underestimated, otherwise the single cross-spectra would scatter more than what is expected from their error bars and produce an excess of low p-values in the histogram.

Concluding, we do not find particular indications of a systematics or error bar underestimation: it is compatible with a statistical fluke.

8.3   Robustness of null power with respect to the threshold

The threshold defining the mode removal is a free parameter of our analysis. Its value is defined before any analysis on the real data and, especially, before seeing the BB power spectrum produced for different thresholds.

We can use the threshold to define an additional null test. The test is based on another null test, the "first v.s. second" in this example. We filter the unbiased null maps that produced the null spectrum with a threshold equal to 0.05, 0.2 and 0.5, we compute their null spectrum and finally we take the difference with the null spectrum obtained with the 0.1 threshold. We do the same analysis on 100 noise simulations and check that the shifts obtained from the data and simulations are compatible. The results are reported in Fig. 8.5 and show a general good agreement: when we vary the threshold the change in the power spectrum is consistent with the expectations from our noise model. These results also support that, as expected, the threshold of our choice—based on completely different and a priori arguments—has no special advantage in passing the null power tests.

8.4   Map-domain null test

The previous studies checked the compatibility of our data with the noise model at the power spectrum level. Since we are capable of both producing null maps and estimating their covariance matrix according to our data model, the null test can be performed directly at the map level, potentially providing more comprehensive diagnostic than the null power test. The map-level test is again made only possible within the unbiased pipeline.

This analysis is performed using the 8 map-making runs of the standard POLARBEAR analysis. We

build a global null map as follows

$$\boldsymbol{m}_{\mathrm{null}} = \boldsymbol{m}_0 - \boldsymbol{m}_1 + \boldsymbol{m}_2 - \boldsymbol{m}_3 + \boldsymbol{m}_4 - \boldsymbol{m}_5 + \boldsymbol{m}_6 - \boldsymbol{m}_7 \tag{8.2}$$

We report both the complete map and its filtered version in Fig. 8.6. While the temperature map shows in both cases clear signs of atmospheric contamination, the polarization map does not show any apparent residual and is compatible with noise from a visual inspection.

In order to make this statement more quantitative, we perform the following analysis. For each map $\boldsymbol{m}_i$ we have the estimated covariance matrix, $(\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A})_i^{-1}$, from which we can compute the covariance matrix of the null map and its eigendecomposition

$$\langle \boldsymbol{m}_{\mathrm{null}} \boldsymbol{m}_{\mathrm{null}}^\top \rangle = \sum_{i=0}^{7} (\boldsymbol{A}^\top \boldsymbol{F_T} \boldsymbol{A})_i^{-1} = \boldsymbol{V} \mathrm{diag}(\boldsymbol{e})^{-1} \boldsymbol{V}^\top. \tag{8.3}$$

The choice of $\mathrm{diag}(\boldsymbol{e})^{-1}$ instead of $\mathrm{diag}(\boldsymbol{e})$ is done for consistency with Eq. (6.19): the large entries of $\boldsymbol{e}$ correspond to the well-observed modes. In order to test if this covariance matrix, based solely on the data model, actually describes the statistical properties of the null maps, we compute the vector of the prewhitened mode amplitudes

$$\mathrm{diag}(\boldsymbol{e})^{1/2} \boldsymbol{V}^\top \boldsymbol{m}_{\mathrm{null}}. \tag{8.4}$$

Its entries are expected to have Gaussian distribution with zero mean and unit sigma. From their histogram in Fig. 8.7a we see that they follow very closely the predicted normal distribution. The K-S test gives indeed a p-value of $\sim 3\%$, which is extremely satisfactory given the sensitivity of the K-S test with such a large number of data points. In Fig. 8.7b we show how this value evolves if we discard all the amplitudes beyond a given threshold on the eigenvalues. Above the ~ 0.2 threshold, the p-value are always above 10% and, in particular, for the 0.1 threshold it is equal to 81%, implying that the filtered null map is fully compatible with the expected map-domain noise.

8.5   Maps from the validated data set

As of late summer 2016, the validation of the analysis of the first and second season is being finalised and we will soon unblind our power spectrum. The preliminary maps produced from the current validated data set are shown in Fig. 8.8 (top panel). These are the first unbiased maps produced by the POLARBEAR experiment. They have only three missing modes, which we identified and described in Sec. 6.6.1. In the same section we also described the large-scale noisy modes, clearly visible in the reconstructed map. In the middle panel we show that most of them are removed if we apply a mild threshold of 10^{-5} on the eigenvalues. Any trace of long modes disappears if we apply our fiducial 0.1 threshold (bottom panel). This mode filtering also removes the noisy pixels at the boundaries, leaving only the inner core, which clearly shows a "grid"-like pattern, horizontal from Q and diagonal from U: the signature of the E modes.

