Hunting B modes in CMB polarization observations

Thesis submitted for the degree of
Doctor Philosophiæ

October 2008

Federico Stivoli
International School for Advanced Studies
Via Beirut 2-4, 34014 Trieste, Italy.
E-mail: *stivoli@sissa.it*

Supervisor:
Professor Carlo Baccigalupi

Dedicated to Mamma e Papà

Contents

The work presented in this thesis can be found in the following publications:

1. Separating polarized cosmological and galactic emissions for CMB B-mode polarization experiments - F. Stivoli, C. Baccigalupi, D. Maino, R. Stompor

   Monthly Notices of the Royal Astronomical Society, 2006 372, 615

2. Component separation with FastICA - F. Stivoli

   CMB and Physics of the Early Universe, 20-22 April 2006 Ischia, Italy

3. CMB signal in WMAP 3yr data with FastICA - D. Maino, S. Donzelli, A. Banday, F. Stivoli, C. Baccigalupi

   Monthly Notices of the Royal Astronomical Society, 2007 374, 1207

4. WMAP 3yr data with the CCA: anomalous emission and impact of component separation on the CMB power spectrum - A. Bonaldi, S.Ricciardi, S. Leach, F. Stivoli, C. Baccigalupi, G. De Zotti

   Monthly Notices of the Royal Astronomical Society, 2007 382 1791B

5. Component Separation Method For The Planck Mission - S. Leach et al. Submitted to Astronomy and Astrophysics, 2008 - astro-ph/0805.0269

6. Maximum Likelihood algorithm for component separation in CMB experiments - R. Stompor, S. Leach, F. Stivoli, C. Baccigalupi

   Submitted to Monthly Notices of the Royal Astronomical Society, 2008, - astro-ph/0804.2645

7. FastICA: Research and Development - R. Stompor, F. Stivoli, C. Baccigalupi, S. Leach

   in preparation.

Introduction

In the last decades cosmologists have seen a substantial growth in the understanding of the early stages and the evolution of the Universe. Major theoretical predictions were either confirmed or ruled out by more and more precise observations, thanks to modern technological progresses. This allowed the construction of a parameterizable and measurable model, which during the years has been tightly constrained by accurate datasets, leading to the establishment of a Cosmological Concordance Model.

The entire picture is based upon four observational pillars: the Big Bang Nucleosynthesis (BBN), the large scale structure, the accelerated expansion of the Universe and Cosmic Microwave Background (CMB). Without committing ourselves in providing an extensive review of these topics, let us briefly introduce them.

The theory of BBN gives a detailed mathematical description of the production of the light elements (namely, deuterium, helium-3, helium-4, and lithium-7) in a well defined epoch a few minutes after the Big Bang. Specifically, the theory yields precise quantitative predictions for the mixture of these elements, that is, the primordial abundances (Alpher, Bethe & Gamow, 1948). The current level of agreement with observations, achieved by measurements of emission lines in the galaxies, is by no means trivial or guaranteed, and represents an impressive success of modern Cosmology: BBN extrapolates the contents and conditions of the present Universe back to times of about one second (Steigman, 2007).

The large scale structure is observed via redshift surveys of million of galaxies, as the Sloan Digital Sky Survey (SDSS[1]) and the 2dF Galaxy Redshift Survey[2]. These surveys allow to construct three dimensional maps of the Universe extending up to hundreds of Mpc and, under the assumption that the luminous matter traces the presence of dark matter, to constrain the spectrum of the fluctuations today, that can be fruitfully compared to the primordial power spectrum, inferred through the CMB. In the cold dark matter (CDM) scenario, where the latter is made by collisionless particles with negligible kinetic energy with respect to their mass, structure grows hierarchically, with small objects collapsing first and merging in a continuous hierarchy to form more and more massive objects. The CDM theory

[1] http://www.sdss.org/dr6/

[2] http://www2.aao.gov.au/ TDFgg/

makes no exact predictions about the nature of dark matter itself, which must be looked for in specific models of the early Universe, or directly in modern particle physics experiments.

In 1998, observations suggested that the expansion of the Universe is surprisingly accelerating (Perlmutter et al., 1999; Riess, 1998). These measurements used the Type Ia as standard candles, which are explosions of accreting carbon-oxygen white dwarves in binary systems reaching the Chandraseker limit of approximately 1.4 solar masses. This breakthrough left cosmologists with more open questions then answers, since the nature of this acceleration is totally obscure and unexpected. Indeed, the most obvious explanation for this effect, a Cosmological Constant coming from the energy density in vacuum, overestimates the measured value of 123 orders of magnitude. The Cosmological Model still says nothing about the fundamental physical origin of this "energy" that, according to the observations, accounts for more then $2/3$ of the all content of the Universe.

Last but not least, here it comes the main topic of this thesis:the CMB, the relic radiation that comes from the early Universe. In the last 10 years, outstanding technological improvements made it possible to measure this signal with high accuracy.

About 300.000 years after the Big Bang, as the Universe was cooling down, the protons and the electrons started to combine to generate atoms, and the photons were able for the first time to decouple and free stream toward us. Since this process was almost instantaneous, this cosmic radiation is a snapshot of the Universe at those early times. The Friedmann-Robertson-Walker (FRW) metric, i.e. the metric commonly used to describe the cosmological framework, naturally leads to the existence at each time of a special scale, called horizon, that defines the largest distance within causal contact is established, as a function of time. The epoch at which a given cosmological scale is equal to the horizon is therefore special, and referred to as the *horizon crossing*. At the moment of decoupling, such a distance subtended an angle that today is about one degree; thus, differences in the CMB coming from greater angles are able to trace those cosmological perturbations which were unaffected by causal processes, giving us directly a picture of primordial perturbations. On the other hand, regions smaller than the horizon underwent dynamical processes before decoupling, mainly resulting in oscillations due to the opposite effect of gravitational infall and photon pressure. These acoustic oscillations are imprinted in the CMB and bring to us most valuable information on the status and the composition of the early Universe.

Remarkably, the initial conditions and the subsequent dynamics give rise to anisotropies as low as 10^{-5} on all scales. This results in a non-trivial simplification of the treatment: metric fluctuations, as well as perturbations in the various components, may be treated linearly. In addition the CMB decoupling occurs at about T$\sim$ 3000K, corresponding to an energy scale where physical processes are well known, i.e. the electron-photon scattering may be analyzed in the Thomson limit.

Together with what we already achieved with the observations, theoretical predictions still have to be verified and tested, like the one which is the subject of

this thesis and that we introduce now. The Thomson scattering allows a fraction of the radiation to be linearly polarized; the CMB polarization field can be conveniently divided into two components: the gradient-like E modes and the curl-like B modes. According to the linear theory of cosmological perturbations, the latter can be activated by gravitational waves coming from the inflation, an epoch of exponential expansion that the Universe underwent just after the Big Bang, driven by a negative-pressure vacuum energy density (Guth, 1981). The inflation magnifies the original quantum fluctuations to cosmic size, that become the seeds for the growth of structure in the Universe. The gravitational waves produced by this process survived until the decoupling. Then a detection of the B modes would be of the most importance for Cosmology and for Physics in general: a strong prediction of the inflation would be confirmed and, even more importantly, the existence of the gravitational waves would be proved.

Hunting B modes is then one of the most exciting frontier in modern Cosmology. Given the tiny level of the signal, its detection will unavoidably require that two conditions are satisfied: the control of instrumental systematics and of the foreground radiation from our own Galaxy. The present work tries to address the second issue. The most advanced operating instrument, i.e. the Wilkinson Microwave Anisotropy Probe (WMAP[3]), a NASA satellite that has been observing the CMB for 5 years, demonstrated that in the microwave band the signals emitted from our own Galaxy may be comparable or higher than the expected CMB B mode power at all the angular scales and frequencies. In less than one year, another CMB satellite, the ESA mission Planck[4], doubling the covered frequency range and reaching a sensitivity about ten times better than WMAP, will fly; together with the main scientific outcome from the CMB observations, Planck will put stringent limit on the level of B modes from primordial gravitational waves, and provide crucial information on the level and properties of the foreground Galactic emission limiting our capability of detecting the signal. Moreover, the community of theorists and instrumentalists which worked in these years for these missions is conceiving a future CMB satellite aiming specifically at the detection of B modes. As part of this effort, a new interdisciplinary area of the CMB data analysis science is born in recent years, concerning the application of algorithms developed in signal processing in a CMB context, aiming at the separation of the background signal from the foregrounds on the basis of multi-frequency data. On an instrumental side, a relevant effort is being made for carrying out several sub-orbital experiments for testing the new technologies for the detection of the signal, as well as applying the new data analysis techniques mentioned above, targeting those regions of the sky that are known to be less contaminated by the Galaxy.

This thesis contains the first steps in this framework. The ultimate goal of this work, which extends well beyond the results exposed here, is the assessment of the minimum detectable level of B modes when the foregrounds are taken into account,

[3] wmap.gsfc.nasa.gov

[4] www.rssd.esa.int/Planck

applying the new data analysis techniques mentioned above to the incoming data.

Here is the outline of this thesis. In Chapter 1, I will introduce the current cosmological model and review the theoretical aspects of the CMB anisotropies and the present status of the observations. In Chapter 2, I will focus on the diffuse emissions of our Galaxy in the microwaves, expected to be a serious contamination for the CMB studies. The separation of these diffuse components and the CMB cleaning will be the main topic of Chapter 3, where I will describe a few algorithms I've worked on during my Phd . Finally, in Chapter 4, I will present a couple of applications of these methods aimed at B mode recovery.

Chapter 1

Theory of CMB anisotropies

1.1 Introduction

Our current knowledge of the beginning of the Universe is mainly based upon the successful theory of the Hot Big Bang. It states that our Universe began about 14 billion years ago as a hot, dense and very uniform sea of elementary particles, mutually interacting and at thermal equilibrium.

After the quasi exponential inflationary expansion, baryonic matter formed within the first second, while the nucleosynthesis lasted a few minutes as the Universe was expanding and cooling. The baryons were in this plasma status until about 300,000 years after the Big Bang, when the Universe reached a temperature sufficiently cool to permit protons to capture free electrons and form atomic hydrogen. This process, called recombination, lowered greatly the density of free electrons and dropped the photon-plasma cross section by 3 orders of magnitude, leading to the decoupling of matter and radiation, and making the Universe transparent to light.

The Cosmic Background Radiation (CBR) released during this era of decoupling has traveled almost unperturbed until today. We can observe the peak of this black body emission in the microwave region, corresponding to a temperature of 2.73 K, about 1000 times smaller than the one at the recombination. This radiation is then called Cosmic Microwave Background and brings us a snapshot of the Universe as it was 14 billion years ago. It was predicted by the Big Bang theory and first discovered in 1964 (Penzias and Wilson, 1965).

The CMB is until now the best tool in our possess to probe the early Universe. Theories of structure formation predict small inhomogeneities in the matter distribution at early times that later became galaxy and galaxy clusters trough gravitational collapse. These density inhomogeneities are imprinted in the CMB black body as temperature anisotropies, detected for the first time by the COsmic Background Explorer (COBE) satellite on angular scales larger than 7^o (Smoot et al., 1992; Mather et al., 1992).

The importance of the CMB in modern Physics is enormous, as it is clearly

confirmed by the the two Nobel prizes awarded in this research field to the works cited above.

The next promise that cosmologists hope to be fulfilled comes from the polarized part of the CMB radiation. The CMB polarization anisotropy, unlike the temperature, can discriminate between different kinds of perturbations, as we will discuss in this Chapter. It has then the potential to reveal us the signature of primordial gravity waves and could show us the "smoking gun" of the Inflation (Zaldarriaga, Seljak, 1997; Kamionkowski et al., 1997).

We will start this Chapter by introducing the cosmological framework in which the model is defined. Then, we will describe the CMB phenomenology and dynamics, taking advantage of the linear perturbations theory. This will naturally lead us to the definition of the CMB power spectra. We will conclude with a short review on the status of the CMB observations and planned experiments.

1.2 Cosmological framework

The starting point is the Cosmological Principle that states that on large spatial scales, the Universe is homogeneous and isotropic. When applied, this principle severely restricts the large variety of possible cosmological theories. The properties of homogeneity and isotropy assumed by the Cosmological Principle suggest that Earth is not at a preferred place in the space and that at very large scales the Universe is smooth (i.e. not fractal). A very strong support to this principle comes from the CMB itself, being isotropic to roughly one part in 100,000.

The casting of the following CMB anisotropy theory is based on the text book by Dodelson (2003) as well as the work by Hu and White (1997).

1.2.1 The FRW metric

Homogeneity and isotropy leave essentially only two degrees of freedom to the system. The first one is the scale factor $a(t)$, that fixes at each time the value of physical lengths and determines the cosmological redshift, defined as $z+1 = a_0/a$, where the subscript 0 marks the present. The second one, K, is related to the global space curvature, as an homogeneous metric can be more or less curved. Then, the fundamental length element is:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a(t)^2\left(\frac{1}{1-Kr^2}dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2\right), \quad (1.0)$$

where r, θ and ϕ are the usual spherical coordinates for radius, polar and azimuth angle, respectively, and $g_{\mu\nu}$ is the metric tensor. In this expression the scale factor a has been factored out of the spatial part. The spatial coordinates, r, θ and ϕ, are then at rest with respect to the cosmological expansion and are called comoving coordinates. Similarly we can introduce the conformal time τ:

$$d\tau = \frac{dt}{a(t)}\,. \quad (1.0)$$

Finally, the FRW metric may be easily written as:

$$g_{\mu\nu} \equiv a^2 \cdot \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & (1-Kr^2)^{-1} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2\theta \end{pmatrix} \equiv a^2 \gamma_{\mu\nu} \,, \tag{1.0}$$

so that the cosmic expansion can be completely factored out, defining the comoving metric $\gamma_{\mu\nu}$. We will indicate the conformal time derivative with a dot on top of the quantity, while those with respect to the ordinary time are indicated with the t subscript. It is also useful to define two different quantities describing the velocity of the expansion, i.e.

$$H = \frac{a_t}{a} \;\;,\;\; \mathcal{H} = \frac{\dot{a}}{a} \,, \tag{1.0}$$

named ordinary and conformal Hubble expansion rates, respectively.