In Fig. 8.9 we focus on the 3°-by-3° centre of the patch and show the good visual agreement between the temperature map of POLARBEAR and the *Planck* full mission map recovered with the SMICA method (Planck Collaboration, 2015c). Even though the POLARBEAR map has a somewhat better resolution—3.5 arcmin v.s. an effective 5 arcmin of the *Planck* SMICA map—and the Planck maps have some small-scale artifacts, probably attributable to the SMICA algorithm, the agreement between the two maps is truly excellent, emphasizing at the same time the quality of our map-making algorithm and the POLARBEAR data. We also note that a similar comparison of the polarized signals is hampered by the high-level of noise on the pixel-scales in Planck maps and it is therefore not shown here.
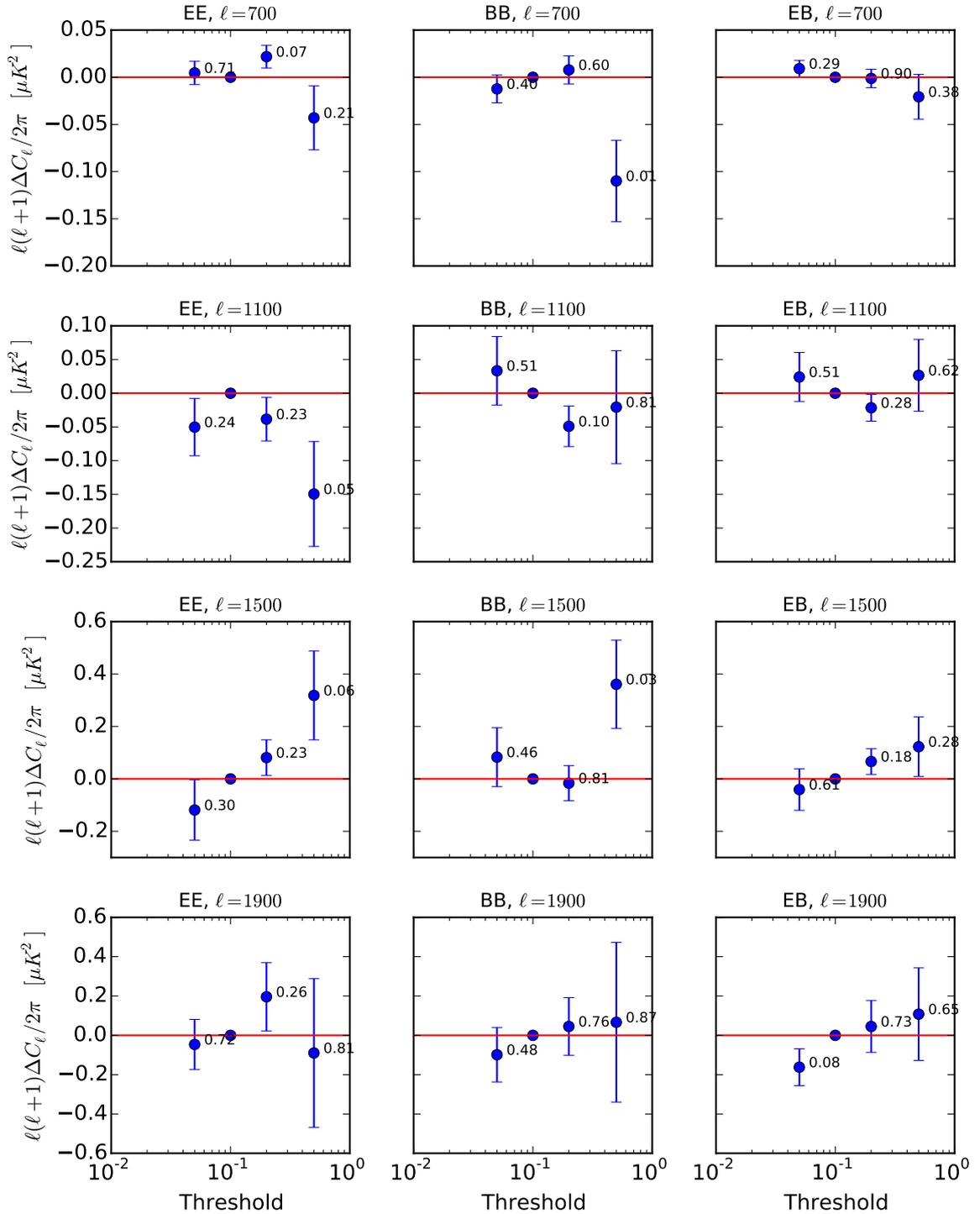
Figure 8.5: Threshold null test for the "first v.s. second" null test. Each point of the power spectrum is computed for different thresholds. The plots show the difference with respect to the reference threshold 0.1. The same procedure is performed on simulations, their standard deviation defines the error bars. The p-value of each point is reported next to the marker.
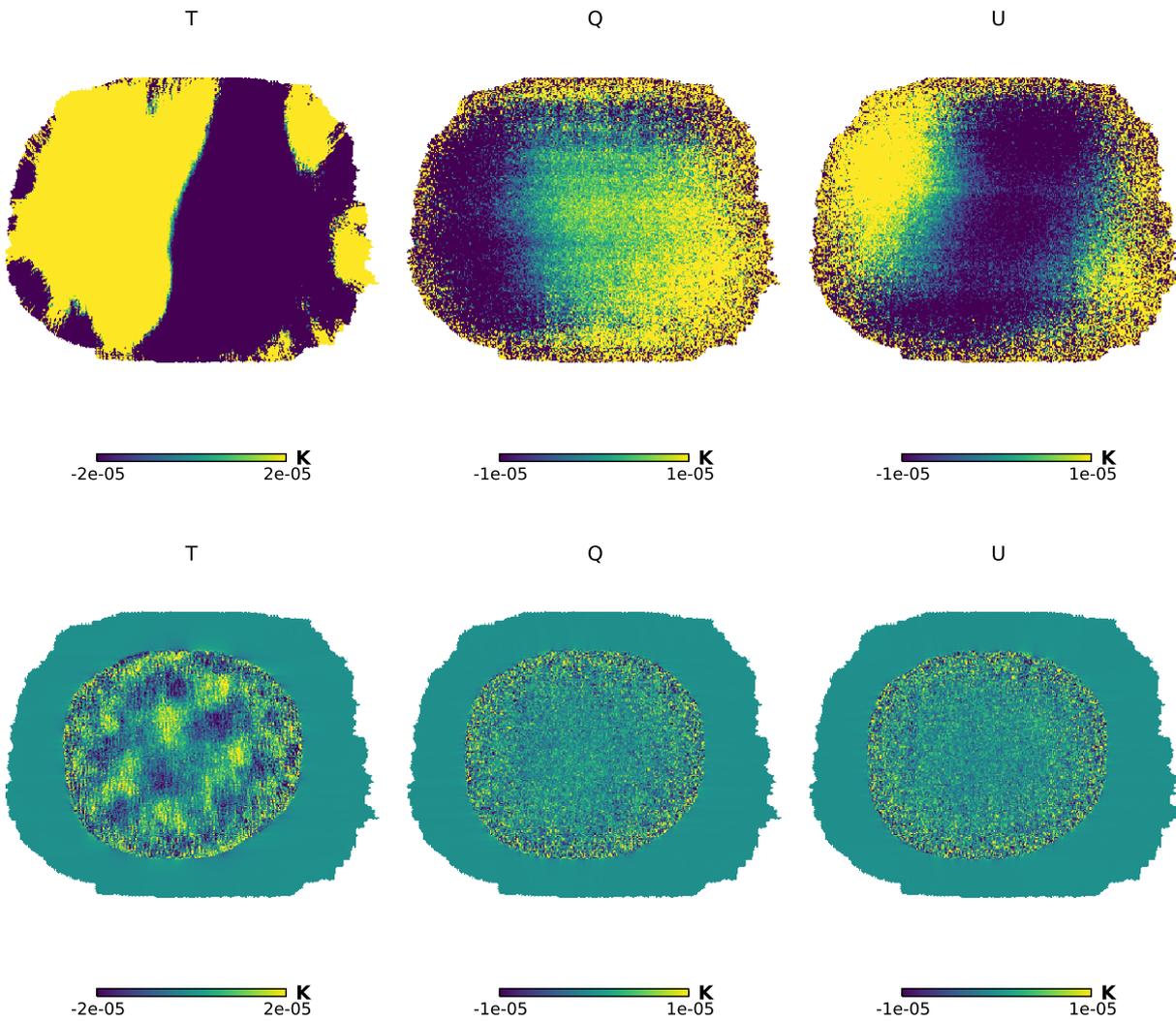
Figure 8.6: Null map of the RA23 patch. The top panel is the full unbiased map while in the bottom panel the 0.1 threshold is applied
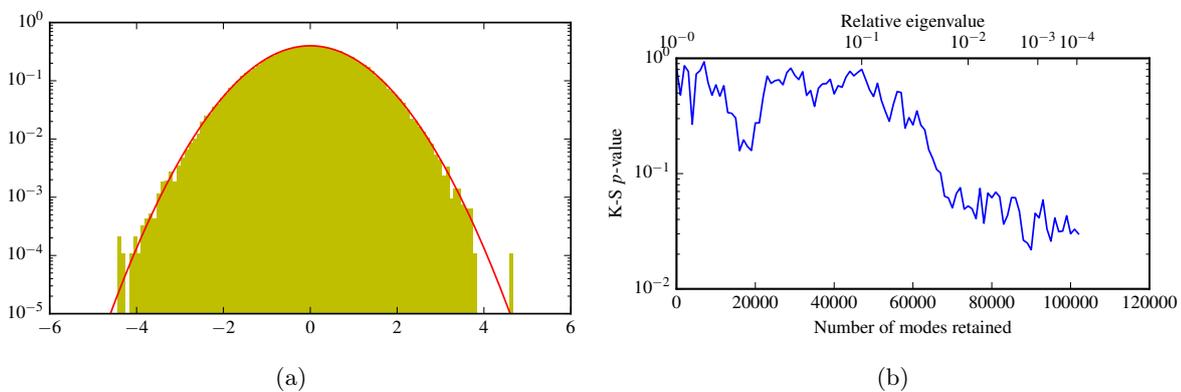
Figure 8.7: Study of the prewhitened amplitudes of the eigenmodes in the null maps. (a) Histogram of the amplitudes, the over-plotted red line is the Gaussian with $\mu = 0$ and $\sigma = 1$. (b) p-value of the Kolmogorov-Smirnov test when only the best observed fraction of the modes is considered.
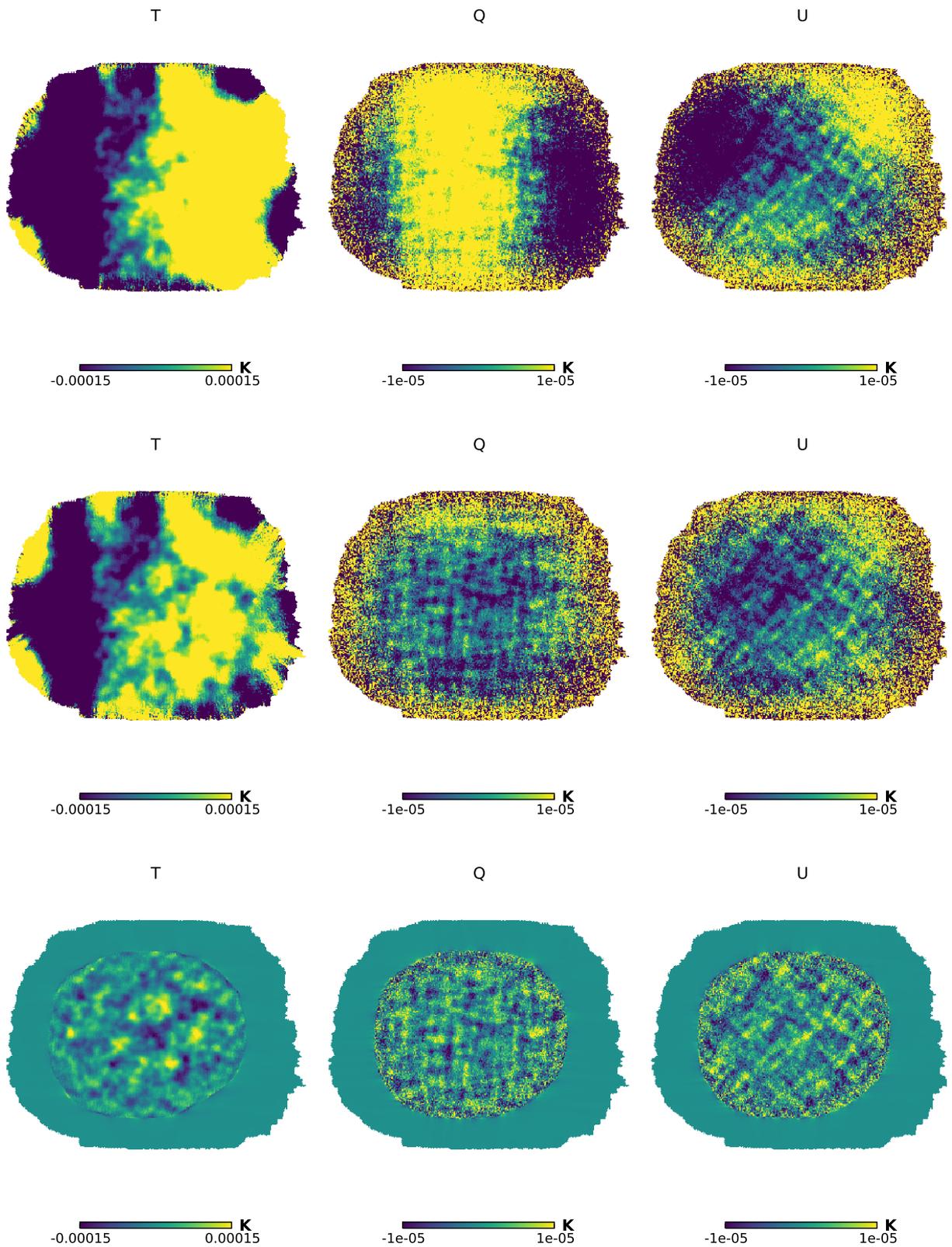
Figure 8.8: Map of the RA23 patch. The top panel shows the actual map produced by the unbiased mapmaker. In the middle panel a mild threshold on the eigenvalues is applied (10^{-5}). The bottom panel shows the actual map utilised for the power spectrum estimation, produced with a threshold of 0.1 on the eigenvalues. (Notice that in the top and middle panel some sky areas saturate the colour scale)
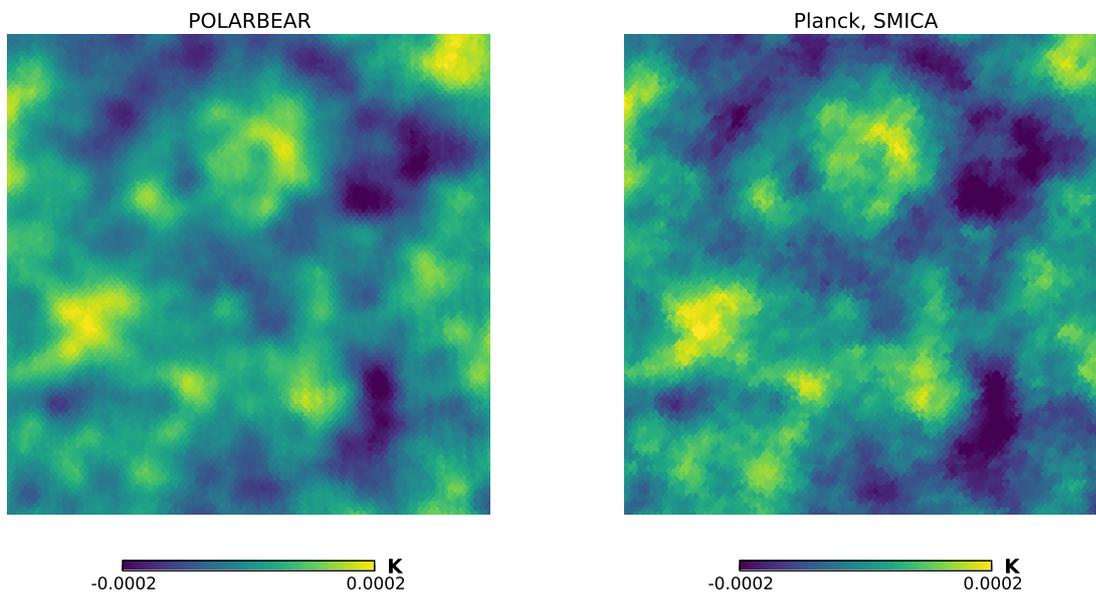
Figure 8.9: The 3°-by-3° centre of the POLARBEAR RA23 temperature map (threshold 5×10^{-4}) compared to the same sky area reconstructed from the *Planck* full mission using the SMICA component separation method. The two maps have respectively a 3.5 and 5 arcmin resolution.

Chapter 9

CONCLUSIONS

In this thesis I address the problem of producing high-fidelity CMB maps from the data of current and forthcoming, ground-based CMB polarization experiments. I describe relevant methodological developments, their implementation and application to data sets of the POLARBEAR experiment. I also discuss what impact these map-making techniques may have for further stages of data analysis, focusing specifically on the power spectrum estimation. This again is done in the general context as well as in the specific one of the POLARBEAR data analysis. Last but not least, I present the POLARBEAR experiment, its first and second season data set and the current status of its analysis.

In spite of years of development, the map-making continues providing significant challenges to data analysis for modern CMB experiments. The time-ordered data they produce are typically composed of 10^{11}-10^{13} samples, which have to be compressed into maps typically composed of 10^5-10^7 pixels, depending on the experiment. This has to be done in a way that the cosmologically-relevant information remains as intact as only possible. Achieving this is exacerbated by the presence of low-frequency noise and contaminating signals, for example due to atmospheric emission or ground pickup. These components are typically removed with some filtering operation that, if not properly devised and accounted for, produces a bias or some contaminant leftover in the reconstructed maps.

In this thesis, following on Poletti et al. (2016), I present a new perspective on the unbiased map-making formalism in the presence of filtering. In particular, the formalism proposed is capable of naturally incorporate arbitrary sets of filters and handle them in a statistically sound manner. It also encompasses previous techniques in a more general framework.

A key problem in the high-fidelity reconstruction of sky maps is the interplay between filters and sky signal. In this work I discuss when it leads to degeneracies in the formalism and how to handle them in practice. Moreover, as a concrete example, I study the problem of ground-pickup filtering, ubiquitous in the ground-based observations, I investigate the degenerate sky modes that it produces and what are the possible degeneracy-breaking effects.

I have implemented the proposed map-making formalism as a massively-parallel C++/MPI code. The implementation not only delivers an outstanding accuracy in the map reconstruction, but also provides the full covariance matrix of the map-estimator, which enables to identify both the degeneracies and the map-domain correlations induced by the time-domain filtering.

Since the beginning of my PhD training, I have been a member of the POLARBEAR Collaboration, a CMB B-mode dedicated experiment located in the Chilean Atacama desert. The main motivation for the map-making methodological and technical developments was indeed their application to the POLARBEAR first and second season observations. The developed method and code have met successfully the challenges the data set has posed.

The map-maker was instrumental in producing (nearly) unbiased maps of all the three Stokes parameters, which were obtained together with a precise characterization of their noise correlation structure. These maps have been used to estimate the polarized power spectra using pure pseudo-spectrum methods. This step has been shown to be compromised by the non-trivial correlation structure of the map noise and stimulated a development of a practical approach to compensate for such deficiencies. I have shown that when properly applied this approach allows matching the performance of the quicker and

simpler biased map-making, based on a simple, noise-weighted binning of the filtered time-domain data. This emphasizes that a full exploitation of the potential of unbiased maps requires the help of more involved and resource consuming techniques—such as those based on maximum likelihood principles.

These results, though obtained in the context of the Polarbear data sets, are expected to be more general and should apply to all the ground-based experiments.

As a member of the Polarbear experiment I have led the development, validation and exploitation of an entire data analysis pipeline that, starting from the calibrated time-ordered data, produces the final power spectra by exploiting unbiased maps. This pipeline incorporates the map-making and power-spectrum estimation techniques and codes described earlier.

Because of the blindness policy that the Polarbear Collaboration adopts, we are not allowed to see our BB spectrum—our main science product—before having defined and validated both the tools and final data set. Instead, the pipeline I have developed has been validated in a series of null tests. Some of them reproduce those performed in the first season results (Polarbear Collaboration, 2014a), others are novel and exploit the new unbiased map-making approach.

As of today, the data set validation is being finalised. The power spectrum will be soon unblinded and we expect that it will bring improvements on the current constraints on the lensing peak of the BB power spectrum. From one hand this will confirm our understanding of structure formation in the Universe and on the other will bring us a step closer to exploiting the cosmological constraints that can be deduced from the measurements of the lensing potential.