As for the metric, also the stress energy tensor, that specifies the content of spacetime in terms of physical entities, is simplified by the assumptions of homogeneity and isotropy. The relevant quantities are just the energy density, ρ, and the pressure p.

$$T^{\nu}_{\mu} \equiv \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \,, \tag{1.0}$$

where the minus to the energy density is due to the choice of our signature. The stress energy tensor may also be written as

$$T_{\mu\nu} = (\rho + p) u_\mu u_\nu + p g_{\mu\nu} \,, \tag{1.0}$$

where u^μ represents the quadri-velocity of a fluid element with respect to a given observer:

$$u^\mu = \frac{dx^\mu}{d\tau} \,. \tag{1.0}$$

The quadri-velocities are normalized as $u_\mu u^\mu = -1$. In comoving coordinates, where the $u^a = 0$, this condition implies:

$$u^\mu \equiv \left(\frac{1}{a}, 0, 0, 0\right) \,. \tag{1.0}$$

The dynamics of the system is driven by:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \;,\; T^{;\nu}_{\mu\nu} = 0 \tag{1.0}$$

respectively the Einstein and the conservation equations. Thanks to the assumptions of the FRW metric, they reduce to two differential equations only, with the

conformal time τ as independent variable; they rule the dynamics of the expansion and the conservation of energy. They are the Friedmann equation:

$$\mathcal{H}^2 = \frac{8\pi G}{3} a^2 \rho - K \,, \tag{1.0}$$

and the simplified conservation equation:

$$\dot{\rho} + 3\mathcal{H}(\rho + p) = 0 \,. \tag{1.0}$$

A further assumption is needed to solve these coupled equations, i.e. a relation between pressure and energy density, $p(\rho)$. For interesting cases, pressure is proportional to the energy density:

$$p = w\rho \,, \tag{1.0}$$

where w is the equation of state of the fluid.

Starting from the Friedmann equation (1.2.1), it is straightforward to define the cosmological critical density

$$\rho_c = \frac{3H^2}{8\pi G} \,, \tag{1.0}$$

that is the density required to make the Universe spatially flat. The amount of each components that make up the Universe is usually defined in terms of their contribution to the critical density, $\Omega_s = \rho_s/\rho_c$, for each component s. We will see in the next Section how they are defined and what are the present day estimated values for each of them.

1.2.2 The Cosmological Concordance Model

The number of parameters needed for describing a given model varies a bit according to anyone's taste and on how simple the considered model is supposed to be. The Cosmological Concordance Model provides a minimum set of parameters sufficient to describe the status of the observations up to date. These are:

$$h \;,\; \Omega_m h^2 \;,\; \Omega_b h^2 \;,\; n_s \;,\; \tau \;,\; \sigma_8 \;, \tag{1.0}$$

which we describe and define below. In the concordance model there is the underlying assumption that the Universe is flat (actually in perfect accordance with CMB observations). Moreover, according to the inflation mechanism, there may be gravitational waves in the perturbation spectrum, and this is parameterized relatively to the density perturbation. Therefore, we add two more parameters:

$$\Omega_{tot} \;,\; r \;, \tag{1.0}$$

describing curvature and gravitational waves amplitude, respectively. Let's now take a quick look at these 8 parameters. We will provide their best fit values (with 2σ errors), using the WMAP data (see Dunkley et al., 2008, and references therein).

- h

As introduced before, the so called Hubble constant H_0 is the derivative of the Cosmological scale factor, measured today. Locally it provides the Hubble expansion law, $v = H_0 r$ and then is measured in units of $\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$. For practical convenience it is sometimes defined as $h = H_0/100$. The most recent estimation gives $h = 0.719^{0.026}_{0.027}$. Since the Hubble constant is a rate, its inverse defines a timescale, that is more or less the age of the Universe (that actually depends also on other parameters of the model).

- $\Omega_m h^2$

It is the contribution given by the total mass density. Before the supernovae observations (Riess, 1998; Perlmutter et al., 1999) this parameter was thought to be the largest among all the densities contributions. Now its best fit value is $\Omega_m h^2 = 0.1326 \pm 0.0063$.

- $\Omega_b h^2$

It defines the fraction of the matter that is accounted by baryons. In CMB observations, it influences the relative height of the acoustic peaks. Current value is $\Omega_b h^2 = 0.02273 \pm 0.00062$.

- Ω_{tot}

The overall curvature of the Universe depends on the sum of all the relative densities of all the components that make up it. It is defined as $\Omega_{tot} = \sum \Omega_s$. The CMB played a crucial role in constraining Ω_{tot}, since the position of the acoustic peaks in the CMB strongly depends on its value. It is now measured to be close to unity, $\Omega_{tot} = 1.02 \pm 0.02$.

- n_s

Slope of the scalar power spectrum of the initial perturbations, defined as the variance of the fluctuation density in the Fourier spaces assuming the latter to be Gaussian distributed, $P(k) \propto k^{n_s}$. Its current best fit is $n_s = 0.963 \pm 0.014$

- τ

We know that the Universe reionized at some redshift higher than 6, probably due to the earliest stars. The understanding of that period, known as the end of the Dark Age, is an open issue in Cosmology and the CMB can help in the task, since the extra Thomson scattering, that occurred at that time, partially erased the anisotropies. The parameter used to describe this effect is the optical depth, whose current value is $\tau = 0.087 \pm 0.017$.

- σ_8

It is the rms mass fluctuation amplitude in spheres of size $8h^{-1}$ Mpc and measures the normalization of the matter power spectrum. The current estimation is $\sigma_8 = 0.796 \pm 0.036$.

- r

It is the ratio between gravitational waves and density perturbation amplitudes, expressed by means of their induced power on the CMB anisotropies at large angular scales, as it will be discussed in the next sections.

1.2.3 Linear perturbations

The most important assumption that, for most of their evolution, perturbations were in linear regime comes from the CMB itself, since the typical temperature anisotropies are of the order of 10^{-5}, with respect to the average temperature, over the whole sky and up to very small scales. Then, the linear approximation should be valid to describe cosmological perturbations on a large interval of time and physical scales, before and after the CMB origin, breaking down only recently and on scales smaller than those of galaxy clusters. In this Section we will briefly describe the cosmological perturbation theory, building the setup for understanding the physics of the CMB anisotropies, that will be treated later on in this Chapter.

Consider now the definition of the perturbed metric tensor:

$$\bar{g}_{\mu\nu} = g_{\mu\nu} + \delta g_{\mu\nu} = a^2(\gamma_{\mu\nu} + h_{\mu\nu}) \,, \tag{1.0}$$

where the bar means background plus perturbations. Since $\gamma_{\mu\nu}$ contains coefficients of order 1 (in the flat FRW it coincides with the Minkowski metric), linearity is satisfied if:

$$h_{\mu\nu} \ll 1, \quad \forall \ \mu, \nu. \tag{1.0}$$

At the same time, the stress energy tensor is perturbed as

$$\bar{T}^{\nu}_{\mu} = T^{\nu}_{\mu} + \delta T^{\nu}_{\mu} \,, \tag{1.0}$$

and in this case linearity may be expressed in terms of the non-zero quantities in the background tensor, $T_{\mu\nu}$:

$$\delta T^{\nu}_{\mu} \ll \rho, p \,, \quad \forall \ \mu, \nu. \tag{1.0}$$

Most importantly, a direct consequence of the linearity is that perturbations do not affect the background dynamics. Therefore, any equation in cosmology splits in two and (1.2.1) become:

$$\bar{G}_{\mu\nu} = 8\pi G \bar{T}_{\mu\nu} \equiv \begin{cases} G_{\mu\nu} = 8\pi G T_{\mu\nu} \\ \delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu} \end{cases} , \tag{1.0}$$

$$\bar{T}^{;\nu}_{\mu\nu} = 0 \equiv \begin{cases} T^{;\nu}_{\mu\nu} = 0 \\ \delta(T^{;\nu}_{\mu\nu}) = 0 \end{cases} . \tag{1.0}$$

1.2.4 Decomposition and Fourier expansion

Perturbations can be classified according to their properties under spatial rotations (Dodelson, 2003). Since we are dealing with a tensor theory, from a very general point of view a function of the space may behave as a scalar, as a vector or as a tensor. Now, while the physical nature of scalars is unique, each vector or tensor function may be divided in full generality in different parts. Indeed vectors can be divided in a scalar-type component, i.e. the part that is the derivative of a scalar, and a vector-type divergenceless component, describing vortex-like perturbations. Similarly, tensors can be decomposed into a scalar-type component (responsible for the trace), a vector-type component and a tensor-type component that satisfies the transverse and traceless conditions. The latter correspond to tensor fluctuation modes in general relativity, i.e. gravitational waves.

Moreover these perturbations are conveniently defined in the Fourier space, taking advantage of the fact that in linear theory any mode evolves independently in that space. A gauge freedom allows us to choose in a family of frames that differ among each other for coordinate change as small as the perturbations themselves.

In this work we adopt the Newtonian gauge, where the only relevant elements are Ψ and Φ, parameterizing the Fourier amplitudes of the generalized scalar Newtonian potentials, associated respectively to δg_{00}, δg_{ii}, and the element H_T, associated to the gravitational waves expressed by the tensor-type component of δg_{ij}. Since their evolution is suppressed by the expansion, we do not consider here vector-type perturbations.

Similarly, the perturbations in the cosmological components are associated with the fluctuations in the corresponding species of the stress energy tensor. For a given species x, density fluctuations $\delta\rho_x/\rho_x$ correspond to $(\delta T^0_0)_x$, peculiar velocities $\vec{v}_x$ to the component $(\delta T^i_0)_x$, isotropic pressure perturbations to the diagonal terms in $(\delta T^i_i)_x$, while viscosity, or shear, to the off diagonal ones.

1.3 The Cosmic Microwave Background

1.3.1 Main phenomenology

The CMB is made by photons that carry information of an epoch in which they were interacting electromagnetically and gravitationally with the other species. During that epoch, the Thomson cross section σ_T between electrons and photons was large enough to keep the latter in thermal equilibrium with the rest of the fluid (tight coupling regime). As the Universe expanded enough and the temperature dropped sufficiently, the electrons started to recombine with nucleons, forming hydrogen and helium atoms and causing the cross section between electrons and photons to rapidly decay. The photons stayed in thermal equilibrium until the mean free path, $\lambda_\gamma = 1/(n_e\sigma_T)$, got larger than the horizon:

$$\lambda_\gamma > H^{-1} \, . \tag{1.0}$$

This epoch is called *decoupling*. The physical quantities in the relation above conspire to set this epoch about $300,000$ years after the Big Bang, corresponding to a redshift of about $1 + z \simeq 1100$. To be noted that the decoupling process was not instantaneous, and took an interval of about $\Delta z \simeq 100$.

Soon after that time, CMB photons may be thought to be free streaming in any direction. Therefore, they represent a sort of snapshot of their distribution at last scattering. They bring to us a picture of the anisotropies, as they were in the Universe at the last scattering surface (LSS).

We may expect to see two classes of records in this image. First, the intensity coming from a given direction should be recording the energy of photons and the metric perturbations at the LSS. Second, since the decoupling was not instantaneous, photons in general had time to generate an anisotropic angular distribution around the last scattering electrons, that were then able to linearly polarize the light through the Thomson scattering, bringing to us extra information in terms of cosmological perturbations. We describe these processes more in detail in the following.

1.3.2 Linearly perturbed black body

Photons are described as a fluctuating perfect relativistic fluid at thermal equilibrium. At the temperature T, the number density dn_γ of photons with energy E and momentum $\vec{p}$ is that of a black body obeying the Bose-Einstein distribution F with $g_\gamma = 2$ degrees of freedom:

$$dn_\gamma = F(E,T)d^3p = \frac{g_\gamma}{\exp(E/k_BT) - 1} d^3p \, . \tag{1.0}$$

The perturbed distribution of the black body can be obtained simply by replacing E with $E + \delta E$:

$$\frac{g_\gamma}{\exp\left(\frac{E+\delta E}{k_BT}\right) - 1} = \frac{g_\gamma}{\exp\left[\frac{E}{k_BT}(1-\Theta)\right] - 1} \, , \tag{1.0}$$

where $\Theta = -\delta E/E$ may be interpreted as the first order correction to the temperature T which would lead to an equivalent perturbation to the black body spectrum if E was unperturbed. As we will see through this Chapter, Θ is the key quantity to characterize the CMB anisotropies. At first order, one can write:

$$\begin{aligned} \delta F &= -\Theta \frac{E}{k_BT} \frac{d}{d(E/k_BT)} \left[\frac{g_\gamma}{\exp(E/k_BT) - 1} \right] = \\ &= \Theta \frac{E}{k_BT} \frac{g_\gamma \exp(E/k_BT)}{[\exp(E/k_BT) - 1]^2} \, . \end{aligned} \tag{1.0}$$

Θ may depend on a generic spacetime point $(\tau, \vec{x})$, as well as on the photon momenta p^μ. Defining $\hat{n}$ as $\vec{p} = E\hat{n}$, we can write $\Theta = \Theta(\tau, \vec{x}, \hat{n}, E)$. On the other hand, we do not allow Θ to depend on E, as that would represent a distortion of the

black body spectrum and the breaking of thermal equilibrium. Within this assumption, the perturbed spectrum is still black blackbody for any spacetime position and photon propagation direction, $\Theta = \Theta(\tau, \vec{x}, \hat{n})$.

1.3.3 Polarization

An homogeneous and isotropic black body distribution is unpolarized by definition, but if the black body temperature varies with position and photon propagation direction, this is not true anymore. Indeed, if Θ is measured on a given direction $\hat{n}$, one may project the intensity onto two perpendicular axes orthogonal to $\hat{n}$ forming the *polarization plane*. We can define:

$$Q(\tau, \vec{x}, \hat{n}) = \Theta_{\parallel} - \Theta_{\vdash} \ , \ U(\tau, \vec{x}) = \Theta_{\nparallel} - \Theta_{\nvdash} \ , \tag{1.0}$$

which are the Stokes parameters describing linear polarization, where the symbols $\nparallel$ and $\nvdash$ represent axes rotated by 45 degrees with respect to the ones defining Q. The last Stokes parameter, V, would be needed to describe circularly polarized radiation, but since the Thomson scattering produces linear polarization only, we will not consider it anymore. The polarization behaves as a rank 2 tensor.

It is possible to show that Q and U represent the amplitude of the decomposition of the polarization tensor into the Pauli matrices σ_1 and σ_3:

$$\mathbf{P} = Q\sigma_3 + U\sigma_1 = \begin{pmatrix} Q & U \\ U & -Q \end{pmatrix} . \tag{1.0}$$

Under a counterclockwise rotation of axes around $\hat{n}$ through an angle ψ, Q and U exchange their roles; on the other hand, it is possible to show that the quantity $Q \pm iU$ transforms as

$$Q \pm iU \to e^{\mp 2i\psi}(Q \pm iU) \ , \tag{1.0}$$

making manifest the spin 2 nature of the polarization field. Defining the matrices

$$\mathbf{M}_{\pm} = \frac{1}{2}\left(\sigma_3 \mp i\sigma_1\right) \ , \tag{1.0}$$

one also obtains

$$\mathbf{P} = (Q + iU)\mathbf{M}_{+} + (Q - iU)\mathbf{M}_{-} \ . \tag{1.0}$$

The polarization tensor is often described through its amplitude $\sqrt{Q^2 + U^2}$ and a direction defined as:

$$2\phi = \arctan\frac{U}{Q} \ . \tag{1.0}$$

If Q and U rotate of a given angle, ϕ remains unchanged, representing the *polarization direction* on the polarization plane.

1.3.4 Harmonic expansion

The dependence of Θ and $Q \pm iU$ on the arguments $\vec{x}$ and $\hat{n}$ makes these quantities suitably expandable in Fourier eigenfunctions and spherical harmonics. Concerning the latter, in this Section we perform two different kind of angular expansions, made in two different frames.

The first case is the natural expansion, made in the laboratory frame, corresponding to the θ and ϕ directions defined in the previous Section. The second choice is less intuitive, but extremely useful to deal with the CMB dynamics, as we will see later. This is the $\hat{k}-$frame expansion, different for each Fourier wavevector and having the direction of $\vec{k}$ coincident with the polar axis. Let us see these two cases separately.

Lab-frame

Spacetime dependence and harmonic expansion are completely independent. Therefore, for any given position x, the angular expansion of Θ and $Q \pm iU$ is

$$\Theta(\tau, \vec{x}, \hat{n}) = \sum_{lm} \Theta_{lm}(\tau, \vec{x}) Y_{lm}(\hat{n}) \,,$$

$$(Q \pm iU)(\tau, \vec{x}, \hat{n})\mathbf{M}_{\pm} = \sum_{lm} (Q \pm iU)_{lm}(\tau, \vec{x})_{\pm 2} Y_{lm}(\hat{n})\mathbf{M}_{\pm} \,, \tag{1.0}$$

where we notice that $Q \pm iU$ has been expanded in *tensor* spherical harmonics, indicated as ${}_{\pm 2}Y_{lm}(\hat{n})$, being the components of a rank 2 tensor. This is the harmonic expansion in the lab-frame, defined by the θ and ϕ directions.

The expansion coefficients in (1.3.4) fully describe the CMB observables; however, for polarization, two more quantities can be defined. The two combinations:

$$Y^E_{lm}\mathbf{M}_E = \frac{1}{2}\left({}_2Y_{lm}\mathbf{M}_+ + {}_2Y_{lm}\mathbf{M}_-\right) \,,$$

$$Y^B_{lm}\mathbf{M}_B = \frac{1}{2}\left({}_2Y_{lm}\mathbf{M}_+ - {}_2Y_{lm}\mathbf{M}_-\right) \tag{1.0}$$

turn out to be extremely useful because of their special connection with the cosmological perturbations. The E and B notation has been chosen because of the parity relations between these two quantities in analogy with the electric and the magnetic fields. It can be showed that they represent the gradient and the curl of a scalar and vector potential, respectively. In Fig.1.1 we sketch typical E and B patterns, generated by different contributions of Q and U. It is easy to see that, for parity inversion $\hat{n} = -\hat{n}$, where $\hat{n}$ is the direction perpendicular to the figure, only B modes are changed.

Using this new definitions, the polarization tensor can be written:

$$\mathbf{P} = \sum_{lm} \left(E_{lm}Y^E_{lm}\mathbf{M}_E + B_{lm}Y^B_{lm}\mathbf{M}_B\right) \,, \tag{1.0}$$

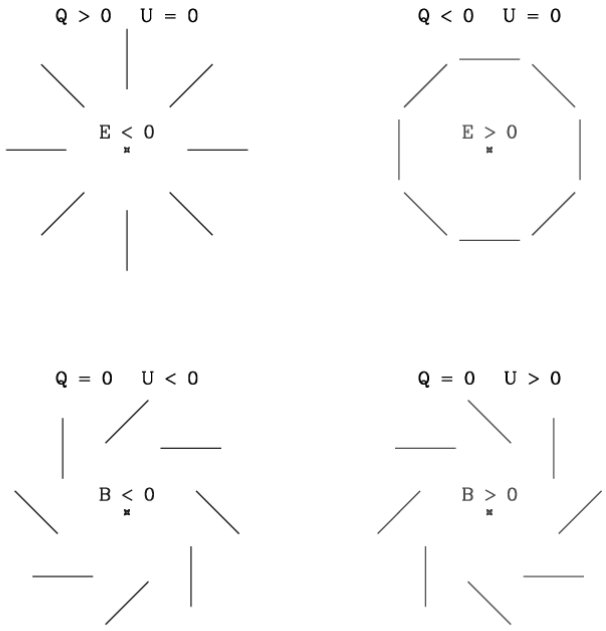

Figure 1.1: Typical patterns for E and B modes. From Zaldarriaga, Seljak (1997).

where the new coefficient are:

$$E_{lm} = \frac{1}{2}\left[(Q+iU)_{lm} + (Q-iU)_{lm}\right] ,$$

$$B_{lm} = \frac{1}{2i}\left[(Q+iU)_{lm} - (Q-iU)_{lm}\right] . \qquad (1.0)$$

The great advantage on using these quantities, as we will see later, is that the E modes are excited by all kinds of cosmological perturbations, while the B modes are activated only by the non-scalar ones.

It is now straightforward to introduce the angular power spectra of CMB anisotropies. At any spacetime location x, they are defined as:

$$\tilde{C}_l^{\Theta\Theta}(\tau,\vec{x}) = \frac{1}{2l+1}\sum_m |\Theta_{lm}(\tau,\vec{x})|^2 ,$$

$$\tilde{C}_l^{\Theta E}(\tau,\vec{x}) = \frac{1}{2l+1}\sum_m \Theta_{lm}(\tau,\vec{x})E^*_{lm}(\tau,\vec{x}) ,$$

$$\tilde{C}_l^{\Theta B}(\tau,\vec{x}) = \frac{1}{2l+1}\sum_m \Theta_{lm}(\tau,\vec{x})B^*_{lm}(\tau,\vec{x}) ,$$

$$\tilde{C}_l^{EE} = \frac{1}{2l+1}\sum_m |E_{lm}(\tau,\vec{x})|^2 ,$$

$$\tilde{C}_l^{BB} = \frac{1}{2l+1}\sum_m |B_{lm}(\tau,\vec{x})|^2 \ ,$$

$$\tilde{C}_l^{EB} = \frac{1}{2l+1}\sum_m E_{lm}(\tau,\vec{x})B^*_{lm}(\tau,\vec{x}) \ . \qquad (1.0)$$

Although this represents a lossy compression of the CMB observables, (the phases of the a_{lm} coefficients are definitely lost), it is still convenient from the point of view of a statistical description. Moreover, if the a^{Θ}_{lm} coefficients are Gaussian variables with known variance, no physical information is coded into the phases of CMB anisotropies, so the compression (1.3.4) is lossless:

$$< \Theta_{lm}\Theta_{l'm'} >= C_l^{\Theta\Theta}\delta_{ll'}\delta_{mm'} \ , < \Theta_{lm}E_{l'm'} >= C_l^{\Theta E}\delta_{ll'}\delta_{mm'} \ ,$$

$$< E_{lm}E^*_{l'm'} >= C_l^{EE}\delta_{ll'}\delta_{mm'} \ , < B_{lm}B_{l'm'} >= C_l^{BB}\delta_{ll'}\delta_{mm'} \ ,$$

$$< \Theta_{lm}B_{l'm'} >= 0 \ , \ < E_{lm}B^*_{l'm'} >= 0 \ . \qquad (1.0)$$

The vanishing spectra above are due to the difference in parity relations. It is important to distinguish between the angular power spectra in (1.3.4), defined on a given realization of the coefficients, and those in (1.3.4), which are supposed to describe the averaged variance of the anisotropy power over the whole Gaussian statistics.

$\hat{k}-$**frame**

We now want to perform the harmonic expansion of the temperature fluctuations $\Theta(\tau,\vec{x},\hat{n})$ in the $\hat{k}-$frame, where the polar direction is parallel to $\vec{k}$. This expansion has to be done after we moved into the Fourier space. Indeed, the harmonic modes for Θ are given by:

$$G_{lm}(\vec{x},\vec{k},\hat{n}) = (-i)^l Y(\vec{k},\vec{x})\sqrt{\frac{4\pi}{2l+1}}Y_{lm}(\hat{n}_{\hat{k}}) \ , \qquad (1.0)$$

where $Y_{lm}(\hat{n}_{\hat{k}})$ are the usual spherical harmonics, the subscript indicates that $\hat{n}$ is expressed in the $\hat{k}-$frame, and the constants in front are purely conventional; $Y(\vec{k},\vec{x})$ represents the Fourier expansion mode, corresponding to plane waves for flat Universes.The harmonic coefficients of Θ along the basis (1.3.4) are given by:

$$\Theta_{lm}(\tau,\vec{k}) = \int d^3x \int d\Omega^{\hat{k}}_{\hat{n}}\Theta(\tau,\vec{x},\hat{n})G_{lm}(\vec{k},\vec{x},\hat{n}) \ , \qquad (1.0)$$

and the full expansion of Θ is

$$\Theta(\tau,\vec{x},\hat{n}) = \int \sum_{lm}\Theta_{lm}(\tau,\vec{k})G_{lm}(\vec{x},\vec{k},\hat{n})d^3k \ . \qquad (1.0)$$

For polarization, the expansion is analogous and proceeds conceptually as outlined in the previous Section for the E and B modes .The polarization tensor in the $\hat{k}-$frame is given by:

$$\mathbf{P}(\tau,\vec{x},\hat{n}) = (Q+iU)(\tau,\vec{x},\hat{n})\mathbf{M}_{+} + (Q-iU)(\tau,\vec{x},\hat{n})\mathbf{M}_{-} =$$

$$= \int d^3k \sum_{lm}\left[E_{lm}(\tau,\vec{k})G^{E}_{lm}(\hat{n}_{\hat{k}})\mathbf{M}^{\vec{k}}_{E} + B_{lm}(\tau,\vec{k})G^{B}_{lm}(\hat{n}_{\hat{k}})\mathbf{M}^{\vec{k}}_{B}\right], \qquad (1.0)$$

where all arguments have been made explicit.

The use of the $\hat{k}-$frame is convenient because, as we will see, cosmological perturbations only activate a few modes in it. At the end of the treatment we will make an explicit connection between the $\hat{k}-$frame and the more intuitive lab-frame.

1.4 CMB dynamics

The CMB distribution is affected by two processes, gravity and Thomson scattering. Therefore, by propagating the total time derivative in all arguments, the Boltzmann equation for the CMB may be written as:

$$\frac{d\bar{F}}{d\tau} = \frac{\partial\bar{F}}{\partial\tau} + \frac{\partial\bar{F}}{\partial x^i}\frac{dx^i}{d\tau} + \frac{\partial\bar{F}}{\partial E}\frac{d\bar{E}}{d\tau} + \frac{\partial\bar{F}}{\partial n^i}\frac{dn^i}{d\tau} = \bar{G} + \bar{C}\,, \qquad (1.0)$$

where the two terms on the right hand side represent the gravitational and Thomson scattering, respectively. The first simplification concerns the first term: since the distribution does not depend explicitely on time, $\partial\bar{F}/\partial\tau$ is null.