Appendix A

BLOCK-WISE GLS ESTIMATOR

In this section we provide more details about the equations in Sec. 4.1.4. We derive the unbiased map estimator in the presence of time domain filtering or, more in general, we write the GLS estimator in a block-wise fashion.

We report here for completeness the data model

$$d = As + Ty + w$$
$$= [A, T]\begin{bmatrix} s \\ y \end{bmatrix} + w, \tag{A.1}$$

of which we want to estimate the parameters s and y. For simplicity we assume a unitary weight matrix $M = 1$. With this assumption, we remind that the filtering operator reads

$$F_X = 1 - X(X^\top X)^{-1}X^\top. \tag{A.2}$$

As in Sec. 4.1.4, we are assuming that the there is no degeneracy: $[A, T]$ is full rank.

A.1   The direct approach

The GLS estimator can be written as

$$\begin{bmatrix} \hat{s} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} A^\top A & A^\top T \\ T^\top A & T^\top T \end{bmatrix}^{-1} \begin{bmatrix} A^\top \\ T^\top \end{bmatrix} d. \tag{A.3}$$

The following matrix

$$\begin{bmatrix} (A^\top F_T A)^{-1} & -(A^\top F_T A)^{-1}A^\top T(T^\top T)^{-1} \\ -(T^\top F_A T)^{-1}T^\top A(A^\top A)^{-1} & (T^\top F_A T)^{-1} \end{bmatrix} \tag{A.4}$$

is the block-wise inverse of the system matrix. To show this we only need to apply this matrix on the system matrix and check that we get the identity (it is trivial, one just need to perform the block-wise multiplication and remind Eq. (A.2)). We can use this matrix in the expression of the estimator and obtain the direct estimates of $\hat{s}$, Eq. (4.29), and $\hat{y}$, Eq. (4.30),

$$\begin{bmatrix} \hat{s} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} (A^\top F_T A)^{-1} & -(A^\top F_T A)^{-1}A^\top T(T^\top T)^{-1} \\ -(T^\top F_A T)^{-1}T^\top A(A^\top A)^{-1} & (T^\top F_A T)^{-1} \end{bmatrix} \begin{bmatrix} A^\top \\ T^\top \end{bmatrix} d \tag{A.5}$$

$$= \begin{bmatrix} (A^\top F_T A)^{-1}A^\top F_T \\ (T^\top F_A T)^{-1}T^\top F_A \end{bmatrix} d. \tag{A.6}$$

A.2   The two-step approach

Also the following matrix is inverse of the system matrix,

$$\begin{bmatrix} (A^\top A)^{-1} + (A^\top A)^{-1}A^\top T(T^\top F_A T)^{-1}T^\top A(A^\top A)^{-1} & -(A^\top A)^{-1}A^\top T(T^\top F_A T)^{-1} \\ -(T^\top F_A T)^{-1}T^\top A(A^\top A)^{-1} & (T^\top F_A T)^{-1} \end{bmatrix}. \tag{A.7}$$

Again, simply apply this matrix on the system matrix to find the identity.

Notice that just the first line has changed. Of course, the expressions are different but the two matrices are equal. The blocks of this matrix can be derived from the former block-wise inverse, Eq. (A.4). The Sherman–Morrison–Woodbury formula says that

$$(\boldsymbol{S} - \boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{Q})^{-1} = \boldsymbol{S}^{-1} - \boldsymbol{S}^{-1}\boldsymbol{R}(\boldsymbol{P} - \boldsymbol{Q}\boldsymbol{S}^{-1}\boldsymbol{R})^{-1}\boldsymbol{Q}\boldsymbol{S}^{-1} \tag{A.8}$$

where $\boldsymbol{S}$, $\boldsymbol{R}$, $\boldsymbol{P}$ and $\boldsymbol{Q}$ are matrices of the correct sizes. Choosing $\boldsymbol{S} \equiv \boldsymbol{A}^{\top}\boldsymbol{A}$, $\boldsymbol{R} \equiv \boldsymbol{A}^{\top}\boldsymbol{T}$, $\boldsymbol{P} \equiv \boldsymbol{T}^{\top}\boldsymbol{T}$ and $\boldsymbol{Q} \equiv \boldsymbol{T}^{\top}\boldsymbol{A}$ we get exactly the conversion from the first diagonal block of Eq. (A.4) to the one of Eq. (A.7). To demonstrate the equivalence of the two off-diagonal blocks we can equivalently multiply both by $(\boldsymbol{A}^{\top}\boldsymbol{F_T}\boldsymbol{A})$ from left and show that $(\boldsymbol{A}^{\top}\boldsymbol{F_T}\boldsymbol{A})(\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1}$ is equal to $\boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{T})^{-1}$, which is indeed demonstrated with the following simple steps,

$$
\begin{aligned}
[\boldsymbol{A}^{\top}\boldsymbol{F_T}\boldsymbol{A}](\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1} &= \left[\boldsymbol{A}^{\top}\boldsymbol{A} - \boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{T})^{-1}\boldsymbol{T}^{\top}\boldsymbol{A}\right](\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1} \\
&= \boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{T})^{-1}\left[\boldsymbol{T}^{\top}\boldsymbol{T} - \boldsymbol{T}^{\top}\boldsymbol{A}(\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}\right](\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1} \\
&= \boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{T})^{-1}.
\end{aligned}
\tag{A.9}
$$

Adopting this new block-wise inverse in the GLS estimator we get the two-step map estimator

$$
\begin{bmatrix} \hat{\boldsymbol{s}} \\ \hat{\boldsymbol{y}} \end{bmatrix} = \begin{bmatrix} (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1} + (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1}\boldsymbol{T}^{\top}\boldsymbol{A}(\boldsymbol{A}^{\top}\boldsymbol{A})^{-1} & -(\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1} \\ -(\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1}\boldsymbol{T}^{\top}\boldsymbol{A}(\boldsymbol{A}^{\top}\boldsymbol{A})^{-1} & (\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1} \end{bmatrix} \times
\tag{A.10}
$$

$$
\begin{aligned}
&\times \begin{bmatrix} \boldsymbol{A}^{\top} \\ \boldsymbol{T}^{\top} \end{bmatrix} \boldsymbol{d} \\
&= \begin{bmatrix} (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top} - (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}(\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1}\boldsymbol{T}^{\top}\boldsymbol{F_A} \\ (\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1}\boldsymbol{T}^{\top}\boldsymbol{F_A} \end{bmatrix} \boldsymbol{d} \\
&= \begin{bmatrix} (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{d} - (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}\boldsymbol{T}\hat{\boldsymbol{y}} \\ (\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{T})^{-1}\boldsymbol{T}^{\top}\boldsymbol{F_A}\boldsymbol{d} \end{bmatrix}.
\end{aligned}
\tag{A.11}
$$

Appendix B

FILTERING THE GROUND-SYNCHRONOUS SIGNAL

In this appendix we study in detail the filtering of ground-synchronous signals, elaborating on and justifying the conclusions presented in Sec. 4.2.4.

We start by considering the data recorded from a single detector, d, during a constant elevation scan, s. According to the data model in Eq. (4.50), the data recoded in a given azimuthal bin, ψ, is

$$
\begin{aligned}
\boldsymbol{d}|_{d,s,\psi} &= \boldsymbol{A}|_{d,s,\psi} \, \boldsymbol{s}|_{d,s,\psi} + \boldsymbol{G}|_{d,s,\psi} \, \boldsymbol{g}|_{d,s,\psi} + \boldsymbol{n}|_{d,s,\psi} \\
&= \boldsymbol{A}|_{d,s,\psi} \, \boldsymbol{s}|_{d,s,\psi} + \boldsymbol{1}|_{d,s,\psi} \, g_\psi + \boldsymbol{n}|_{d,s,\psi} \,,
\end{aligned}
\tag{B.1}
$$

where $\boldsymbol{1}|_{d,s,\psi}$ is a vector of ones of the appropriate length and g_ψ denotes a ground template amplitude common to all selected samples.

Let us now focus on the shape of the sky patch corresponding to $\boldsymbol{s}|_{d,s,\psi}$. The geometry of the problem is depicted in Fig. 4.3 and, for the time being, we neglect the role of the sky pixels. We consider a small azimuthal change of the pointing direction at a point on the sky at which the parallactic angle is η. The change in horizontal coordinates $\Delta(\mathrm{Az}, \mathrm{El}) = (\delta, 0)$ corresponds to an interval in the equatorial ones given by,

$$
\Delta(\mathrm{RA}, \mathrm{Dec}) = (-\delta \cos\eta, -\delta \sin\eta)
\tag{B.2}
$$

In particular, at the South Pole the horizontal coordinates correspond to the equatorial ones after flipping the y (El) axis, thus we have always $\eta = \pi$ and $\Delta(\mathrm{RA}, \mathrm{Dec}) = (\delta, 0)$. For a given elevation, the parallactic angle is always the same for a given azimuth but depends on the azimuth's value, so although it changes across a single ground template bin the changes are very small. Consequently, each constant elevation scan crossing a bin will draw a line interval on the sky given approximately by Eq. (B.2). Because of the Earth's rotation, if we keep on crossing the bin multiple times the intervals will cover a trapezoidal shape in sky coordinates, as shown in Fig. 4.3. The size of the trapezoid depends on the bin width but also on the parallactic angle, Eq. (B.2). However, the lines traced by the azimuthal bin end always follow the constant declination direction on the sky. Note that if $|\sin\eta| = 0$, which is always the case at the South Pole or whenever the instrument is pointed straight to the South or the North, the width of the trapezoid in declination is zero.

These patches, narrow in declination and elongated in azimuth, are degenerate (or ill-conditioned) modes. In the following we discuss this statement in detail for both temperature and polarization and investigate possible degeneracy-breaking effects.

B.1 Total intensity measurements.

Let us start with the total intensity measurements and consider a data subset that has the same ground contribution. It is described by (B.1), with the pointing matrix, $\boldsymbol{A}$, merely composed of ones and zeros. For this single scan the sky modes, $\tilde{\boldsymbol{s}}$, that are degenerate with the ground template signal have to fulfil the following relation, stemming from (4.45),

$$
\boldsymbol{A}|_{d,s,\psi} \, \tilde{\boldsymbol{s}} = \boldsymbol{G}|_{d,s,\psi} \, \tilde{\boldsymbol{g}} = \boldsymbol{1}|_{d,s,\psi} \, g_\psi.
\tag{B.3}
$$

Thus for this subset of measurements, the ground template can only give rise to a constant offset in the time-domain for all samples of the scan. Given that $\boldsymbol{A}$ here simply assigns the pixel amplitudes to the respective time samples without changing their values, there is only one sky mode that reproduces this behaviour: a constant offset in the corresponding sky map, i.e.,

$$\tilde{\boldsymbol{s}} \propto \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \equiv \boldsymbol{1}\,\boldsymbol{s}|_{(d,s,\psi)} \tag{B.4}$$

This demonstrates that, as intuitively expected, the absolute offset of the map produced from these measurements is unavoidably lost as it is degenerate with the ground signal, g_ψ.

Let us consider another data subset taken by detector d', during scan s' and with the azimuth coordinate corresponding to bin ψ'. The data for this scan can be expressed by a relation analogous to (B.1). If either the scan, the detector or the azimuthal bin is different between these two data subsets, then the subsets will have independent, and *a priori* different, ground-pickup amplitudes g_ψ and $g_{\psi'}$. If the sky observed during these scans overlaps, then the combined data set, $d_{\{d,s,\psi\}\cup\{d',s',\psi'\}}$, will have again only one degenerate mode pair,

$$\tilde{\boldsymbol{s}} \propto \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \boldsymbol{1}\,\boldsymbol{s}|_{(d,s,\psi)} \cup \boldsymbol{s}|_{(d',s',\psi')} \quad \text{and} \quad \tilde{\boldsymbol{g}} \propto \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \boldsymbol{1}\,\boldsymbol{g}|_{(d,s,\psi)} \cup \boldsymbol{g}|_{(d',s',\psi')}. \tag{B.5}$$

This reflects the fact that the relative offset of the two sky maps recovered from each of the subsets can be constrained internally owing to the fact they overlap on the sky and have to recover the same sky signal in each common pixel.

Otherwise, if no overlap exists, the data set made of two subsets will have two degenerate pairs of modes, which can be cast as either the absolute offsets of both of the sky patches or as an absolute offset of both of them and a relative offset between them.