At the perturbation level, Eq.(1.4) determines the evolution of $\Theta(\tau,\vec{x},\hat{n})$. In a flat FRW background the last term in (1.4) contributes only to second order and we neglect it. Moreover, the term $dx^i/d\tau$ may be directly related to the photon propagation direction, n^i, $dx^i/d\tau = (dx^i/d\eta)\cdot(d\eta/d\tau) = p^i/p^0 = n^i$. Therefore, exploiting the perturbed expression (1.3.2) the perturbed Boltzmann equation may be written as:

$$\dot{\Theta} + \hat{n}\cdot\vec{\nabla}\Theta = \delta G + \delta C^{\Theta}\,, \qquad (1.0)$$

where the gravitational and Thomson scattering terms will now be defined.

Gravitational scattering

The time derivative of $\bar{E}$ in (1.4) is determined by the the geodesics propagation of photons in a linearly perturbed FRW metric. As we have introduced before, the latter may be written as

$$\bar{g}_{\mu\nu} = a^2(\gamma_{\mu\nu} + h_{\mu\nu})\,, \qquad (1.0)$$

where $h_{\mu\nu}$ represents the perturbation.

The complete computation of the gravitational term can be found in Sachs & Wolfe (1967). We report here the final result for arbitrary $h_{\mu\nu}$, useful for the following discussion:

$$-\frac{d\Theta}{d\tau} = \frac{d(\delta E/E)}{d\tau} = -\frac{1}{2}h_{ij,0}n^i n^j + \frac{1}{2}h_{00,i}n^i - h_{0i,0}n^i \ . \tag{1.0}$$

These relations determine the δG term in (1.4).

Thomson scattering

The CMB temperature today is $T \simeq 3K$ which corresponds to ~ 1 eV at decoupling. The most relevant interaction for photons at these energies is Compton scattering onto non-relativistic electrons, indicated by the following relation:

$$e^-(q^\mu) + \gamma(p^\mu) \rightarrow e^-(q^{\mu\prime}) + \gamma(p^{\mu\prime}) \ , \tag{1.0}$$

where q^μ, p^μ, $q^{\mu\prime}$, $p^{\mu\prime}$ are the electron and photon incoming and outcoming quadrimomenta, respectively. We are interested in the change in the photon distribution $\bar{F}$ from Compton scattering.

We don't write explicitely the computation here and just point the reader to Hu and White (1997). The final result for the collision term from Compton scattering is:

$$\delta C^\Theta = a n_e \sigma_T E \frac{\partial d}{\partial E}(\Theta - \Theta_0 - \hat{n}\cdot\vec{v}_e) \ , \tag{1.0}$$

where $\vec{v}$ represents the peculiar velocity of electrons, and we defined the differential optical depth:

$$\dot{\vartheta} = a n_e \sigma_T \ . \tag{1.0}$$

Eq.(1.4) represents the Thomson scattering term in (1.4). Analogous terms, which we do not write here explicitely, drive the evolution of the $Q \pm iU$ terms, characterizing the linear polarization of CMB photons.

1.4.1 The Boltzmann equation in the harmonic domain

We now expand the perturbed Boltzmann equation in spherical harmonics in the $\hat{k}-$frame. This will lead us to important conclusions.

The first thing to notice is that the gradient term $\hat{n}\cdot\vec{\nabla}$ becomes a dipole term given by

$$i\hat{n}\cdot\vec{k} = i\sqrt{\frac{4\pi}{3}}kY_1^0 \ . \tag{1.0}$$

The $l = 1$, $m = 0$ spherical harmonic function above multiplies the coefficients of the $\hat{k}-$frame expansion for Θ and $\mathbf{P}$ in the expansions (1.3.4) and (1.3.4), coupling different angular terms between each other. This is actually the term which describes the *free streaming*, i.e. the projections of local anisotropies on larger

and larger angular scales as time goes. In the harmonic expression for the Boltzmann equation this means that the power is transferred dynamically from l to $l+1$ (this can be seen easily applying the Clebsh-Gordan decomposition to the product of (1.4.1) with the spherical harmonics in Θ), i.e. during the free streaming the anisotropy power at higher and higher multipoles is populated. As we will see later, the anisotropy pattern we see today in the sky on all angular scales is essentially given by the monopole and dipole of the radiation field at decoupling. Similarly, one can see that, in the $\hat{k}-$frame, the scalar metric perturbations activate $m = 0$ terms only in Θ_l^m and E_l^m, while tensor perturbations excite in Θ_l^m, E_l^m and B_l^m the $m = \pm 2$ modes only

The expansion of the Boltzmann equation proceeds straightforwardly with spherical harmonic algebra, becoming an infinite set of evolution equations for the harmonic multipoles in Θ and $Q \pm iU$. Its expression for Θ_l^m is:

$$\dot{\Theta}_l^m + \dot{\vartheta}\Theta_l^m - k\left(\frac{{}_0\kappa_l^m}{2l-1}\Theta_{l-1}^m - \frac{{}_0\kappa_{l+1}^m}{2l+3}\Theta_{l+1}^m\right) = S_l^m \,, \tag{1.0}$$

where ${}_s\kappa_{l+1}^m = \sqrt{(l^2-m^2)(l^2-s^2)/l^2}$. In this relation, the second term on the left hand side represents the friction due to Thomson scattering, causing a damping of anisotropies because of the scattering itself, while the third one is the free streaming; the source term S_l^m collects all the contribution from the metric perturbations and the remaining ones from Thomson scattering. Let us write them explicitely:

$$S_0^0 = \dot{\vartheta}\Theta_0^0 - \dot{\Phi} \,,\; S_1^0 = \dot{\vartheta}v_b^0 + k\Psi \,,\; S_2^0 = \dot{\vartheta}P^0 \,, \tag{1.0}$$

$$S_1^{\pm 1} = \dot{\vartheta}v_b^{\pm 1},\; S_2^{\pm 1} = \dot{\vartheta}P^{\pm 1} \tag{1.0}$$

$$S_2^{\pm 2} = \dot{\vartheta}P^{\pm 2} - \dot{H}_T^{\pm 2} \,, \tag{1.0}$$

where:

$$P^m = \Theta_2^m\frac{1}{10} - \frac{\sqrt{6}}{10}E_2^m \,. \tag{1.0}$$

An analogous computation gives the evolution equations for the E and B modes:

$$\dot{E}_l^m + \dot{\vartheta}E_l^m - k\left(\frac{{}_2\kappa_l^m}{2l-1}E_{l-1}^m - \frac{2m}{l(l+1)}B_l^m - \frac{{}_2\kappa_{l+1}^m}{2l+3}E_{l+1}^m\right) =$$

$$= -\dot{\vartheta}\sqrt{6}P^m\delta_{l2} \,, \tag{1.0}$$

$$\dot{B}_l^m + \dot{\vartheta}B_l^m - k\left(\frac{{}_2\kappa_l^m}{2l-1}B_{l-1}^m + \frac{2m}{l(l+1)}E_l^m - \frac{{}_2\kappa_{l+1}^m}{2l+3}B_{l+1}^m\right) = 0 \,. \tag{1.0}$$

Several interesting features may be noticed here. First, it is evident how the B modes may exist only for non-scalar perturbations: the anisotropy power activates B only if $m \neq 0$ in Eq.(1.4.1), since the right hand side is null. This is not true for E modes. Second, the free streaming mixes E and B dynamically for each given pair of l and m values, as it is represented by the presence of B_l^m in (1.4.1) and

E_l^m in (1.4.1). Third, the power transfer from total intensity to polarization and back is represented by the P^m term defined in (1.4.1) and shows that the polarization perturbations are activated by the quadrupole term. As we shall see later, the anisotropy power is sourced by the metric perturbations as well as the monopole and dipole in the Θ function; by free streaming, the power is transferred to the quadrupole and the higher moments, activating the polarization and projecting the whole anisotropy spectrum at higher and higher multipoles.

1.4.2 Acoustic oscillations and large scale power

The integral solution to the Boltzmann equation is easily achieved in the $\hat{k}-$frame (Hu and White, 1997):

$$\Theta_{lm}(\tau_0, \vec{k}) = (2l+1) \int_0^{\tau_0} d\tau e^{\mathcal{T}-\mathcal{T}_0} \sum_{l'} S_{l'm}(\tau, \vec{k}) j_l^{l'm}[k(\tau_0 - \tau)] \,, \qquad (1.0)$$

where

$$\mathcal{T}_0 - \mathcal{T} = \int_\tau^{\tau_0} \dot{\vartheta} d\tau' \,, \qquad (1.0)$$

represents the optical depth for Compton scattering between the epoch corresponding to τ and the present. Similar solutions are found for E and B modes.

This solution fully accounts for the relevant CMB anisotropy phenomenology. The integral over the conformal time may be seen as a line of sight weighting of the different contributions. More in detail, let us consider scalar-type perturbations. Plugging into Eq.(1.4.2) the expressions for the sources of the perturbations in Eq.(1.4.1), one obtains:

$$\frac{\Theta_{l0}(\tau_0, \vec{k})}{2l+1} = \int_0^{\tau_0} d\tau e^{\mathcal{T}-\mathcal{T}_0} \{[\dot{\vartheta}\Theta_{00}(\tau, \vec{k}) + \dot{\vartheta}\Psi(\tau, \vec{k}) + \dot{\Psi}(\tau, \vec{k}) - \dot{\Phi}(\tau, \vec{k})] \cdot$$

$$\cdot j_l[k(\tau_0 - \tau)] + \dot{\vartheta} v_b^0(\tau, \vec{k}) j_l^{10}[k(\tau_0 - \tau)] + \dot{\vartheta} P^0(\tau, \vec{k}) j_l^{20}[k(\tau_0 - \tau)]\} \,. \qquad (1.0)$$

It can be shown that the terms proportional to the combination $\dot{\vartheta} e^{\mathcal{T}-\mathcal{T}_0}$ selects contributions at the last scattering only, since that is the product of $\dot{\vartheta}$, a sort of step function with is non-zero at decoupling and earlier, with $e^{\mathcal{T}-\mathcal{T}_0}$, that is one up to decoupling and then exponentially suppressed. The term j_l^{20} is a combination of the Bessel functions and their derivatives (Hu and White, 1997), that are sharply peaked at multipoles, $l \simeq k(\tau_0 - \tau)$. The product between j_l^{20} and $\dot{\vartheta} e^{\mathcal{T}-\mathcal{T}_0}$ creates a one to one link between l and k, $l \simeq k(\tau_0 - \tau_{dec})$. This relates the particular scale k to the angle at which it is probed, that simply corresponds to the one subtended at last scattering:

$$\theta \simeq \frac{\pi}{l} \simeq \frac{k^{-1}}{\tau_0 - \tau_{dec}} \,. \qquad (1.0)$$

For scalar-type fluctuations, the effects which are picked up at decoupling are four. The first one is the monopole of CMB anisotropies, Θ_{00}, that represents the

contribution from the intrinsic photon density fluctuations at last scattering. The second one is the gravitational potential Ψ, known as Sachs-Wolfe effect, representing the contribution from the metric fluctuations at last scattering, simply due to the redshift or blueshift which photons undergo coming out of a gravitational potential well or hill, respectively. The third one is a Doppler shift, due to the electron peculiar velocity v_b^0. Finally the fourth one is a combination of the total intensity and polarization quadrupoles, $P^0 = (\Theta_{20} - \sqrt{6}E_{20})/10$, related to the anisotropic character of the Thomson scattering.

A fifth term is represented by the time derivatives of the gravitational potentials, $\dot{\Psi} - \dot{\Phi}$. Along the line of sight, this is multiplied by the quantity $e^{\mathcal{T}-\mathcal{T}_0}$, which is essentially equal to 1 from decoupling to the present. This means that $\dot{\Psi} - \dot{\Phi}$ is integrated along the whole line of sight of CMB photons and takes into account the variation of the gravitational fields experienced by the photon traveling to us. It is known as Integrated Sachs-Wolfe (ISW) effect.

As we stated above, integral solution exists also for polarization. It may be shown that the latter contains only the first kind of contributions, i.e. the one which comes from last scattering, as polarization is generated by the local quadrupole in Θ at decoupling.

Tight coupling and acoustic oscillations

Most of the phenomenology of the CMB dynamics in Θ occurs because of the competition between gravitational infall and photon pressure. This can be easily seen in the limit of infinitely strong coupling between photons and baryons (tight coupling). In this regime, the hierarchical equation system (1.4.1) reduces to two equations for Θ_0^0 and Θ_1^0, which correspond to the baryon-photon fluid density and peculiar velocity perturbations, respectively (Hu and White, 1997). The equation for Θ_0^0 is simply:

$$\ddot{\Theta}_0^0 + \frac{k^2}{3}\Theta_0^0 = -\frac{k^2}{3}\Psi \, . \tag{1.0}$$

It can be shown that Ψ is slowly varying when sub-horizon angular scales are considered. Therefore, the solution is made by a superposition of cosine and sine modes, with offset given by Ψ. This setup corresponds precisely to acoustic oscillations activated by gravity. Although in this simple example cosmological quantities remain rather hidden, these oscillations markedly depend on the various abundances, density and velocity perturbations, as well as the kind of primordial fluctuations in Ψ, giving us important insights on the physical conditions at those times, as we will see later.

1.5 The CMB anisotropy angular power spectrum

In this Section, after introducing the primordial power spectrum of the perturbations, we will focus on the CMB power spectra, expliciting the link between the

$\hat{k}-$frame and the *lab*−frame. We will then give a phenomenological description of the different features that the power spectra show, both in total intensity and polarization.