This latter situation can happen if the two subsets correspond to the same detector, d, and the same scan, s, but to two different though adjacent azimuthal bins. As described above, their corresponding sky areas will be indeed strictly speaking disjoint. However, as the sky maps are necessarily pixelized, there will be some pixels on the border between two patches which will straddle both of them, constraining their relative offset. The constraint will not be very strong though, and the corresponding degeneracy only weakly broken. An upshot of this is that a map produced from the single constant elevation scan data of a single detector will have as many ill-constrained modes as there are azimuthal bins used to represent the ground pick-up. These modes will become even more ill-constrained if the number of bins increases because the relative offset uncertainty for sky patches corresponding to two extreme bins increases in proportion to $\sqrt{n}_\psi$, while the uncertainty on the offset of two adjacent bins remains roughly unchanged. This scaling reflects the fact that the relative offset between two non-overlapping sky patches with n intermediaries is a result of a random walk of the adjacent patches' offsets, each subject to the same uncertainty (Stompor and White, 2004). Consequently, ill-constrained modes can be suppressed if fewer azimuthal bins are used. Similarly, the uncertainty on the offset of two adjacent bins can be decreased if larger sky pixels are adopted, and more samples from both azimuth bins fall into them. However, the pixel size is typically set by the beam size, while the size of the azimuthal bins is driven by our preconceptions about the ground signal and the structure of the far side lobes. Consequently, whatever freedom is left should be used with care, as potential improvements in statistical uncertainty can be translated into increased ground pick-up residual.

There are two reasons why these ill-constrained modes may be further suppressed in the final maps, combining the data of all the detector and all of the scans. First, for a single scan the additional constraints on the relative offsets of these patches typically also come from the data collected by different detectors. This is because the azimuthal bins are often defined differently for different detectors so the sky patch corresponding to an azimuthal bin of one detector will often overlap with two sky patches corresponding to two different bins of the other detector, thus providing an extra leverage on their relative offset. Second, the sky patches corresponding to fixed azimuthal bins of different scans can have different width because the parallactic angle is different in the two scans (see the previous section). This introduces additional overlaps between the patches of different scans, helping to constrain their relative offsets.

In general the global offset of the final map is expected to be the only truly degenerate mode in the total intensity maps derived from data that are contaminated by ground pickup that requires explicit filtering. Notwithstanding this, the constraints that can be set between the relative offsets of different patches with the same ground pickup are usually inferior to those between different parts of the same

patch, and some ill-constrained large-scale modes should be expected, predominantly in the declination direction.

We note that these conclusions apply qualitatively to any observational site on Earth, including the poles, with the difference that as the single ground bin patches become very narrow in declination the relative offset degeneracies in this direction are broken only by the pixel effects. By contrast, the bin size only plays a role in breaking the degeneracies in the azimuthal direction.

It is important to appreciate the role of the assumptions in breaking these (near) degeneracies. The choices made about the sky pixel size, the pixelization itself, the binning, and the size of the bins, all impact the degeneracies and can be used, or abused, to break them. In addition, the offset degeneracies can be broken if a less flexible model for the ground signal is used, for instance if we impose a prior constraint on the relative change of the ground signal from one bin to the next. The key parameter in such cases would be the assumed coherence length for the ground signal.

B.2 Polarization-sensitive measurements.

The situation for polarization-sensitive observations is potentially more complex due to the more complex form of the pointing matrix. However, it is qualitatively similar to the total intensity case. For concreteness, we discuss the case with three Stokes parameters contributing to the measurements, and thus with the pointing matrix as defined in (4.1). As the polarization orientation may change during the operations, i.e., if some polarization modulator is employed, we will introduce a different ground pick-up amplitude not only for each azimuthal orientation of the instrument but also for each different position of the polarizer, as defined in the instrument coordinates. For simplicity, we will however keep on using a single azimuthal bin number, ψ, to distinguish between the ground signal amplitudes.

We again focus on a single constant elevation scan and a single detector. Our data model is then again given by (B.1), where the same ground pickup is added to each measurement. The major qualitative difference from the total intensity case is that in the polarization-sensitive case the pointing matrix elements may be different from sample to sample, even for samples falling into the same sky pixel on a single crossing, as could be the case if fast rotating half-wave plate were employed to modulate the signal. However, for the data subset selected above the angle of the polarizer is fixed in the instrument frame, so the change of the pointing matrix elements, defined by the polarizer orientation but with respect to the sky coordinates, can be only related to the parallactic angle change with the azimuth of the observation. Typically the angle change within a range of azimuths corresponding to a single sky pixel can be safely neglected and we may assign a single polarizer angle as measured with respect to the sky coordinates for each pixel observed with the data subset. These angles may be somewhat different for two different pixels if these are observed at different azimuths, but as the latter have to fall within a single ground template bin, the bins would need to be rather broad to make such an effect important. Nonetheless, henceforth we assume that for the data subset as defined earlier and characterized by the same ground-pickup amplitude, the polarizer's angle in the sky coordinates, and thus the mixing matrix elements, may at most depend on the observed sky pixels and will have a unique value for all observations falling within the same pixel. We note that such small angle variations do not appear if the observations are conducted from the Earth's poles.

As in the total intensity case, the degenerate sky modes have to be able to mimic an offset in time-domain data. This can be the case for three linearly independent sky defined as,

$$\tilde{s}_I \propto \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \tilde{s}_Q \propto \begin{bmatrix} 0 \\ c_0^{-1} \\ 0 \\ \vdots \\ 0 \\ c_{n_p-1}^{-1} \\ 0 \end{bmatrix}, \qquad \tilde{s}_U \propto \begin{bmatrix} 0 \\ 0 \\ s_0^{-1} \\ \vdots \\ 0 \\ 0 \\ s_{n_p-1} \end{bmatrix}, \qquad (B.6)$$

where,

$$s_p \equiv \sin 2\varphi_p \qquad (B.7)$$
$$c_p \equiv \cos 2\varphi_p, \qquad (B.8)$$

and φ_p stands for a polarizer angle in the sky coordinate in pixel $p \,(= 0, \dots, n_p - 1)$. Each of these sky modes is a vector of n_p triples where the elements of each triple correspond to the I, Q, and U Stokes

parameters. We note that if the sky rotation is negligible across the sky patch covered by the scan these three modes correspond to map offsets of the maps of the respective Stokes parameters.

Within our data subset, each pixel is observed with only a single orientation of the polarizer and we thus cannot estimate all three Stokes parameters separately, but merely their linear combination, $I + Q\cos 2\varphi_p + U\sin 2\varphi_p$, even if no ground pickup is considered. The corresponding two-dimensional degeneracy space is a sub-space of the three-dimensional space spanned by the vectors defined in Eq. (B.6). Consequently, adding the ground pick-up merely adds one degenerate vector to the map-making problem, corresponding to the the total offset of the linear combination of the Stokes parameters, i.e., $I + Q\cos 2\varphi_p + U\sin 2\varphi_p$[1].

To recover all the Stokes parameters from data modeled as in (4.1), we need at least three visits to each pixel with a different orientation of the polarizer. These can be provided by other detectors in the focal plane during the same or different constant elevation scans, or come from the same detector if its polarizer direction is modulated either on short or long timescales. In all these cases the new data will have not only a different polarization angle but also potentially a different ground-pickup. Each of these extra data sets can likewise have up to three degenerate sky modes, which for data subset $i\,(=0,1,2)$, we denote as $\tilde{\boldsymbol{s}}_{\boldsymbol{I}}^{(i)}$, $\tilde{\boldsymbol{s}}_{\boldsymbol{Q}}^{(i)}$ and $\tilde{\boldsymbol{s}}_{\boldsymbol{U}}^{(i)}$, respectively. For these data sets considered together, however only the intensity offset, $\tilde{\boldsymbol{s}}_{\boldsymbol{I}} = \tilde{\boldsymbol{s}}_{\boldsymbol{I}}^{(0)} = \tilde{\boldsymbol{s}}_{\boldsymbol{I}}^{(1)} = \tilde{\boldsymbol{s}}_{\boldsymbol{I}}^{(2)}$, always leads to degeneracy, while $\tilde{\boldsymbol{s}}_{\boldsymbol{Q}}^{(i)}$ and $\tilde{\boldsymbol{s}}_{\boldsymbol{U}}^{(i)}$ will only do so if the polarizer angles for each data subset are effectively the same for all observed pixels, and therefore

$$\tilde{\boldsymbol{s}}_{\boldsymbol{Q}}^{(i)} \propto \tilde{\boldsymbol{s}}_{\boldsymbol{Q}} \equiv \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \tilde{\boldsymbol{s}}_{\boldsymbol{U}}^{(i)} \propto \tilde{\boldsymbol{s}}_{\boldsymbol{U}} \equiv \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \text{for } i = 0, 1, 2. \tag{B.9}$$

In this case each of the recovered maps of the Stokes parameters will have an arbitrary offset corresponding to three degenerate vectors, $\tilde{\boldsymbol{s}}_{\boldsymbol{I}}$, $\tilde{\boldsymbol{s}}_{\boldsymbol{Q}}$, and $\tilde{\boldsymbol{s}}_{\boldsymbol{U}}$, as defined.

If the angles do change somewhat from pixel to pixel within a single data subset, i.e., when the change of the parallactic angle within the azimuthal bin is not negligible, only the total intensity map will have an arbitrary offset. This is because in this case $\boldsymbol{A}^{(i)}\,\tilde{\boldsymbol{s}}_{\boldsymbol{Q}}^{(j)}$ and $\boldsymbol{A}^{(i)}\,\tilde{\boldsymbol{s}}_{\boldsymbol{U}}^{(j)}$ are time-domain vectors with elements which depend on time in a non-trivial way. Here, $\boldsymbol{A}^{(i)}$ is a pointing matrix specific to subset i, while the combined pointing matrix for the three subset is given by,

$$\boldsymbol{A} \equiv \begin{bmatrix} \boldsymbol{A}^{(0)} \\ \boldsymbol{A}^{(1)} \\ \boldsymbol{A}^{(2)} \end{bmatrix}. \tag{B.10}$$

We therefore also see that the time-domain vectors, $\boldsymbol{A}\,\tilde{\boldsymbol{s}}_{\boldsymbol{Q}}^{(j)}$ and $\boldsymbol{A}\,\tilde{\boldsymbol{s}}_{\boldsymbol{U}}^{(j)}$, are non-trivial and therefore cannot typically be mimicked by three ground template offsets and the degeneracy condition in (4.45) can not be fulfilled.

However, as the angle change due to the sky rotation is typically small, $\tilde{\boldsymbol{s}}_{\boldsymbol{Q}}$ and $\tilde{\boldsymbol{s}}_{\boldsymbol{U}}$ may be potentially ill-constrained, even if not strictly singular, and the offsets of the Q and U maps may be very uncertain.

The offsets between sky patches corresponding to adjacent ground template bins during the same constant elevation scan can be further constrained as in the case of the total intensity only measurements. The potential degeneracies can then be suppressed with the help of data from the other detectors and/or different scans, although again a natural expectation is that there will be long sky modes in the declination direction which may be ill-constrained.

If the observation is taken from the Earth's poles, the maps recovered from the three subsets of the data taken at the same elevation will have all three degenerate offsets, which will propagate to the final maps combining all the data. In addition, the relative offsets between the sky patches taken at different elevation will only be set by the presence of pixels common to both patches and therefore will lead to long modes in declination which will be ill-constrained.

---

[1] We also note that even if one of the cosines, c_p, or sines, s_p, happens to be zero, and therefore only two of the three modes in (B.6) are indeed degenerate, the latter statement remains true and the loss of information is the same in all these cases.

Bibliography

R. Adam, B. Comis, I. Bartalucci, et al. High angular resolution Sunyaev-Zel'dovich observations of MACS J1423.8+2404 with NIKA: multi-wavelength analysis. pages 1–16, 2015. ISSN 14320746. doi: 10.1051/0004-6361/201527616.

R. Adam, B. Comis, J.-F. Macías-Pérez, et al. Pressure distribution of the high-redshift cluster of galaxies CL J1226.9+3332 with NIKA. *A&A*, 576:A12, Apr. 2015. doi: 10.1051/0004-6361/201425140.

R. Adam, I. Bartalucci, G. W. Pratt, et al. Mapping the kinetic Sunyaev-Zel'dovich effect toward MACS J0717.5+3745 with NIKA. *ArXiv e-prints*, June 2016.

N. Aghanim, S. Majumdar, and J. Silk. Secondary anisotropies of the CMB. *Reports on Progress in Physics*, 71(6):066902, 2008. ISSN 0034-4885. doi: 10.1088/0034-4885/71/6/066902.

R. A. Alpher and R. C. Herman. On the relative abundance of the elements. *Phys. Rev.*, 74:1737–1742, Dec 1948. doi: 10.1103/PhysRev.74.1737.

E. Anderson, Z. Bai, C. Bischof, et al. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999. ISBN 0-89871-447-8 (paperback).

D. Araujo, C. Bischoff, A. Brizius, et al. SECOND SEASON QUIET OBSERVATIONS: MEASUREMENTS OF THE COSMIC MICROWAVE BACKGROUND POLARIZATION POWER SPECTRUM AT 95 GHz. *The Astrophysical Journal*, 760(2):145, dec 2012. ISSN 0004-637X. doi: 10.1088/0004-637X/760/2/145.

K. Arnold, P. A. R. Ade, A. E. Anthony, et al. The bolometric focal plane array of the POLARBEAR CMB experiment. In W. S. Holland, editor, *Spie . . .*, page 84521D, sep 2012. ISBN 9780819491534. doi: 10.1117/12.927057.

K. Arnold, N. Stebor, P. A. R. Ade, et al. The Simons Array: expanding POLARBEAR to three multi-chroic telescopes. page 91531F, aug 2014. doi: 10.1117/12.2057332.

J. Aumont, L. Conversi, C. Thum, et al. Measurement of the Crab nebula polarization at 90 GHz as a calibrator for CMB experiments. *A&A*, 514:A70, May 2010. doi: 10.1051/0004-6361/200913834.

J. M. Bardeen. Gauge-invariant cosmological perturbations. *Physical Review D*, 22(8):1882–1905, oct 1980. ISSN 0556-2821. doi: 10.1103/PhysRevD.22.1882.

B. A. Bassett, S. Tsujikawa, and D. Wands. Inflation Dynamics and Reheating. *Reviews of Modern Physics*, 78(2):537–589, jul 2005. ISSN 00346861. doi: 10.1103/RevModPhys.78.537.

D. Baumann. INFLATION. In *Physics of the Large and the Small*, pages 523–686. WORLD SCIENTIFIC, mar 2011. ISBN 978-981-4327-17-6. doi: 10.1142/9789814327183_0010.

D. Baumann and H. V. Peiris. Cosmological Inflation: Theory and Observations. *Advanced Science Letters*, 2(2):105–120, oct 2008. ISSN 19366612. doi: 10.1166/asl.2009.1019.

E. J. Baxter, R. Keisler, S. Dodelson, et al. A Measurement of Gravitational Lensing of the Cosmic Microwave Background by Galaxy Clusters Using Data from the South Pole Telescope. *ApJ*, 806:247, June 2015. doi: 10.1088/0004-637X/806/2/247.

C. L. Bennett, D. Larson, J. L. Weiland, et al. NINE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP) OBSERVATIONS : FINAL MAPS AND RESULTS. 20, 2013. doi: 10.1088/0067-0049/208/2/20.

M. Béthermin, E. Le Floc'h, O. Ilbert, et al. HerMES: deep number counts at 250 μm, 350 μm and 500 μm in the COSMOS and GOODS-N fields and the build-up of the cosmic infrared background. *A&A*, 542:A58, June 2012.

M. Betoule, R. Kessler, J. Guy, et al. Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *Astronomy & Astrophysics*, submitted:30, 2014. ISSN 14320746. doi: 10.1051/0004-6361/201423413.

BICEP2 Collaboration. Detection of B-Mode Polarization at Degree Angular Scales by BICEP2. *Physical Review Letters*, 112 (24):241101, June 2014. doi: 10.1103/PhysRevLett.112.241101.

M. Birkinshaw. The Sunyaev-Zel'dovich effect. *Phys. Rep.*, 310:97–195, Mar. 1999. doi: 10.1016/S0370-1573(98)00080-5.

C. Bischoff, L. Hyatt, J. J. McMahon, et al. New Measurements of Fine-Scale CMB Polarization Power Spectra from CAPMAP at Both 40 and 90 GHz. *The Astrophysical Journal*, 684(2):771, 2008. ISSN 0004-637X. doi: 10.1086/590487.

C. Bischoff, A. Brizius, I. Buder, et al. FIRST SEASON QUIET OBSERVATIONS: MEASUREMENTS OF COSMIC MICROWAVE BACKGROUND POLARIZATION POWER SPECTRA AT 43 GHz IN THE MULTIPOLE RANGE 25 < ell < 475. *The Astrophysical Journal*, 741(2):111, nov 2011. ISSN 0004-637X. doi: 10.1088/0004-637X/741/2/111.

L. S. Blackford, J. Choi, A. Cleary, et al. *ScaLAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997. ISBN 0-89871-397-8 (paperback).

N. W. Boggess, J. C. Mather, R. Weiss, et al. The COBE mission - Its design and performance two years after launch. *ApJ*, 397: 420–429, Oct. 1992. doi: 10.1086/171797.

M. Bonamente, M. Joy, S. LaRoque, et al. Determination of the Cosmic Distance Scale from Sunyaev-Zel'dovich Effect and Chandra X-Ray Measurements of High-Redshift Galaxy Clusters. \apj, 647:25–54, 2006. ISSN 0004-637X. doi: 10.1086/505291.

J. R. Bond and G. Efstathiou. Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter. *ApJ*, 285:L45–L48, Oct. 1984. doi: 10.1086/184362.

J. R. Bond and G. Efstathiou. The statistics of cosmic background radiation fluctuations. *Monthly Notices of the Royal Astronomical Society*, 226(3):655–687, jun 1987. ISSN 0035-8711. doi: 10.1093/mnras/226.3.655.

J. R. Bond, a. H. Jaffe, and L. Knox. Estimating the Power Spectrum of the Cosmic Microwave Background. 57(4):23, 1997. ISSN 05562821. doi: 10.1103/PhysRevD.57.2117.

M. L. Brown, P. G. Castro, and a. N. Taylor. Cosmic microwave background temperature and polarization pseudo- C estimators and covariances. *Monthly Notices of the Royal Astronomical Society*, 360(4):1262–1280, jul 2005. ISSN 00358711. doi: 10.1111/j.1365-2966.2005.09111.x.

M. L. Brown, P. Ade, J. Bock, et al. IMPROVED MEASUREMENTS OF THE TEMPERATURE AND POLARIZATION OF THE COSMIC MICROWAVE BACKGROUND FROM QUaD. *The Astrophysical Journal*, 705(1):978–999, 2009. ISSN 0004-637X. doi: 10.1088/0004-637X/705/1/978.

E. F. Bunn, M. Zaldarriaga, M. Tegmark, and A. De Oliveira-Costa. E/B decomposition of finite pixelized CMB maps. *Physical Review D*, 67(2):023501, jan 2003. ISSN 05562821. doi: 10.1103/PhysRevD.67.023501.

C. Burigana, M. Malaspina, N. Mandolesi, et al. A preliminary study on destriping techniques of PLANCK/LFI measurements versus observational strategy. (November):27, jun 1999.

N. G. Busca, T. Delubac, J. Rich, et al. Baryon Acoustic Oscillations in the Ly-\alpha\ forest of BOSS quasars. *eprint arXiv*, 1211:2616, 2012. ISSN 14320746. doi: 10.1051/0004-6361/201220724.

M. Calabrese, C. Carbone, G. Fabbian, M. Baldi, and C. Baccigalupi. Multiple Lensing of the Cosmic Microwave Background anisotropies. page 33, 2014. ISSN 1475-7516. doi: 10.1088/1475-7516/2015/03/049.

C. M. Cantalupo, J. D. Borrill, A. H. Jaffe, T. S. Kisner, and R. Stompor. MADmap: A Massively Parallel Maximum Likelihood Cosmic Microwave Background Map-maker. *ApJS*, 187:212–227, Mar. 2010. doi: 10.1088/0067-0049/187/1/212.

C. Carbone, C. Baccigalupi, M. Bartelmann, S. Matarrese, and V. Springel. Lensed CMB temperature and polarization maps from the Millennium Simulation. *MNRAS*, 396:668–679, June 2009. doi: 10.1111/j.1365-2966.2009.14746.x.

J. E. Carlstrom, G. P. Holder, and E. D. Reese. Cosmology with the Sunyaev-Zel'dovich Effect. 2002.

A. Challinor. CMB anisotropy science: a review. *IAU Symp.*, 288:42, 2013. doi: 10.1017/S1743921312016663.

H. C. Chiang, P. A. R. Ade, D. Barkats, et al. MEASUREMENT OF COSMIC MICROWAVE BACKGROUND POLARIZATION POWER SPECTRA FROM TWO YEARS OF BICEP DATA. *The Astrophysical Journal*, 711(2):1123–1140, mar 2010. ISSN 0004-637X. doi: 10.1088/0004-637X/711/2/1123.

J. Chluba. Which spectral distortions does ΛCDM actually predict? *MNRAS*, 460:227–239, July 2016. doi: 10.1093/mnras/stw945.

A. T. Crites, J. W. Henning, P. A. R. Ade, et al. MEASUREMENTS OF E-MODE POLARIZATION AND TEMPERATURE-E-MODE CORRELATION IN THE COSMIC MICROWAVE BACKGROUND FROM 100 SQUARE DEGREES OF SPTPOL DATA. 36, 2015. doi: 10.1088/0004-637X/805/1/36.

T. Culverhouse, P. Ade, J. Bock, et al. The QUaD Galactic Plane Survey. I. Maps and Analysis of Diffuse Emission. *ApJ*, 722: 1057–1077, Oct. 2010. doi: 10.1088/0004-637X/722/2/1057.

R. H. Cyburt, B. D. Fields, and K. A. Olive. An update on the big bang nucleosynthesis prediction for ^{7}Li: the problem worsens. *J. Cosmology Astropart. Phys.*, 11:012, Nov. 2008. doi: 10.1088/1475-7516/2008/11/012.

S. Das and P. Bode. A Large Sky Simulation of the Gravitational Lensing of the Cosmic Microwave Background. *ApJ*, 682:1-13, July 2008. doi: 10.1086/589638.

S. Das, B. D. Sherwin, P. Aguirre, et al. Detection of the power spectrum of cosmic microwave background lensing by the atacama cosmology telescope. *Physical Review Letters*, 107(July):1–5, 2011. ISSN 00319007. doi: 10.1103/PhysRevLett.107.021301.

S. Das, T. Louis, M. R. Nolta, et al. The Atacama Cosmology Telescope: temperature and gravitational lensing power spectrum measurements from three seasons of data. *Journal of Cosmology and Astroparticle Physics*, 2014:014, 2014. ISSN 1475-7516. doi: 10.1088/1475-7516/2014/04/014.

G. de Gasperis, A. Balbi, P. Cabella, P. Natoli, and N. Vittorio. ROMA: A map-making algorithm for polarised CMB data sets. *A&A*, 436:1159–1165, June 2005. doi: 10.1051/0004-6361:20042512.

G. de Zotti, R. Ricci, D. Mesa, et al. Predictions for high-frequency radio surveys of extragalactic sources. *A&A*, 431:893–903, Mar. 2005. doi: 10.1051/0004-6361:20042108.

M. Dobbs, E. Bissonnette, and H. Spieler. Digital Frequency Domain Multiplexer for Millimeter-Wavelength Telescopes. *IEEE Transactions on Nuclear Science*, 55:21–26, 2008. doi: 10.1109/TNS.2007.911601.

S. Dodelson. *Modern Cosmology*. Academic Press. Academic Press, 2003. ISBN 9780122191411.

C. Dragone. Offset multireflector antennas with perfect pattern symmetry and polarization discrimination. *Bell System Technical Journal*, 57(7):2663–2684, 1978. ISSN 1538-7305. doi: 10.1002/j.1538-7305.1978.tb02171.x.

C. C. Dyer. The gravitational perturbation of the cosmic background radiation by density concentrations. *MNRAS*, 175:429–447, May 1976. doi: 10.1093/mnras/175.2.429.

S. Eales, L. Dunne, and D. Clements. The Herschel ATLAS. *Publications of the ...*, pages 499–515, 2010. doi: 10.1086/653086.

G. F. R. Ellis and M. Bruni. Covariant and gauge-invariant approach to cosmological density fluctuations. *Phys. Rev. D*, 40: 1804–1818, Sep 1989. doi: 10.1103/PhysRevD.40.1804.

J. Errard, F. Stivoli, and R. Stompor. Publisher's Note: Framework for performance forecasting and optimization of CMB B-mode observations in the presence of astrophysical foregrounds [Phys. Rev. D 84, 063005 (2011)]. *Phys. Rev. D*, 84(6):069907, Sept. 2011. doi: 10.1103/PhysRevD.84.069907.

J. Errard, S. M. Feeney, H. V. Peiris, and A. H. Jaffe. Robust forecasts on fundamental physics from the foreground-obscured, gravitationally-lensed CMB polarization. *J. Cosmology Astropart. Phys.*, 3:052, Mar. 2016. doi: 10.1088/1475-7516/2016/03/052.

P. C. Farese, G. Dall'Oglio, J. O. Gundersen, et al. COMPASS: An Upper Limit on Cosmic Microwave Background Polarization at an Angular Scale of 20'. *ApJ*, 610:625–634, Aug. 2004. doi: 10.1086/421837.

P. G. Ferreira and A. H. Jaffe. Simultaneous estimation of noise and signal in cosmic microwave background experiments. *MNRAS*, 312:89–102, Feb. 2000. doi: 10.1046/j.1365-8711.2000.03108.x.

A. Ferté, J. Grain, M. Tristram, and R. Stompor. Efficiency of pseudospectrum methods for estimation of the cosmic microwave background B-mode power spectrum. *Phys. Rev. D*, 88(2):023524, July 2013. doi: 10.1103/PhysRevD.88.023524.