1.5.1 Primordial perturbation spectra

As we have seen the Boltzmann equation, Eq.(1.4.1), drives the evolution of the photons fluid. The perturbations are activated by the source term on the right hand side. They depend on the gravity terms Ψ, Φ and H_T, defined in Sec.1.2.4, the latter term being the amplitude of the gravitational waves.

The inflation excites all the metric components, but, while the vector contributions are washed away by the expansion, scalar and tensor components keep being relevant. Historically, cosmological perturbations have been classified according to two distinct families of initial conditions. In the iso-curvature scheme, the super-horizon metric perturbations are arranged in order to nullify the linear perturbations to the curvature. On the other hand the adiabatic (or iso-entropic) perturbations do possess large scale curvature fluctuations and the different species are arranged together as dictated by thermal equilibrium. In the inflationary paradigm, where perturbations were born inside the horizon, thermalized, and stretched outside by the expansion, a negligible or null contribution from iso-curvature perturbations is predicted. Therefore, only adiabatic perturbation will be considered in this work.

In addition, assuming Gaussianity, the density perturbations are characterized by the following relation:

$$< \delta_k \delta_{k'} >= P(k)\delta(\vec{k} - \vec{k}') \, , \tag{1.0}$$

where $P(k)$ is the power spectrum and completely characterizes the fluctuations. It may be shown that the request of having the same amount of logarithmic power (i.e. $k^3 P(k)$) on all scales at the horizon crossing implies that $P(\tau, k) \propto k$ (Harrison-Zel'dovich spectrum).

Deviations from scale invariance are parameterized through the introduction of the scalar spectral index for perturbations, n_s, such that

$$P(\tau, k) \propto k^{n_s} \, , \tag{1.0}$$

where $n_s = 1$ reduces to the pure scale invariance.

The power spectrum of gravitational waves is also characterized by Gaussianity and defined similarly:

$$< H_T(\vec{k}) H_T(\vec{k'}) >= P_T(k)\delta(\vec{k} - \vec{k}') \, . \tag{1.0}$$

$P_T(k)$ is also a power law in most models of inflation, characterized by an amplitude r and a spectral index n_t. In the simplest model, where the inflation energy density is represented by a single scalar field with a definite mass, r and n_T are related by:

$$n_t = r/6.8 \, . \tag{1.0}$$

Measuring these parameters is equivalent to constraining the inflationary mechanism.

1.5.2 Total intensity, E and B mode angular power spectra

Now that we have the solution for the Boltzmann equation we can compute the C_l of the CMB. The last step to be made is the connection between an observed quantity in the *lab*−frame with the one evaluated in the $\hat{k}$−frame. For the angular expansion coefficients of the temperature anisotropy it is (Hu and White, 1997):

$$C_l = \frac{2}{\pi} \int k^2 dk |\Theta_l(\tau, k)|^2 \ . \tag{1.0}$$

In Fig.1.2 we show the theoretical CMB power spectra for total intensity and polarization, according to the current Cosmological Concordance Model.

Let us start with the TT power spectrum (black line). At large scales ($\ell < 100$) the spectrum is almost flat because the angular scales correspond to wavelengths which were outside the horizon at decoupling, reflecting scale invariance. At lower multipoles, the power spectrum grows slightly because of the Integrated Sachs-Wolfe effect, that enhances fluctuations at large scales, but rapidly decays as contributions at smaller scales tend to average out.

As the perturbations enter the horizon in the tight coupling regime, they start to oscillate as described before. The position of the peaks correspond either to maximum compression or maximum rarefaction in the fluid. The phase, the height and the relative amplitudes of the peaks of these *acoustic oscillations* in the CMB power spectrum are most important tools to infer many parameters in the model. Finally the damping tail of the spectrum is due to the loss of coherence in the anisotropies because of the finite width of the LSS: small perturbations are washed away as the different photons decoupled in different regions within the LSS.

For what concerns the EE power spectrum (red line) the first feature to note is that the acoustic oscillations are in counterphase with respect to TT. Indeed, in the tight coupling regime, at the time in which the acoustic oscillations take place, it may be shown that the quadrupole is proportional to the dipole of the local CMB anisotropy distribution, Θ_1, and that the latter is essentially the velocity of the photon baryon fluid. The maximum of the velocity is obviously reached in the middle of an oscillation, between compression and rarefaction, i.e. when the temperature power spectrum is in a well (to be noted here that Θ_1 is proportional to the derivative of Θ_0). Another feature is the bump at $\ell < 10$. It is due to reionization: the photons, in their travel to us, interacted with the local reionized medium in the Universe. With a mechanism analogous to the one on the LSS, the quadrupole of the photon radiation scatters with free electrons and power is then generated on large scales (corresponding to the horizon at that time). Finally, the TE power spectrum (blue line) expresses the high level of correlation between total intensity and polarization. This reflects the fact that polarization is mostly due to the temperature anisotropies.

The BB power spectrum (green line) is actually the collage of three different physical processes. First of all it shows the same bump at low multipoles as the EE power spectrum. Second, there is a peak at $l \sim 100$. As we know, at the decoupling the quadrupole contains the term due to the gravitational waves, H_T, that is almost monochromatic in ℓ, i.e it selects multipoles that correspond to perturbations that were entering the horizon at the decoupling. Indeed, for scales larger then the horizon no polarization is expected since super-horizon perturbations cannot interact by definition, while, at smaller scales, once the perturbations entered the horizon they are rapidly washed away, because they are massless and do not experience any pressure. This explains the rising shape of the spectrum up to $\ell \sim 100$ that is quickly damped at larger multipoles. The rise we actually see in the plot for $\ell > 100$ is the contribution of a third effect: the gravitational lensing. Indeed, photons traveling to us are perturbed by local gravitational fields that twist the polarization patterns of the CMB. Reminding Fig.1.1, it is easy to understand that even a small change in the patterns causes a mixing between E and B modes. Since the former are much larger then the latter, the contamination in the BB power spectrum is expected to be high. The power at multipoles around 1000 is then due to power coming from the E modes, as the evident correlation between the two lines demonstrates. The gravitational lensing is a valuable tool to map the matter in the Universe and probe structure formation and dark energy, as the cross section of the lensing is exactly half a way between the LSS and us, at about $z \sim 1$ (Lewis & Challinor, 2005; Acquaviva et al., 2006).

1.6 Status of CMB data and future expectations

Since its discovery by Penzias and Wilson in 1965, the CMB has been the target of many experiments. The first one to find evidence for the CMB anisotropies was the COsmic Background Explorer (COBE) satellite that observed the CMB on all the sky on angular scales larger than 7^o (Smoot et al., 1992).

While another satellite mission was being prepared, many sub-orbital experiments measured the CMB radiation on limited portions of the sky. Among these we quote here BOOMERANG (de Bernardis et al., 2000) and MAXIMA (Hanany et al., 2000) who were the first to detect the acoustic peaks and to infer the flatness of the Universe, while the first experiment to find evidence for the polarized E modes was DASI (Kovac et al., 2002).

A second generation satellite, WMAP, was launched in 2001 and successfully arrived to the fifth year of observations. WMAP was able to observe the whole sky at 5 different frequencies (from 23 up to 94 GHz) with an angular resolution about 10 times better than COBE (see Tab.1.1 for more details). The 5-year release of the data is currently the best available all-sky measure of the CMB, both in total intensity and polarization. In Fig.1.3, we report the current status of the observations for what concern the total intensity and E mode power spectra, compared with the best-fit theoretical model (red line). Together with this, it is worth quoting the work

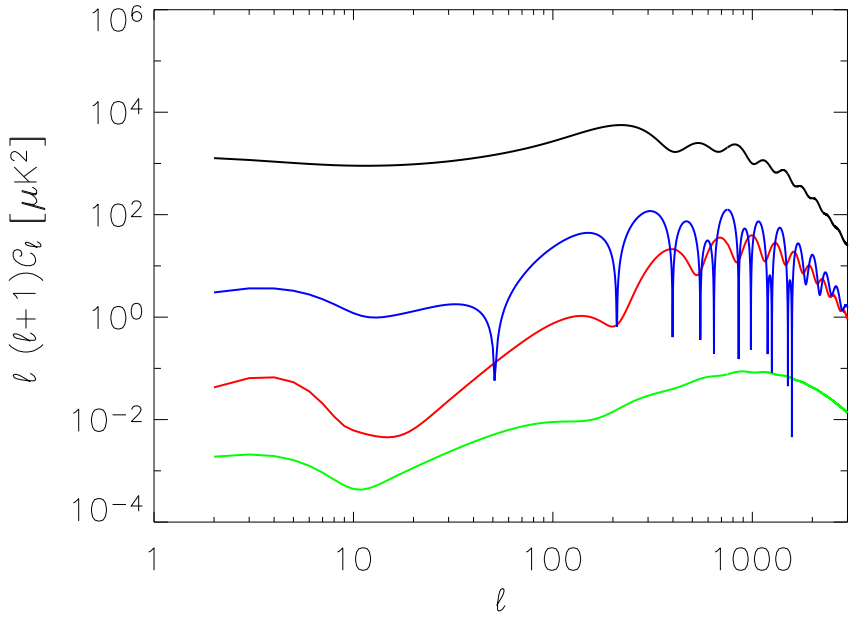

Figure 1.2: Theoretical CMB power spectra for the λCDM model described in this Chapter: TT (black line), TE (blue line), EE (red line) and BB (green line). Values for the parameters are the current best-fit to the data. The tensor to scalar ratio is $r = 0.1$.

by Reichardt et al. (2008), who saw evidence for weak gravitational lensing of the CMB at more than 3σ significance, by comparing the likelihood for the best fit of lensed and unlensed models to the ACBAR+WMAP 5-year data.

Currently, many CMB experiments are being prepared and planned. Let us see into details some of them.

1.6.1 Planck

The Planck satellite is a mission of the European Space Agency aimed to provide a measure of the microwave sky with an unprecedented sensitivity in a very broad frequency band, both in total intensity and polarization. The launch is supposed to be at the beginning of 2009. Once the satellite will have reached the observation point, Planck will start a 14 months survey of the all sky, using the two instruments on board: the Low Frequency Instrument (LFI), that uses radiometers in the range 30-70 GHz, and the High Frequency Instrument (HFI), that uses bolometers in the range 100-857 GHz.

The Planck mission has many different objectives. The first, most obvious one is to provide an all-sky high sensitivity CMB map, both in total intensity and polarization, with an angular resolution down to 5 arc-minutes and sensitivity of $\sim 10\mu K$. This will allow to tighten the constraints on the most important cosmological parameters, as well as investigate open problems such as discriminating between different inflation models, the nature of dark energy and dark matter. Finally it will set stringent upper limits on the presence of gravitational B modes.

Beside these important aims, Planck will also provide astrophysical informations on the Galactic and extra-galactic emissions, enormously expanding our knowledges on the physics of these processes. As we will see later, this is extremely important for the study of the CMB itself.

In Fig.1.4 and 1.5, we report the comparison of the expected Planck performances with respect to COBE and WMAP. Compare the latter of these plots with the one showing the current status of the observations for the power spectrum estimation in Fig.1.3. Planck promises to be a big leap forward in the knowledge of the CMB.

	No. channels	Frequency coverage	Angular resolution	Sensitivity ($mK\sqrt{s}$)
WMAP	5	23-94 GHz	(up to) 15 arc-min	(up to) 0.8
Planck LFI	3	30-70 GHz	(up to) 14 arc-min	(up to) 0.17
Planck HFI	6	100-853 GHz	(up to) 5 arc-min	(up to) 0.05
EBEX	3	150-420 GHz	(up to) 8 arc-min	(up to) 0.01

Table 1.1: Comparison between the nominal performances of WMAP, Planck (LFI+HFI) and EBEX.

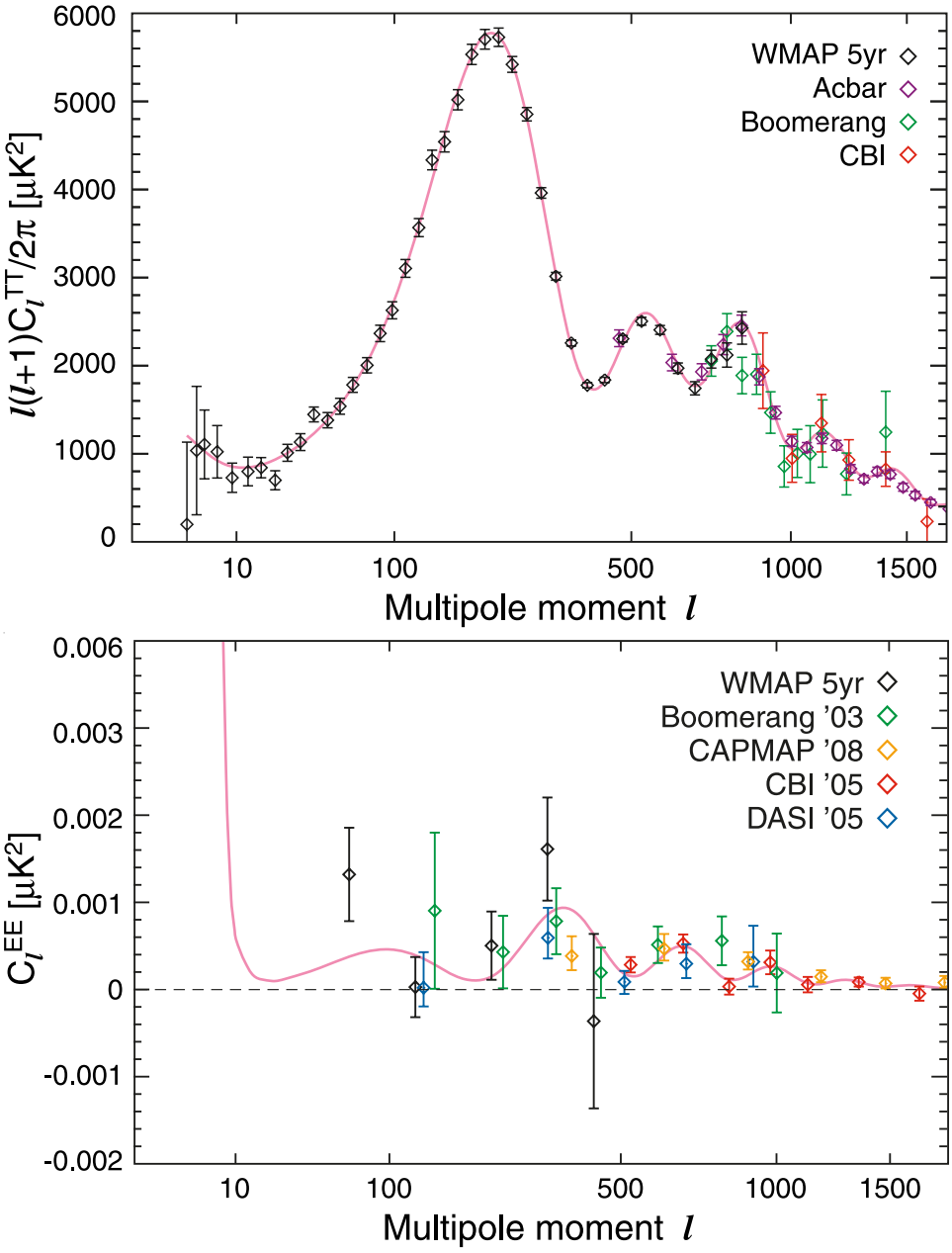

Figure 1.3: Current status of the observations of the total intensity and E mode power spectra. The solid red line is the best-fit theoretical model. From Nolta et al. (2008).

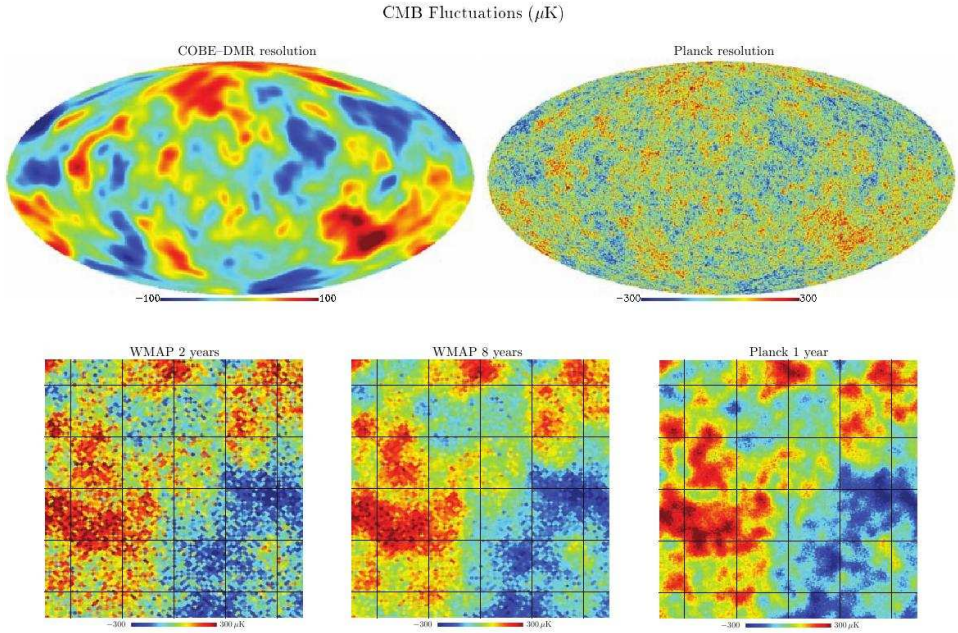

Figure 1.4: Comparison between the expected resolution of Planck and the ones achieved by COBE-DMR (upper panel) and WMAP (lower panel).

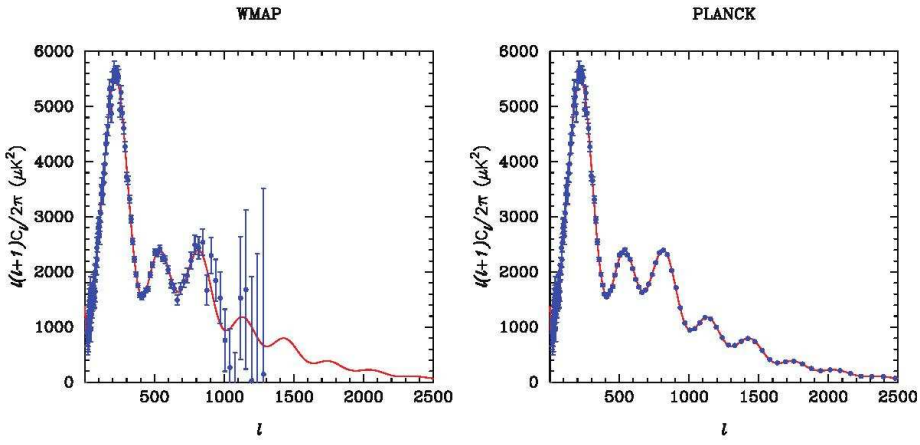

Figure 1.5: Reconstruction of a realization of the CMB power spectrum, for the ΛCDM model (red line), by WMAP on the left and Planck on the right.

1.6.2 Sub-orbital B mode probes

The cosmological B modes are the holy grail of the forthcoming CMB observations. The faintness of the curl component and the presence of diffuse emission from our own Galaxy (that will be the central topic of the next Chapter) make the detection of this signal extremely difficult in an all-sky survey. Sub-orbital experiments are being planned[1] to deeply observe small, clean patches of the sky, to achieve at the same time two important results: a high signal to noise level and a low contaminated CMB map.

Among these experiments we cite here: Spider[2], the Polarization of Background Radiation (PolarBear[3]), Clover[4] and the E and B EXperiment (EBEX[5]). Since I'm a collaborator of the last one, I will introduce it briefly. It is a NASA funded balloon borne instrument optimized for the measurement of the linearly polarized component of the CMB anisotropies, and scheduled for a first flight from North America in spring 2009, and a second one from Antarctica one year later; the angular resolution of 8 arc-minutes, sensitivity on the same angular scale of about 0.5 μK, as well as the frequency coverage between 150 and 410 GHz, have been optimized for the observation of the curl component of CMB polarization. In the Antarctica flight, EBEX will exploit 1406 bolometer detectors to observe an area of $\sim$ 1% of the sky in a region expected to be have a very low Galactic contamination (see Fig.1.6). We will talk more about EBEX in the last Chapter.

[1] for a complete list: http://lambda.gsfc.nasa.gov/

[2] http://www.astro.caltech.edu/ lgg/spider_front.htm

[3] http://bolo.berkeley.edu/polarbear/

[4] http://www-astro.physics.ox.ac.uk/research/expcosmology/groupclover.html

[5] http://groups.physics.umn.edu/cosmology/ebex

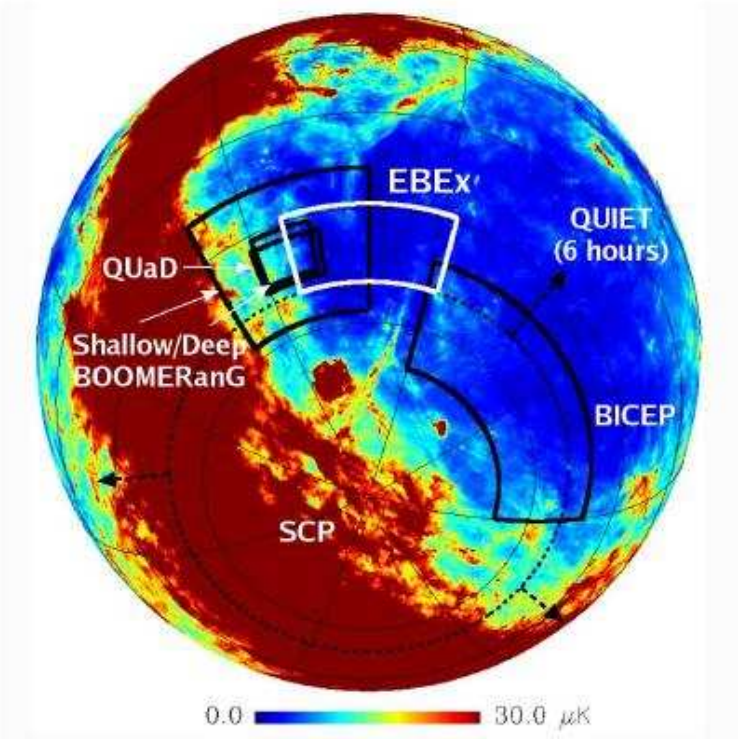

Figure 1.6: Target areas of some CMB experiments. The red region show the presence of the Galaxy emission and is expected to be highly contaminated (see next Chapter for more details).

Chapter 2

Foregrounds

2.1 Introduction

In this Chapter we will review the current knowledge about the Galactic emission in the microwave band. As pointed out before, our capability in removing the different emissions that cover and pollute the background signal will be one of the ultimate limitation in the CMB science. These *foregrounds* are indeed known to be a serious contaminant at all the frequencies of cosmological interest, both in total intensity and in polarization.

Foregrounds can be divided into two main categories: diffuse and point-like. The former are due to diffuse emission from our Galaxy, while the latter come from unresolved (in the telescope beam) sources, both Galactic and extra-galactic[1]. Obviously, since they enter into the game at different angular scales, these two types of foreground have different impact on the CMB power spectrum and have to be addressed accordingly in the data analysis. During my PhD I have mostly focused on the diffuse foregrounds cleaning, developing and testing component separation techniques, as I will show in the next Chapter. For this reason, we will focus mainly on the diffuse foregrounds and point-like sources will be treated only marginally.

We will start by introducing the physics of the most important diffuse Galactic emissions and their observations, both in total intensity and polarization, outside the microwave band. Then, we will review the main results of WMAP for what concern the foregrounds: what that data-set can tell us and what is still missing. Finally, we will conclude the Chapter with a discussion on how the diffuse foregrounds contaminate the CMB signal.

[1] N.b.: In this distinction, we are neglecting a few galaxy clusters that are still resolved in a typical CMB experiments

2.2 Diffuse Galactic foregrounds

The diffuse emission in the microwave sky, beside the CMB, comes mainly from three Galactic components: the synchrotron radiation from electrons spinning in the magnetic field, the free-free radiation associated with ionized interstellar gas and the thermal radiation of interstellar dust. Moreover, as we will see later, there are also several clues on the existence of another diffuse Galactic component that is still not clearly identified today but it is likely to be related to spinning dust grains. We will conservatively call it *anomalous emission*.

Before entering in the discussion, it is useful to introduce the following quantity that we are going to use to define the intensity of the signals: the antenna (or brightness) temperature. It is the temperature at which a blackbody would have to be in order to mimic the observed intensity I_ν of an object at a frequency ν in the Rayleigh-Jeans limit:

$$T_b = \frac{I_\nu c^2}{2\nu^2 k} \tag{2.0}$$

where c is the speed of light and k is Boltzmann's constant.

2.2.1 Synchrotron

Synchrotron emission arises from the acceleration of cosmic ray electrons in magnetic fields. The morphology of the observed emission depends both on the distribution of the relativistic electrons in the Galaxy and on the Galactic magnetic field structure. The latter is generally weak in our Galaxy (tipically of the order of a few microgauss) and then the diffuse synchrotron emission is mostly observed in the radio band. Another small contribution is expected from discrete supernova remnants, but they contribute only $\sim$ 10% of the total synchrotron emission at 1.5 GHz (Ulvestad, 1982).

Across the Galactic plane, the Galactic magnetic field exhibits a large-scale ordering with the field parallel to the spiral arms (regular component). Superimposed to this, there is an irregular component with small scale structures which shows variations between the arm and inter-arm regions and with gas phase. The regular and irregular components seem to be of comparable magnitude. At high latitudes, there is a contribution from the Galactic halo, and specific nearby structures, i.e. the North Polar Spur (Rybicki & Lightman, 1979) .

One of the most important properties to characterize the synchrotron emission is its behaviour with frequency, usually characterized by a spectral index, β_s. Indeed, over a large frequency range, the intensity, in antenna temperature, of this component can be modeled by means of a single power law:

$$T_s(\nu) = A_s \left(\frac{\nu}{\nu_0}\right)^{-\beta_s} \quad , \tag{2.0}$$

where A_s is the synchrotron amplitude, in a given position of the sky, at the frequency ν_0. It exhibits a quick decrease with frequency, (typical values are $\beta_s \sim 3$),

reflecting the energy distribution of electrons.
Variations in the frequency spectral index of the synchrotron continuum emission arise from variations in the cosmic ray electron energy spectrum, which has a range of distributions depending on age and the environment of origin (e.g., supernova explosions or diffuse shocks in the interstellar medium). Reich & Reich (1988) used radio continuum surveys of the Northern sky at 408 and 1420 MHz, from Haslam et al. (1982)[2] to demonstrate a range of spectral index values between 2.3 and 3.0. The steepest spectra were observed towards the North Polar Spur, and there was a flattening in spectral index towards higher latitudes in the Galactic anti-center direction. Such behavior has been confirmed over the full sky by Reich et al. (2003), who find that the flattening is particularly pronounced in the Southern hemisphere. At higher frequencies, the antenna temperature spectral index is expected to steepen by 0.5 due to electron energy losses (Platania et al., 1998). Banday et al. (2003) derived a mean spectral index between 408 MHz and 19.2 GHz from the Cottingham (1987) survey and between 31.5, 53, and 90 GHz from the COBE-DMR data: they found a steep spectral index of ∼3.1 for Galactic latitudes $|b| > 15^\circ$, consistent with expectations. Bennett et al. (2003) claim that the spectral break occurs near the WMAP K-band. Spectral indices above 10 GHz are then likely between 2.7 to 3.2.