A. Ferté, J. Peloton, J. Grain, and R. Stompor. Detecting the tensor-to-scalar ratio with the pure pseudospectrum reconstruction of B -mode. *Phys. Rev. D*, 92(8):083510, Oct. 2015. doi: 10.1103/PhysRevD.92.083510.

P. Fosalba, E. Gaztañaga, F. J. Castander, and M. Manera. The onion universe: all sky lightcone simulations in spherical shells. *MNRAS*, 391:435–446, Nov. 2008. doi: 10.1111/j.1365-2966.2008.13910.x.

W. L. Freedman, B. F. Madore, V. Scowcroft, et al. Carnegie Hubble Program: a Mid-Infrared Calibration of the Hubble Constant. *The Astrophysical Journal*, 758(1):24, 2012. ISSN 0004-637X. doi: 10.1088/0004-637X/758/1/24.

S. Fukuda, Y. Fukuda, M. Ishitsuka, et al. Tau Neutrinos Favored over Sterile Neutrinos in Atmospheric Muon Neutrino Oscillations. *Physical Review Letters*, 85(19):3999–4003, nov 2000. ISSN 0031-9007. doi: 10.1103/PhysRevLett.85.3999.

E. M. George, C. L. Reichardt, K. A. Aird, et al. A Measurement of Secondary Cosmic Microwave Background Anisotropies from the 2500 Square-degree SPT-SZ Survey. *ApJ*, 799:177, Feb. 2015. doi: 10.1088/0004-637X/799/2/177.

M. Gervasi, M. Zannoni, A. Tartari, G. Boella, and G. Sironi. TRIS. II. Search for CMB Spectral Distortions at 0.60, 0.82, and 2.5 GHz. *ApJ*, 688:24-31, Nov. 2008. doi: 10.1086/592134.

E. B. Gliner. Algebraic Properties of the Energy-momentum Tensor and Vacuum-like States o$^+$ Matter. *Soviet Journal of Experimental and Theoretical Physics*, 22:378, Feb. 1966.

J. N. Goldberg. Spin-s Spherical Harmonics and e. *Journal of Mathematical Physics*, 8(11):2155, 1967. ISSN 00222488. doi: 10.1063/1.1705135.

G. H. Golub and C. F. van Loan. *Matrix computations*. Johns Hopkins University Press, 1996.

K. M. Górski, E. Hivon, A. J. Banday, et al. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *ApJ*, 622:759–771, Apr. 2005. doi: 10.1086/427976.

J. Grain, M. Tristram, and R. Stompor. Polarized CMB power spectrum estimation using the pure pseudo-cross-spectrum approach. *Physical Review D*, 79(12):123515, June 2009. ISSN 1550-7998. doi: 10.1103/PhysRevD.79.123515.

J. Grain, M. Tristram, and R. Stompor. CMB EB and TB cross-spectrum estimation via pseudospectrum techniques. *Phys. Rev. D*, 86(7):076005, Oct. 2012. doi: 10.1103/PhysRevD.86.076005.

H. P. Gush, M. Halpern, and E. H. Wishnow. Rocket measurement of the cosmic-background-radiation mm-wave spectrum. *Physical Review Letters*, 65:537–540, July 1990. doi: 10.1103/PhysRevLett.65.537.

A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, Jan 1981. doi: 10.1103/PhysRevD.23.347.

S. Hanany, P. Ade, A. Balbi, et al. MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on Angular Scales of 10 5deg. *ApJ*, 545:L5–L9, Dec. 2000. doi: 10.1086/317322.

N. Hand, G. E. Addison, E. Aubourg, et al. Evidence of galaxy cluster motions with the kinematic Sunyaev-Zel'dovich effect. *Physical Review Letters*, 109(4), 2012. ISSN 00319007. doi: 10.1103/PhysRevLett.109.041101.

N. Hand, A. Leauthaud, S. Das, et al. First measurement of the cross-correlation of CMB lensing and galaxy lensing. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 91(6):1–11, 2015. ISSN 15502368. doi: 10.1103/PhysRevD.91.062001.

M. Hanke. *Conjugate gradient type methods for ill-posed problems*, volume 327. CRC Press, 1995.

D. Hanson, A. Challinor, and A. Lewis. Weak lensing of the CMB. *General Relativity and Gravitation*, 42(9):2197–2218, 2010. ISSN 00017701. doi: 10.1007/s10714-010-1036-y.

D. Hanson, S. Hoover, A. Crites, et al. Detection of B -Mode Polarization in the Cosmic Microwave Background with Data from the South Pole Telescope. 141301(October):1–6, 2013. doi: 10.1103/PhysRevLett.111.141301.

M. Hasselfield, M. Hilton, T. A. Marriage, et al. The Atacama Cosmology Telescope: Sunyaev-Zel'dovich selected galaxy clusters at 148 GHz from three seasons of data. *J. Cosmology Astropart. Phys.*, 7:008, July 2013. doi: 10.1088/1475-7516/2013/07/008.

M. G. Hauser and E. Dwek. The Cosmic Infrared Background: Measurements and Implications. *ARA&A*, 39:249–307, 2001.

M. G. Hauser and P. J. E. Peebles. Statistical analysis of catalogs of extragalactic objects. II - The Abell catalog of rich clusters. 1973. ISSN 0004-637X. doi: 10.1086/152453.

A. D. Hincks, V. Acquaviva, P. A. R. Ade, et al. The Atacama Cosmology Telescope (ACT): Beam Profiles and First SZ Cluster Maps. *ApJS*, 191:423–438, Dec. 2010. doi: 10.1088/0067-0049/191/2/423.

J. R. Hinderks, P. Ade, J. Bock, et al. QUaD: A HIGH-RESOLUTION COSMIC MICROWAVE BACKGROUND POLARIMETER. *The Astrophysical Journal*, 692(2):1221–1246, 2009. ISSN 0004-637X. doi: 10.1088/0004-637X/692/2/1221.

G. Hinshaw, A. J. Banday, C. L. Bennett, et al. Band power spectra in the cobe dmr four-year anisotropy maps. *The Astrophysical Journal Letters*, 464(1):L17, 1996.

G. Hinshaw, D. N. Spergel, L. Verde, et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: The Angular Power Spectrum. *The Astrophysical Journal Supplement Series*, 148(1):135–159, sep 2003. ISSN 0067-0049. doi: 10.1086/377225.

G. Hinshaw, M. R. Nolta, C. L. Bennett, et al. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature Analysis. *ApJS*, 170:288–334, June 2007. doi: 10.1086/513698.

C. M. Hirata, S. Ho, N. Padmanabhan, U. Seljak, and N. A. Bahcall. Correlation of CMB with large-scale structure. II. Weak lensing. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 78(4):1–33, 2008. ISSN 15507998. doi: 10.1103/PhysRevD.78.043520.

E. Hivon, K. M. Gorski, C. B. Netterfield, et al. MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets. *The Astrophysical Journal*, 567(1):2–17, Mar. 2002. ISSN 0004-637X. doi: 10.1086/338126.

W. Hu. Weak lensing of the CMB: A harmonic approach. 62(January):1–17, 2000.

W. Hu and T. Okamoto. Mass Reconstruction with Cosmic Microwave Background Polarization. *The Astrophysical Journal*, 574 (2):566–574, 2002. ISSN 0004-637X. doi: 10.1086/341110.

W. Hu and N. Sugiyama. Anisotropies in the cosmic microwave background: an analytic approach. *The Astrophysical Journal*, 444:489–506, 1995.

W. Hu, N. Sugiyama, and J. Silk. The Physics of Microwave Background Anisotropies. 1996. ISSN 0028-0836. doi: 10.1038/386037a0.

E. Hubble. a Relation Between Distance and Radial Velocity Among Extra-Galactic Nebulae. *Proceedings of the National Academy of Sciences of the United States of America*, 15(3):168–173, 1929. ISSN 0027-8424. doi: 10.1073/pnas.15.3.168.

G. Hurier, N. Aghanim, M. Douspis, and E. Pointecouteau. Measurement of the T_{CMB} evolution from the Sunyaev-Zel'dovich effect. *A&A*, 561:A143, Jan. 2014. doi: 10.1051/0004-6361/201322632.

F. Iocco, G. Mangano, G. Miele, O. Pisanti, and P. D. Serpico. Primordial nucleosynthesis: From precision cosmology to fundamental physics. *Phys. Rep.*, 472:1–76, Mar. 2009. doi: 10.1016/j.physrep.2009.02.002.

M. A. Janssen and S. Gulkis. Mapping the sky with the COBE differential microwave radiometers. In M. Signore and C. Dupraz, editors, *NATO Advanced Science Institutes (ASI) Series C*, volume 359 of *NATO Advanced Science Institutes (ASI) Series C*, pages 391–408, 1992.

M. Kamionkowski. *Dark matter and dark energy.* 2007. ISBN 0000000000000. doi: 10.1007/978-3-540-31535-3_5.

A. Kashlinsky. Cosmic infrared background and early galaxy evolution [review article]. *Phys. Rep.*, 409:361–438, Apr. 2005.

B. G. Keating, M. Shimon, and A. P. S. Yadav. SELF-CALIBRATION OF COSMIC MICROWAVE BACKGROUND POLARIZATION EXPERIMENTS. *ApJ*, 762(2):L23, jan 2013. ISSN 2041-8205. doi: 10.1088/2041-8205/762/2/L23.

E. Keihänen, H. Kurki-Suonio, T. Poutanen, D. Maino, and C. Burigana. A maximum likelihood approach to the destriping technique. *A&A*, 428:287–298, Dec. 2004. doi: 10.1051/0004-6361:200400060.

E. Keihänen, R. Keskitalo, H. Kurki-Suonio, T. Poutanen, and A.-S. Sirviö. Making cosmic microwave background temperature and polarization maps with MADAM. *A&A*, 510:A57, Feb. 2010. doi: 10.1051/0004-6361/200912813.

R. Keisler, S. Hoover, N. Harrington, et al. Measurements of Sub-degree B-mode Polarization in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data. *astro-ph.CO*, 2015.

Z. D. Kermish, P. Ade, A. Anthony, et al. The POLARBEAR experiment. In W. S. Holland, editor, *Spie . . .*, page 84521C, sep 2012. doi: 10.1117/12.926354.

T. Kitayama. Cosmological and astrophysical implications of the Sunyaev-Zel'dovich effect. *Progress of Theoretical and Experimental Physics*, 2014(6):1–26, 2014. ISSN 20503911. doi: 10.1093/ptep/ptu055.

a. Kogut, D. N. Spergel, C. Barnes, et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature-Polarization Correlation. *The Astrophysical Journal Supplement Series*, 148(1):161–173, sep 2003. ISSN 0067-0049. doi: 10.1086/377219.

A. Kogut, D. N. Spergel, C. Barnes, et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature-Polarization Correlation. *ApJS*, 148:161–173, Sept. 2003. doi: 10.1086/377219.

A. Kogut, D. J. Fixsen, D. T. Chuss, et al. The Primordial Inflation Explorer (PIXIE): a nulling polarimeter for cosmic microwave background observations. *J. Cosmology Astropart. Phys.*, 7:025, July 2011. doi: 10.1088/1475-7516/2011/07/025.

E. Kolb and M. Turner. *The Early Universe.* Frontiers in physics. Westview Press, 1994. ISBN 9780813346458.

A. Kompaneets. . *Zh. Eksp. Teor. Fiz.*, 31:876, 1957.

A. Kosowsky. Cosmic microwave background polarization. *Annals of Physics*, 246:49–85, Feb. 1996. doi: 10.1006/aphy.1996.0020.

J. M. Kovac, E. M. Leitch, C. Pryke, et al. Detection of polarization in the cosmic microwave background using DASI. *Nature*, 420(6917):772–787, dec 2002. ISSN 0028-0836. doi: 10.1038/nature01269.

C. T. Kowal. Absolute magnitudes of supernovae. *The Astronomical Journal*, 73:1021, dec 1968. ISSN 00046256. doi: 10.1086/110763.

N. Krachmalnicoff, C. Baccigalupi, J. Aumont, M. Bersanelli, and A. Mennella. Characterization of foreground emission at degree angular scale for CMB B-modes observations. Thermal Dust and Synchrotron signal from Planck and WMAP data. *ArXiv e-prints*, Nov. 2015.

E. Krause and C. M. Hirata. Weak lensing power spectra for precision cosmology. Multiple-deflection, reduced shear, and lensing bias corrections. *A&A*, 523:A28, Nov. 2010. doi: 10.1051/0004-6361/200913524.

L. La Porta, C. Burigana, W. Reich, and P. Reich. The impact of Galactic synchrotron emission on CMB anisotropy measurements. I. Angular power spectrum analysis of total intensity all-sky surveys. *A&A*, 479:641–654, Mar. 2008. doi: 10.1051/0004-6361:20078435.