Given its physical origin, synchrotron radiation can be strongly polarized in the direction perpendicular to the Galactic magnetic field, in principle up to 75% (Rybicki & Lightman, 1979). The polarization has been measured at a number of frequencies from Leiden between 408 MHz to 1.4 GHz (Wolleben et al., 2006), from Parkes at 2.4 GHz (Duncan et al., 1999), and by the Medium Galactic Latitude Survey at 1.4 GHz (Uyaniker et al., 1999). The morphology of the signal showed a substantial power on small angular scales ($C_\ell \sim \ell^{-2}$) and no dependence on the Galactic latitude up intermediate latitudes. Moreover, even if the theoretical upper limit is higher, typically measured values for the fraction of polarization reached a few percent.

Indeed, these observations are effected by the Faraday depolarization. Electrons in the Galactic magnetic field rotate the direction of polarization because the left and right circular polarizations propagate with different velocities in the interstellar medium (Faraday rotation). Such rotation is a function of the electron density and the component of the Galactic magnetic field along the line of sight. In our Galaxy, given the strength and properties of the Galactic magnetic field, a typical value for the net rotation is $\Delta\theta \sim 420^\circ/\nu^2$, with the frequency ν expressed in GHz. This means that at WMAP frequencies this effect is negligible, while it becomes important in the radio band. For these reasons a crude extrapolation from the low frequencies can be problematic in polarization and great care must be taken when using low frequency maps as tracers of the synchrotron polarized emission at microwave frequencies.

[2]Haslam et al. (1982) represents one of the most important synchrotron surveys, covering the all sky, with a nominal resolution of 52 arc-minutes.

2.2.2 Free-free

The free-free is a non thermal emission due to bremsstrahlung of free electrons in presence of Hydrogen ions, in the so called Hα regions. The scatter of the electrons produces microwaves with an antenna temperature spectrum $T_{ff} \sim \nu^{-2.14}$ (Rybicki & Lightman, 1979), given a temperature of the electrons, $T_e \approx 8000$K. Radio astronomy provides no free-free emission maps because it does not dominate the sky at any frequency. Nevertheless, as discussed in Bennett et al. (2003), high-resolution maps of Hα emission can serve as approximate tracers of free-free emission and several large-area Hα surveys of the sky are indeed available. Among them we list here: the Wisconsin H-Alpha Mapper (WHAM[3]) that used Fabry-Perot spectroscopy on an angular scale of 1° covering the northern sky ($\delta > -30^\circ$), the Virginia Tech Spectral-Line Survey (VTSS) that exploited a filter system with an angular resolution of 1 arc-minute and covers the southern sky (Gaustad et al., 2001), and the Southern H-Alpha Sky Survey Atlas (SHASSA) that used a similar approach in the northern sky on a 3 arc-minutes scale Dennison et al. (1998).

The major issue in estimating the free-free emission out of these cited surveys is the fact that the absorption by interstellar dust must be taken properly into account; this correction is significant at intermediate and lower Galactic latitudes, as reported by Dickinson et al. (2003).

2.2.3 Thermal Dust

The thermal dust emission that contributes to the frequencies of interest for CMB analysis arises mostly from grains large enough to be in thermal equilibrium with the interstellar radiation field, and is known from analysis of the Infrared Astronomical Satellite (IRAS) and COBE-DIRBE data to peak at a wavelength of approximately 140 μm. The dust is mainly composed by a mixture of compact graphite and silicate grains. The spectrum of the dust radiation is a so called grey body, that is a not isolated black body warmed by the star light. At higher frequencies, there is a contribution from the optically active modes of the polycyclic aromatic hydrocarbons (PAH) molecules, but these are not interesting for our discussion. In the Galactic plane where several clouds overlap on the line of sight, results are easier to interpret in term of physical properties of dust. Nonetheless, studying the dust as a CMB contaminant requires a good knowledge of dust emission also outside the Galactic plane. At these high latitudes, the determination of the Galactic component relies on the existence of a spatial correlation between gas and dust and thus of gas emission lines with the associated dust emission. The correlation which has been the most extensively investigated is the one between IR emission and the 21 cm line from atomic hydrogen. It turns out that the dust spectrum is well fitted by a single Planck curve with an emissivity proportional to ν^2 and a temperature of 17.5 K (Draine & Lee, 1984).

[3] http://www.astro.wisc.edu/wham/index.html

In the microwaves, using the Far Infrared Astronomical Satellite (FIRAS) data, Finkbeiner et al. (1999) showed that the one temperature fit can be significantly improved by including a second emission component with a low temperature (T≈9 K) and an IR/visible emissivity ratio one order of magnitude larger than that of the warmer component. It is unclear yet what is the physical origin of such a cold component and in particular if it represents the emission from grains that are cold due to large sub-mm emissivity. This second component with a temperature of about 9 K can be seen as a sub-mm excess with respect to the T≈17 K modified black body. Such an excess has also been detected in the Archeops data (Benoit et al., 2004) and confirmed by Boomerang (Masi et al., 2001) and MAXIMA (Jaffe et al., 2004). The model by Finkbeiner gives for the dust flux, at the frequency ν:

$$\text{Dust flux} \propto \frac{f_1(q_1/q_2)(\nu/3000\ \text{GHz})^{\alpha_1+3}}{e^{h\nu/kT_1}-1} + \frac{f_2(\nu/3000\ \text{GHz})^{\alpha_2+3}}{e^{h\nu/kT_2}-1}\,, \qquad (2.0)$$

where $f_1 = 0.0363$, $f_2 = 1 - f_1$, $\alpha_1 = 1.67$, $\alpha_2 = 2.70$, $q_1/q_2 = 13$, $T_1 = 9.4$ K, $T_2 = 16.2$ K, and h and k are the Planck and Boltzmann constants, respectively.

Lots of experimental data on dust absorption of starlight have been gathered and show that this absorption generates polarization. The natural explanation of this is that dust grains are aspherical and that they are mainly aligned. A compilation of these measurements and a mapping of the polarization direction started by Serkowski et al. (1975) show that this alignment is highly correlated with the Galactic magnetic field and that optical polarization tends to be aligned with the magnetic field. It therefore implies that grains are orthogonal to the magnetic field on average. Qualitatively, the incident radiation (optical wavelength) polarizes the grains and generates electric dipoles that will in turn radiate. Millimetric or sub-millimetric measurements of the Galactic dust polarization are usually concentrated in Galactic clouds and starforming regions with arc-minutes angular resolution (Hildebrand et al., 1999). They show a few per cent polarization and no clear frequency dependence. The first observations on large angular scales of the polarization of the Galactic dust emission are provided by Archeops at 353GHz (Benoit et al., 2004). The Archeops data show a significantly large scale polarized emission in the Galactic Plane, with a polarization degree of 45%, with several clouds of few square degrees appearing to be polarized at more than 10%.

2.2.4 Anomalous emission

In recent years evidence has been reported for an additional, thermal dust correlated component called anomalous microwave emission, that may dominate, at least in some Galactic regions, in the 20-40 GHz range where it has a frequency spectrum similar to synchrotron. It may be caused by tiny, fast spinning, dust grains (Draine & Lee, 1998; de Oliveira-Costa et al., 2004). Indications of this emission are found by analyses of several CMB experiments (see Bonaldi et al., 2007, and eferences therein).

Although Bennett et al. (2003) concluded that spinning dust emission contributes less than 5% of the WMAP Ka-band antenna temperature, evidence of the anomalous emission was uncovered by combining WMAP data with other measurements, especially at lower frequencies (Lagache et al., 2003; de Oliveira-Costa et al., 2004; Finkbeiner, 2003; Watson et al., 2005 ; Davies et al., 2006; Bonaldi et al., 2007). Moreover, according to Page et al. (2007), the polarization properties of this dust-correlated low-frequency component differ from those of the synchrotron, suggesting a different emission mechanism. Nevertheless, the exact nature, the spectral properties and the spatial distribution of this foreground remain uncertain. There is not even general agreement over its existence; a combination of free-free emission and strongly self-absorbed synchrotron could also account for the data, according to Hinshaw et al. (2006). In the next Chapter, we will come back on this topic in the discussion related to the work by Bonaldi et al. (2007), in which we found evidences of an anomalous emission, strongly correlated with thermal dust, in the WMAP 3-year data.

2.3 WMAP observations

Up to now, the most powerful all-sky measure of Galactic foregrounds in the microwave band came from WMAP. In this Section we will review their analysis, both in total intensity and polarization.

To disentangle the different sources that are present in a sky map, it is necessary a dedicated study called *component separation*. This will be the central topic of our discussion in the next two chapters, but we need to anticipate some of concepts here, altough the WMAP methodology differs from the ones we will study. The WMAP team used a Maximum Entropy Method (MEM) to separate foregrounds in total intensity, as proposed by Bennett et al. (2003). This is a component separation method, not specialized for cleaning the CMB, but useful in separating the different Galactic components exploiting their different emission mechanisms and frequency behaviours. The basic idea is to minimize a χ^2 of the modelled data, combined with some prior informations about the components.

The starting point of the method is to model the components that are expected in the data:

$$T_{\rm m}(\nu, p) \equiv T_{\rm cmb}(p) + S_{\rm s}(\nu, p)\, T_{\rm s}(p) + S_{\rm ff}(\nu, p)\, T_{\rm ff}(p) + S_{\rm d}(\nu, p)\, T_{\rm d}(p), \qquad (2.0)$$

where the subscripts cmb, s, ff, and d denote the CMB, synchrotron, free-free, and thermal dust components, respectively. $T_c(p)$ is the spatial distribution of the component c at the pixel p, and $S_c(\nu, p)$ is the spectrum of the emission, which is not assumed to be uniform across the sky.

The model is then fit in each pixel by minimizing the functional $H = A + \lambda B$, where $A = \sum_\nu [T(\nu, p) - T_m(\nu, p)]^2/\sigma_\nu^2$, is the standard χ^2 of the model fit, and $B = \sum_c T_c(p)\, \ln[T_c(p)/P_c(p)]$ is the so called MEM functional, that contains the

prior informations (Press et al., 1992). The parameter λ controls the relative weight between the data and the priors in the fit. In the functional B, the sum over c is restricted to Galactic emission components, and $P_c(p)$ is a prior estimate of $T_c(p)$. The form of B ensures the positivity of the solution $T_c(p)$ for the Galactic components, which greatly alleviates degeneracy between the different foregrounds.

For the prior informations, Hinshaw et al. (2006) exploited foreground templates built using observations out of the microwave band, as discussed above in this Chapter.

A prior estimate for dust emission, $P_{\rm d}(p)$, was done using Model 8 of Finkbeiner et al. (1999), evaluated at 94 GHz. The dust spectrum is modeled as a straight power law, $S_{\rm d}(\nu) = (\nu/\nu_W)^{+2.0}$. For free-free emission, the prior, $P_{\rm ff}(p)$, was estimated using the extinction-corrected Hα mosaic (Finkbeiner, 2003). The spectrum was modelled as a straight power law, $S_{\rm ff}(\nu) = (\nu/\nu_K)^{-2.14}$. For the synchrotron emission, they constructed a prior estimate, $P_{\rm s}(p)$, using the Haslam 408 MHz map and scaling it to K-band assuming $\beta_s = -2.9$.

Fig.2.1 shows the three input prior maps, $P_c(p)$, and the corresponding output component maps, $T_c(p)$, obtained from the 3-year data. These maps are available to the public as part of the WMAP 3-year data release. The maps are displayed using a logarithmic color stretch to highlight a range of intensity levels. The morphology and amplitude of both the thermal dust and free-free emission are well predicted by the prior maps.

The most notable discrepancy between prior and output maps is seen in the synchrotron emission. Specifically, the K-band signal has a much more extended Galactic longitude distribution than the 408 MHz emission, and it is remarkably well correlated with the thermal dust emission. Whether this is K-band non-thermal component due to anomalous dust emission or mostly flat-spectrum synchrotron emission that dominates at microwave frequencies is not clear. Hinshaw et al. (2006) claim that the answer to this question was not possible with the WMAP data alone because the covered frequency range does not extend low enough to disentangle the two. We will be back on this at the end of the next Chapter.

Polarization observations are available at all the WMAP channels. The measured polarization amplitude P and the angle γ are shown in Fig.2.2. The polarization vectors have length that is logarithmically dependent on the magnitude of P and direction according to γ. These vectors are plotted whenever the signal to noise (P/N) is greater than unity (see the figure caption).

The analysis of these data is much more difficult with respect to the total intensity case. The signal to noise ratio is much lower and our prior knowledge of the polarized foregrounds much poorer. Then, the Maximum Entropy Method could not be used on this data set. In the analysis made by Page et al. (2007), the WMAP team proposed a model for the polarized Galactic emission to be used to explain the general features in the observations.