D. Langlois. Inflation, quantum fluctuations and cosmological perturbations. pages 1–44, may 2004.

A. T. Lee, H. Tran, P. Ade, et al. POLARBEAR: Ultra-high Energy Physics with Measurements of CMB Polarization. In *AIP Conference Proceedings*, volume 1040, pages 66–77. AIP, 2008. ISBN 9780735405677. doi: 10.1063/1.2981555.

E. M. Leitch, J. M. Kovac, N. W. Halverson, et al. Degree Angular Scale Interferometer 3 Year Cosmic Microwave Background Polarization Results. *The Astrophysical Journal*, 624(1):10–20, 2005. ISSN 0004-637X. doi: 10.1086/428825.

a. Lewis and a. Challinor. Weak gravitational lensing of the CMB. *Physics Reports*, 429(1):1–65, jun 2006. ISSN 03701573. doi: 10.1016/j.physrep.2006.03.002.

A. Liddle and D. Lyth. *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press, 2000. ISBN 9780521575980.

A. R. Liddle and D. H. Lyth. COBE, gravitational waves, inflation and extended inflation. *Physics Letters B*, 291(4):391–398, oct 1992. ISSN 03702693. doi: 10.1016/0370-2693(92)91393-N.

A. R. Liddle, P. Parsons, and J. D. Barrow. Formalizing the slow-roll approximation in inflation. *Physical Review D*, 50(12): 7222–7232, dec 1994. ISSN 0556-2821. doi: 10.1103/PhysRevD.50.7222.

A. D. Linde. Inflation and quantum cosmology. (SU-ITP-87-8. PRINT-86-0888-CAL-TECH):59 p, Jun 1986.

M. Madhavacheril, N. Sehgal, R. Allison, et al. Evidence of Lensing of the Cosmic Microwave Background by Dark Matter Halos. *Physical Review Letters*, 114(15):151302, Apr. 2015. doi: 10.1103/PhysRevLett.114.151302.

J. G. Magnum. A telescope pointing algorithm for alma, apr 2014.

D. Maino, C. Burigana, M. Maltoni, et al. The Planck-LFI instrument: Analysis of the 1/f noise and implications for the scanning strategy. *Astronomy & Astrophysics Supplement Series*, 140(3):383–391, 1999. ISSN 0365-0138. doi: 10.1051/aas:1999429.

D. Maino, C. Burigana, K. M. Górski, N. Mandolesi, and M. Bersanelli. Removing $1/f$ noise stripes in cosmic microwave background anisotropy observations. *Astronomy & Astrophysics*, 387(1):356–365, may 2002. ISSN 0004-6361. doi: 10.1051/0004-6361:20020242.

K. A. Malik and D. Wands. Cosmological perturbations. (June):1–51, sep 2008. ISSN 03701573. doi: 10.1016/j.physrep.2009.03.001.

T. A. Marriage, V. Acquaviva, P. A. R. Ade, et al. The Atacama Cosmology Telescope: Sunyaev-Zel'dovich-Selected Galaxy Clusters at 148 GHz in the 2008 Survey. *ApJ*, 737:61, Aug. 2011. doi: 10.1088/0004-637X/737/2/61.

J. C. Mather, E. S. Cheng, R. E. Eplee, Jr., et al. A preliminary measurement of the cosmic microwave background spectrum by the Cosmic Background Explorer (COBE) satellite. *ApJ*, 354:L37–L40, May 1990. doi: 10.1086/185717.

T. Matsumura, Y. Akiba, K. Arnold, et al. Litebird: Mission overview and focal plane layout. *Journal of Low Temperature Physics*, 184(3):824–831, 2016. ISSN 1573-7357. doi: 10.1007/s10909-016-1542-8.

P. D. Mauskopf, P. A. R. Ade, P. de Bernardis, et al. Measurement of a Peak in the Cosmic Microwave Background Power Spectrum from the North American Test Flight of Boomerang. *ApJ*, 536:L59–L62, June 2000. doi: 10.1086/312743.

A. Melchiorri, P. A. R. Ade, P. de Bernardis, et al. A Measurement of Ω from the North American Test Flight of Boomerang. *The Astrophysical Journal*, 536(2):L63–L66, jun 2000. ISSN 0004637X. doi: 10.1086/312744.

Y. Mizugutch, M. Akagawa, and H. Yokoi. Offset Dual Reflector Antenna. *ISA Proceedings*, pages 2–5, 1976.

T. Mroczkowski, M. J. Devlin, S. R. Dicker, et al. New high-resolution Sunyaev-Zel'dovich observations with GBT+MUSTANG. *Mem. Soc. Astron. Italiana*, 82:485, 2011.

V. Mukhanov, H. Feldman, and R. Brandenberger. Theory of cosmological perturbations. *Physics Reports*, 215(5):203 – 333, 1992. ISSN 0370-1573. doi: http://dx.doi.org/10.1016/0370-1573(92)90044-Z.

T. Murphy, E. M. Sadler, R. D. Ekers, et al. The Australia Telescope 20 GHz Survey: the source catalogue. *MNRAS*, 402: 2403–2423, Mar. 2010. doi: 10.1111/j.1365-2966.2009.15961.x.

S. Naess, M. Hasselfield, J. McMahon, et al. The Atacama Cosmology Telescope: CMB Polarization at $200 < \ell < 9000$. *astro-ph preprint*, ArXiv:1405:14, 2014a. ISSN 14757516. doi: 10.1088/1475-7516/2014/10/007.

S. Naess, M. Hasselfield, J. McMahon, et al. The Atacama Cosmology Telescope: CMB Polarization at $200</<9000$. *Astro-Ph Preprint*, ArXiv:1405:14, 2014b. ISSN 14757516. doi: 10.1088/1475-7516/2014/10/007.

R. O'Brient, P. Ade, K. Arnold, et al. A dual-polarized broadband planar antenna and channelizing filter bank for millimeter wavelengths. *Applied Physics Letters*, 102(6):063506, 2013. ISSN 00036951. doi: 10.1063/1.4791692.

T. Okamoto and W. Hu. Cosmic microwave background lensing reconstruction on the full sky. *Physical Review D*, 67(8):083002, 2003. ISSN 0556-2821. doi: 10.1103/PhysRevD.67.083002.

K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014. doi: 10.1088/1674-1137/38/9/090001.

L. Page, G. Hinshaw, E. Komatsu, et al. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis. *The Astrophysical Journal Supplement Series*, 170(2):335–376, 2007. ISSN 0067-0049. doi: 10.1086/513699.

P. J. E. Peebles and J. T. Yu. Primeval Adiabatic Perturbation in an Expanding Universe. *ApJ*, 162:815, Dec. 1970. doi: 10.1086/150713.

J. Peloton. *Data analysis and scientific exploitation of the CMB B-modes experiment, POLARBEAR*. PhD thesis, Université Paris Diderot, 9 2015. Ph.D. Thesis.

a. a. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *The Astrophysical Journal*, 142:419, 1965. ISSN 0004-637X. doi: 10.1086/148307.

S. Perlmutter, G. Aldering, G. Goldhaber, et al. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *ApJ*, 517:565–586, June 1999.

F. Piacentini, P. A. R. Ade, J. J. Bock, et al. *The Astrophysical Journal*, (2):833–839, aug . ISSN 0004-637X. doi: 10.1086/505557.

Planck Collaboration. Planck 2013 results. XX. Cosmology from Sunyaev-Zeldovich cluster counts. *A&A*, 571:A20, Nov. 2014a. doi: 10.1051/0004-6361/201321521.

Planck Collaboration. Planck 2013 results. I. Overview of products and scientific results. 2014b. doi: 10.1051/0004-6361/201321529.

Planck Collaboration. Planck 2013 results. XXIX. The Planck catalogue of Sunyaev-Zeldovich sources. *A&A*, 571:A29, Nov. 2014c. doi: 10.1051/0004-6361/201321523.

Planck Collaboration. Planck 2013 results. XIII. Galactic CO emission. *A&A*, 571:A13, Nov. 2014d. doi: 10.1051/0004-6361/201321553.

Planck Collaboration. Planck 2015 results. VIII. High Frequency Instrument data processing: Calibration and maps. *ArXiv e-prints*, Feb. 2015a.

Planck Collaboration. Planck 2015 results. VI. LFI mapmaking. *ArXiv e-prints*, Feb. 2015b.

Planck Collaboration. Planck 2015 results. IX. Diffuse component separation: CMB maps. *ArXiv e-prints*, Feb. 2015c.

Planck Collaboration. Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. *ArXiv e-prints*, July 2015d.

Planck Collaboration. Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts. *ArXiv e-prints*, Feb. 2015e.

Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. feb 2015f.

Planck Collaboration. Planck 2015 results. I. Overview of products and scientific results. *Physics*, 6:107, feb 2015g. ISSN 1943-2879. doi: 10.1051/0004-6361/201321529.

Planck Collaboration. Planck 2015 results. XXVII. The Second Planck Catalogue of Sunyaev-Zeldovich Sources. *ArXiv e-prints*, Feb. 2015h.

Planck Collaboration. Planck 2015 results. XXII. A map of the thermal Sunyaev-Zeldovich effect. *ArXiv e-prints*, Feb. 2015i.

Planck Collaboration. Planck 2015 results. XIII. Cosmological parameters. *ArXiv e-prints*, Feb. 2015j.

Planck Collaboration. Planck 2016 intermediate results. XLVII. Planck constraints on reionization history. page 19, 2016a.

Planck Collaboration. Planck intermediate results. XXX. The angular power spectrum of polarized dust emission at intermediate and high Galactic latitudes. *A&A*, 586:A133, Feb. 2016b. doi: 10.1051/0004-6361/201425034.

Planck Collaboration, P. A. R. Ade, N. Aghanim, et al. Planck intermediate results. V. Pressure profiles of galaxy clusters from the Sunyaev-Zeldovich effect. *A&A*, 550:A131, Feb. 2013. doi: 10.1051/0004-6361/201220040.

Planck Collaboration, P. A. R. Ade, N. Aghanim, et al. Planck 2015 results. XXV. Diffuse low-frequency Galactic foregrounds. *ArXiv e-prints*, June 2015a.

Planck Collaboration, P. A. R. Ade, N. Aghanim, et al. Planck 2015 results. XX. Constraints on inflation. *ArXiv e-prints*, Feb. 2015b.

Planck Collaboration, P. A. R. Ade, N. Aghanim, et al. Planck 2015 results. XVII. Constraints on primordial non-Gaussianity. *ArXiv e-prints*, Feb. 2015c.

G. Polenta, D. Marinucci, a. Balbi, et al. Unbiased Estimation of an Angular Power Spectrum. 11, 2004. ISSN 14757516. doi: 10.1088/1475-7516/2005/11/001.

D. Poletti, G. Fabbian, M. Le Jeune, et al. Making maps of Cosmic Microwave Background polarization for B-mode studies: the POLARBEAR example. *ArXiv e-prints*, Aug. 2016.

T. Poutanen, D. Maino, H. Kurki-Suonio, E. Keihänen, and E. Hivon. Cosmic microwave background power spectrum estimation with the destriping technique. *MNRAS*, 353:43–58, Sept. 2004. doi: 10.1111/j.1365-2966.2004.08043.x.

G. Pratten and A. Lewis. Impact of post-Born lensing on the CMB. page 25, may 2016.

C. Pryke, P. Ade, J. Bock, et al. SECOND AND THIRD SEASON QUaD COSMIC MICROWAVE BACKGROUND TEMPER-ATURE AND POLARIZATION POWER SPECTRA. *The Astrophysical Journal*, 692(2):1247–1270, 2009. ISSN 0004-637X. doi: 10.1088/0004-637X/692/2/1247.

J.-L. Puget, A. Abergel, J.-P. Bernard, et al. Tentative detection of a cosmic far-infrared background with COBE. *A&A*, 308:L5, Apr. 1996.

QUIET Collaboration. First Season QUIET Observations: Measurements of Cosmic Microwave Background Polarization Power Spectra at 43 GHz in the Multipole Range 25 <= ell <= 475. *ApJ*, 741:111, Nov. 2011. doi: 10.1088/0004-637X/741/2/111.

M. J. Rees and D. W. Sciama. Large-scale Density Inhomogeneities in the Universe. *Nature*, 217:511–516, Feb. 1968. doi: 10.1038/217511a0.

C. L. Reichardt, B. Stalder, L. E. Bleem, et al. Galaxy Clusters Discovered Via the Sunyaev-Zel'Dovich Effect in the First 720 Square Degrees of the South Pole Telescope Survey. *The Astrophysical Journal*, 763(2):127, 2013. ISSN 0004-637X. doi: 10.1088/0004-637X/763/2/127.

A. G. Riess, A. V. Filippenko, P. Challis, et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ*, 116:1009–1038, Sept. 1998.

R. K. Sachs and A. M. Wolfe. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *ApJ*, 147:73, Jan. 1967. doi: 10.1086/148982.

A. D. Sakharov. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Soviet Journal of Experimental and Theoretical Physics Letters*, 5:24, Jan. 1967.

J. Sayers, T. Mroczkowski, M. Zemcov, et al. A Measurement of the Kinetic Sunyaev-Zel'dovich Signal Toward MACS J0717.5+3745. *ApJ*, 778:52, Nov. 2013. doi: 10.1088/0004-637X/778/1/52.

K. K. Schaffer, T. M. Crawford, K. A. Aird, et al. The First Public Release of South Pole Telescope Data: Maps of a 95 deg^2 Field from 2008 Observations. *ApJ*, 743:90, Dec. 2011. doi: 10.1088/0004-637X/743/1/90.

D. N. Schramm and M. S. Turner. Big-bang nucleosynthesis enters the precision era. *Reviews of Modern Physics*, 70:303–318, Jan. 1998. doi: 10.1103/RevModPhys.70.303.

M. Seiffert, D. J. Fixsen, A. Kogut, et al. Interpretation of the ARCADE 2 Absolute Sky Brightness Measurement. *ApJ*, 734:6, June 2011. doi: 10.1088/0004-637X/734/1/6.

U. Seljak and M. Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *The Astrophysical Journal*, 469:437, oct 1996. ISSN 0004-637X. doi: 10.1086/177793.

J. L. Sievers, C. Achermann, J. R. Bond, et al. Implications of the Cosmic Background Imager Polarization Data. *The Astrophysical Journal*, 660(2):15, 2005. ISSN 0004-637X. doi: Doi10.1086/510504.

K. M. Smith. Pseudo-C? estimators which do not mix E and B modes. *Phys. Rev. D*, 74(8):083002, Oct. 2006. doi: 10.1103/PhysRevD.74.083002.

K. M. Smith and M. Zaldarriaga. General solution to the E-B mixing problem. *Phys. Rev. D*, 76(4):043001, Aug. 2007. doi: 10.1103/PhysRevD.76.043001.

K. M. Smith, O. Zahn, and O. Doré. Detection of gravitational lensing in the cosmic microwave background. *Physical Review D*, 76:43510, 2007. ISSN 15507998. doi: 10.1103/PhysRevD.76.043510.

R. E. Smith, J. A. Peacock, A. Jenkins, et al. Stable clustering, the halo model and non-linear cosmological power spectra. *MNRAS*, 341:1311–1332, June 2003. doi: 10.1046/j.1365-8711.2003.06503.x.

G. F. Smoot, C. L. Bennett, A. Kogut, et al. Structure in the COBE differential microwave radiometer first-year maps. *ApJ*, 396: L1–L5, Sept. 1992. doi: 10.1086/186504.

Z. Staniszewski, P. A. R. Ade, K. A. Aird, et al. GALAXY CLUSTERS DISCOVERED WITH A SUNYAEV-ZEL'DOVICH EFFECT SURVEY. *The Astrophysical Journal*, 701(1):32–41, 2009. ISSN 0004-637X. doi: 10.1088/0004-637X/701/1/32.

A. A. Starobinskiǐ. Spectrum of relict gravitational radiation and the early state of the universe. *Soviet Journal of Experimental and Theoretical Physics Letters*, 30:682, Dec. 1979.

G. Steigman. Primordial Nucleosynthesis in the Precision Cosmology Era. *Annual Review of Nuclear and Particle Science*, 57: 463–491, Nov. 2007. doi: 10.1146/annurev.nucl.56.080805.140437.

R. Stompor. Numerical analysis of the cosmic microwave background anisotropies within framework of linearized gravitational instability model. *A&A*, 287:693–712, July 1994.

R. Stompor and M. White. The effects of low temporal frequency modes on minimum variance maps from PLANCK. *A&A*, 419: 783–792, May 2004. doi: 10.1051/0004-6361:20034185.

R. Stompor, M. Abroe, P. Ade, et al. Cosmological Implications of the MAXIMA-1 High-Resolution Cosmic Microwave Background Anisotropy Measurement. *The Astrophysical Journal*, 561(1):L7–L10, 2001a. ISSN 0004637X. doi: 10.1086/324438.

R. Stompor, A. Balbi, J. Borrill, et al. Making maps of the cosmic microwave background: The MAXIMA example. *Physical Review D*, 65(2):022003, Dec. 2001b. ISSN 0556-2821. doi: 10.1103/PhysRevD.65.022003.

R. Stompor, J. Errard, and D. Poletti. Forecasting performance of CMB experiments in the presence of complex foreground contaminations. *ArXiv e-prints*, Sept. 2016.

N. Sugiyama. Cosmic Background Anisotropies in Cold Dark Matter Cosmology. *ApJS*, 100:281, Oct. 1995. doi: 10.1086/192220.

R. A. Sunyaev and I. B. Zeldovich. Microwave background radiation as a probe of the contemporary structure and history of the universe. *ARA&A*, 18:537–560, 1980. doi: 10.1146/annurev.aa.18.090180.002541.

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

A. Suzuki, P. Ade, Y. Akiba, et al. The POLARBEAR-2 Experiment. *Journal of Low Temperature Physics*, 176(5-6):719–725, sep 2014a. ISSN 0022-2291. doi: 10.1007/s10909-014-1112-x.

A. Suzuki, K. Arnold, J. Edwards, et al. Multi-Chroic Dual-Polarization Bolometric Detectors for Studies of the Cosmic Microwave Background. *Journal of Low Temperature Physics*, 176(5-6):650–656, sep 2014b. ISSN 0022-2291. doi: 10.1007/s10909-013-1049-5.

M. Szydlarski, L. Grigori, and R. Stompor. Accelerating the cosmic microwave background map-making procedure through preconditioning. *A&A*, 572:A39, Dec. 2014. doi: 10.1051/0004-6361/201323210.

H. Tashiro. CMB spectral distortions and energy release in the early universe. *Progress of Theoretical and Experimental Physics*, 2014(6):06B107, June 2014. doi: 10.1093/ptep/ptu066.

M. Tegmark. How to measure CMB power spectra without losing information. *Physical Review D*, 64(6):5895–5907, nov 1996. ISSN 05562821. doi: 10.1103/PhysRevD.55.5895.

M. Tegmark. CMB mapping experiments: A designer's guide. *Phys. Rev. D*, 56:4514–4529, Oct. 1997. doi: 10.1103/PhysRevD.56.4514.

T. Polarbear Collaboration. A MEASUREMENT OF THE COSMIC MICROWAVE BACKGROUND B-MODE POLARIZATION POWER SPECTRUM AT SUB-DEGREE SCALES WITH POLARBEAR. *The Astrophysical Journal*, 794(2):171, oct 2014a. ISSN 1538-4357. doi: 10.1088/0004-637X/794/2/171.

T. Polarbear Collaboration. Evidence for gravitational lensing of the cosmic microwave background polarization from cross-correlation with the cosmic infrared background. *Physical Review Letters*, 112(13):1–6, 2014b. ISSN 00319007. doi: 10.1103/PhysRevLett.112.131302.

T. Polarbear Collaboration. POLARBEAR constraints on cosmic birefringence and primordial magnetic fields. *Physical Review D*, 92(12):123509, 2015. ISSN 1550-7998. doi: 10.1103/PhysRevD.92.123509.

The BICEP2/*Keck Array* and *Planck* Collaborations. Joint Analysis of BICEP2/ Keck Array and Planck Data. *Physical Review Letters*, 114(10):101301, mar 2015. ISSN 0031-9007. doi: 10.1103/PhysRevLett.114.101301.

The BICEP2/*Keck Array* Collaboration. Improved Constraints on Cosmology and Foregrounds from BICEP2 and Keck Array Cosmic Microwave Background Data with Inclusion of 95 GHz Band. *Physical Review Letters*, 116(3):1–17, 2016. ISSN 10797114. doi: 10.1103/PhysRevLett.116.031302.

The BICEP2/*Keck Array* Collaborations. BICEP2 / Keck Array VIII: Measurement of gravitational lensing from large-scale B-mode polarization. 2016.

The COrE Collaboration. COrE (Cosmic Origins Explorer) A White Paper. *ArXiv e-prints*, Feb. 2011.

The Planck Collaboration. Planck 2013 results. XVII. Gravitational lensing by large-scale structure. *Astronomy & Astrophysics*, 571:A17, 2014.

The Polarbear Collaboration. a Measurement of the Cosmic Microwave Background B -Mode Polarization Power Spectrum At Sub-Degree Scales With Polarbear. *The Astrophysical Journal*, 794(2):171, 2014. ISSN 1538-4357. doi: 10.1088/0004-637X/794/2/171.

The Polarbear Collaboration. Measurement of the cosmic microwave background polarization lensing power spectrum with the POLARBEAR experiment. *Physical Review Letters*, 113(2):1–7, 2014. ISSN 10797114. doi: 10.1103/PhysRevLett.113.021301.

E. Torbet, M. J. Devlin, W. B. Dorwart, et al. A Measurement of the Angular Power Spectrum of the Microwave Background Made from the High Chilean Andes. *ApJ*, 521:L79–L82, Aug. 1999. doi: 10.1086/312197.

M. Tristram, J. F. Macias-Perez, C. Renault, and D. Santos. XSPECT, estimation of the angular power spectrum by computing cross-power spectra with analytical error bars. *Monthly Notices of the Royal Astronomical Society*, 358(3):833–842, apr 2005. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2005.08760.x.

M. Tristram, C. Filliard, O. Perdereau, et al. Iterative destriping and photometric calibration for Planck-HFI, polarized, multi-detector map-making. *A&A*, 534:A88, Oct. 2011. doi: 10.1051/0004-6361/201116871.

J. P. Uzan, N. Aghanim, and Y. Mellier. Distance duality relation from x-ray and Sunyaev-Zel'dovich observations of clusters. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 70(8), 2004. ISSN 05562821. doi: 10.1103/PhysRevD.70.083533.

A. van Engelen, R. Keisler, O. Zahn, et al. a Measurement of Gravitational Lensing of the Microwave Background Using South Pole Telescope Data. *The Astrophysical Journal*, 756(2):142, 2012. ISSN 0004-637X. doi: 10.1088/0004-637X/756/2/142.

A. van Engelen, B. D. Sherwin, N. Sehgal, et al. the Atacama Cosmology Telescope: Lensing of Cmb Temperature and Polarization Derived From Cosmic Infrared Background Cross-Correlation. *The Astrophysical Journal*, 808(1):7, 2015. ISSN 1538-4357. doi: 10.1088/0004-637X/808/1/7.

K. Vanderlinde, T. M. Crawford, T. de Haan, et al. GALAXY CLUSTERS SELECTED WITH THE SUNYAEV-ZEL'DOVICH EFFECT FROM 2008 SOUTH POLE TELESCOPE OBSERVATIONS. *The Astrophysical Journal*, 722(2):1180–1196, 2010. ISSN 0004-637X. doi: 10.1088/0004-637X/722/2/1180.

C. G. R. Wallis, A. Bonaldi, M. L. Brown, and R. A. Battye. A new map-making algorithm for CMB polarisation experiments. 12(July):1–12, mar 2015.

B. D. Wandelt, D. L. Larson, and A. Lakshminarayanan. Global, exact cosmic microwave background data analysis using Gibbs sampling. *Phys. Rev. D*, 70(8):083511, Oct. 2004. doi: 10.1103/PhysRevD.70.083511.

J. L. Weiland, N. Odegard, R. S. Hill, et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Planets and Celestial Calibration Sources. *ApJS*, 192:19, Feb. 2011. doi: 10.1088/0067-0049/192/2/19.

R. Williamson, B. A. Benson, F. W. High, et al. A Sunyaev-Zel'dovich-selected Sample of the Most Massive Galaxy Clusters in the 2500 deg^2 South Pole Telescope Survey. *ApJ*, 738:139, Sept. 2011. doi: 10.1088/0004-637X/738/2/139.

E. L. Wright, G. Hinshaw, and C. L. Bennett. Producing Megapixel Cosmic Microwave Background from Differential Radiometer Data. *ApJ*, 458:L53, Feb. 1996. doi: 10.1086/309927.

J. M. Wrobel, A. R. Patnaik, I. W. A. Browne, and P. N. Wilkinson. A Catalog of 2118 Compact Radio Sources in the Northern Sky. In *American Astronomical Society Meeting Abstracts*, volume 30 of *Bulletin of the American Astronomical Society*, page 1308, Dec. 1998.

M. Zaldarriaga and U. Seljak. All-sky analysis of polarization in the microwave background. *Physical Review D*, 55(4):1830–1840, feb 1997. ISSN 0556-2821. doi: 10.1103/PhysRevD.55.1830.

Y. B. Zel'dovich. Special Issue: the Cosmological Constant and the Theory of Elementary Particles. *Soviet Physics Uspekhi*, 11: 381–393, Mar. 1968. doi: 10.1070/PU1968v011n03ABEH003927.

Y. B. Zeldovich and R. A. Sunyaev. The Interaction of Matter and Radiation in a Hot-Model Universe. *Ap&SS*, 4:301–316, July 1969. doi: 10.1007/BF00661821.