The two dominant components of the polarized foreground emission in the 23 – 94 GHz range are synchrotron and thermal dust (Bennett et al., 2003). The

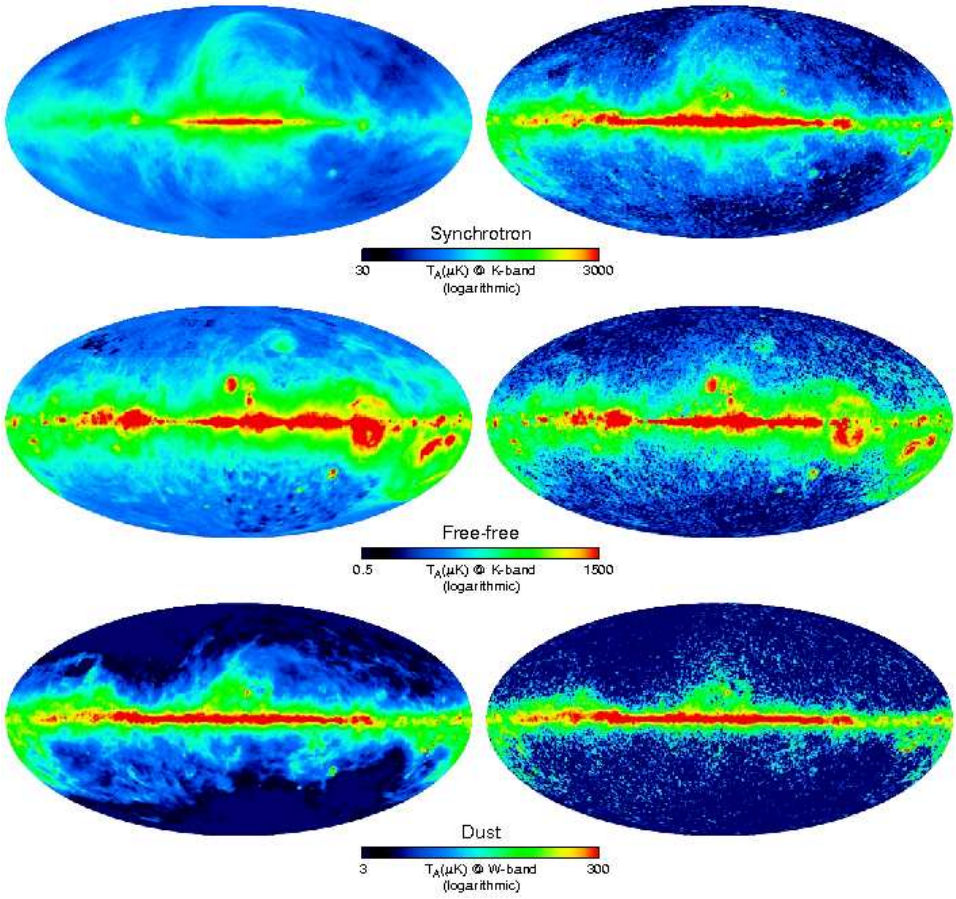

Figure 2.1: Galactic signal component maps from the Maximum Entropy Method (MEM) analysis. *top-bottom*: synchrotron, free-free, and dust emission with logarithmic temperature scales. *left*: Input prior maps for each component. *right*: Output maps based on WMAP 3-year data. From Hinshaw et al. (2006).

ingredients for the polarization model are a description of the Galactic magnetic field and some rule to define both the fraction of polarization and the polarization angle across the sky. The model can be written as:

$$Q_c = I_c(\hat{n})\,\Pi_c\,g_c(\hat{n})\,cos(2\gamma_c) \qquad (2.1)$$
$$U_c = I_c(\hat{n})\,\Pi_c\,g_c(\hat{n})\,sin(2\gamma_c),$$

where I_c is the total intensity template of the component c (either synchrotron or dust) and Π_c is its polarization fraction multiplied by an effective function $g_c(\hat{n})$ that takes into account cancellation effects along the line of sight . Finally γ_c is the polarization angle.

In order to estimate the polarization angle it is necessary to model the Galactic magnetic field properly. As explained above, the angle γ_c is expected to be the same

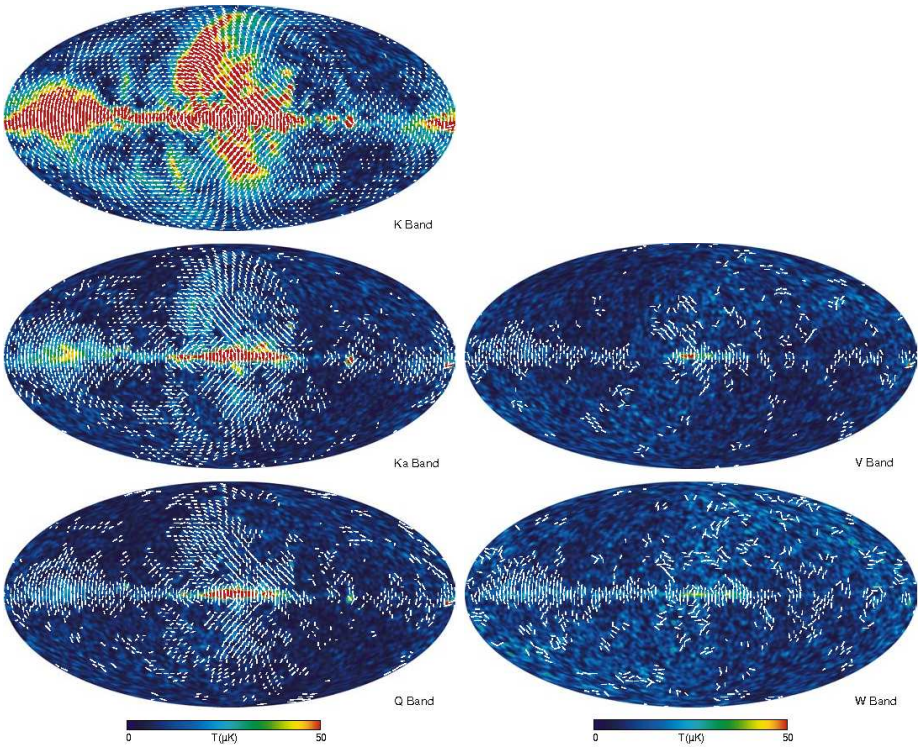

Figure 2.2: P and γ maps for the WMAP bands in Galactic coordinates. The polarization vectors are plotted whenever a pixel, (roughly 4 deg $\times$4 deg) and three of its neighbors has a signal to noise (P/N) greater than unity. The length of the arrow is logarithmically dependent on the magnitude of *P*. From Page et al. (2007).

both for synchrotron and dust. It is computed as a function of the two components of the Galactic magnetic field that are perpendicular to the line of sight. We are not entering into details of such a model here (see Page et al., 2007, for the full explanation), we just point out that they exploited informations from the WMAP K band itself to partly fit such a model. In Fig.2.3, we report a comparison between the angle of the magnetic field, $\gamma_M = \gamma_c + 90°$, derived from the synchrotron radiation in the K-band map (smoothed with a 4° beam) and the predicted magnetic field direction given by the simple model adopted by Page et al. (2007).

The ultimate purpose of these foreground models was CMB cleaning, more than extracting good templates for the Galactic emissions, because, as we will see in the next Section, the diffuse foreground contamination turns out to be remarkably more severe in polarization than in total intensity. Nevertheless, some global physical features still can be drawn out of the data and the model constructed for them. First of all, it was possible to measure the large scale structure of the Galactic magnetic field, mainly thanks to the strong synchrotron emission in the K band, and with that to infer the polarization angle at those scales for synchrotron. More-

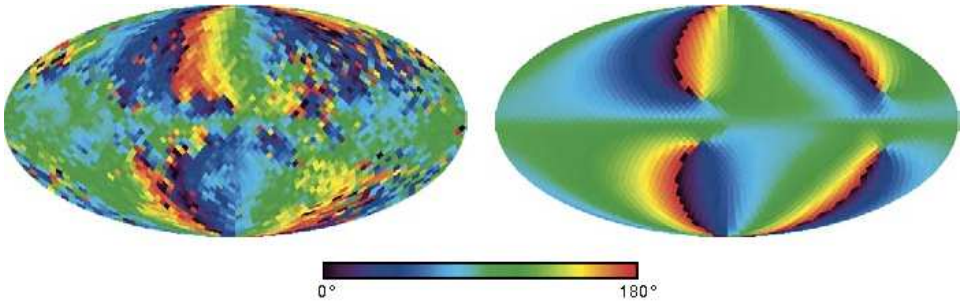

Figure 2.3: *Left:* The angle of the magnetic field, $\gamma_M = \gamma_{PA} + 90°$, derived from the synchrotron radiation in the K-band map (smoothed with a 4° beam). *Right:* The predicted magnetic field direction given by the simple model adopted by WMAP. From Page et al. (2007)

over, assuming the Galactic magnetic field is 100% efficient in determining the polarization direction of both synchrotron and dust, the latter was also modelled accordingly, for what concerns teh poalrization angle.

Moreover, the WMAP data provided a measure for the polarization fraction. It turned out to be $\Pi_s \sim 15\%$ for the synchrotron and $\Pi_d \sim 5 - 10\%$ for the dust, confirming previous results (Uyaniker et al., 1999; Benoit et al., 2004). These measures refer mostly to low and intermediate latitudes. For synchrotron, Π_s is lower than theoretical expectation ($\sim 75\%$) because of cancellation along the line of sight.

2.4 Contamination to the CMB

In this Section we study the foreground properties as contaminants to the CMB.

The first, macroscopic distinction to make is the one between total intensity and polarization. In the former case, the CMB is expected to dominate the sky emission, out of the galactic plane, at least in a certain range of frequencies. This is no more true for polarization where, as we will see in a while, the foreground emission highly contaminates the background signal at any frequency of interest. Because of this difference, the WMAP team during the analysis had to cut out two different portions of the sky to avoid the major contamination of foregrounds. The masks they used are called Kp2 and P06, for total intensity and polarization respectively, and are shown in Fig.2.4.

A first consideration to be made concerns the behaviour of the different emissions in frequency and the their relative level with respect to the CMB. In total intensity the situation is sketched in Fig.2.5, where the spectra of the CMB and of the other sources of contamination are shown. These levels refer to an angular resolution of 1°, out of the Kp2 mask. The high frequency part of the spectrum (frequencies greater than 90 GHz) is dominated by the thermal emission of the

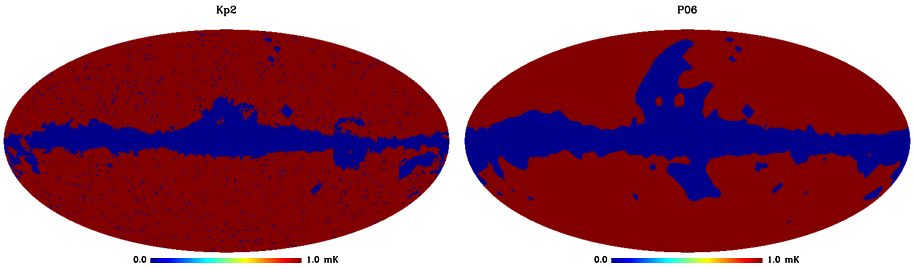

Figure 2.4: Mask used in the WMAP analysis of the CMB, Kp2 for total intensity (left) and P06 for polarization (right). The sky fraction that is cut out is 18% and 27% respectively.

dust, while lower frequencies are dominated by free-free and synchrotron, showing different slopes as explained above. Grey bands are the frequency coverage of the Planck channels.

The difference in the behaviour of the components in the frequency domain is the key feature that allows to disentangle them. Component separation techniques exploits exactly this handle in order to separate the different emissions, as we will see in next Chapter.

The contamination level is also a function of the angulare scale. The next two plots, in Fig.2.6 and 2.7, report the total intensity and polarized emissions in the sky at the different WMAP channels and their comparison with the CMB power spectra, outside the Galactic plane, using the Kp2 and the P06 masks resectively. These plots clearly show that, while in total intensity the contamination is significantly lower outside the plane, in polarization, even if most of the foreground emission is cut out with the mask, the Galactic contribution to the total microwave signal is expected to highly contaminate the CMB E modes and to totally dominate the CMB B modes at any frequency in the band.

In the work by Page et al. (2007), they propose a very simple model to parametrize the foreground emission outside the P06 mask region:

$$\ell(\ell+1)C_\ell^{fore}/2\pi = (\mathcal{B}_s(\nu/65)^{2\beta_s} + \mathcal{B}_d(\nu/65)^{2\beta_d})\ell^m. \qquad (2.0)$$

where we have introduced the notation $\mathcal{B}^{XX} \equiv \ell(\ell+1)C_\ell^{XX}/2\pi$ to simplify the expression. From an unweighted fit to all the raw $\ell < 100$ data with the dust index fixed at $\beta_d = 1.5$, they find for EE $\mathcal{B}_s = 0.36\ (\mu\mathrm{K})^2$, $\beta_s = -3.0$, $\mathcal{B}_d = 1.0\ (\mu\mathrm{K})^2$ and $m = -0.6$; and for BB $\mathcal{B}_s = 0.30\ (\mu\mathrm{K})^2$, $\beta_s = -2.8$, $\mathcal{B}_d = 0.50\ (\mu\mathrm{K})^2$ and $m = -0.6$. This model obviously gives an approximate guide. In Fig.2.7 the dashed red line represents this equation evaluated at the minumum (60 GHz), for B modes.

The high level of the polarized foregrounds makes it very difficult the detection of the cosmological B modes in the CMB, even at the most optimistic levels of T/S. Cleaning two different foregrounds, that are, at least, one or two orders

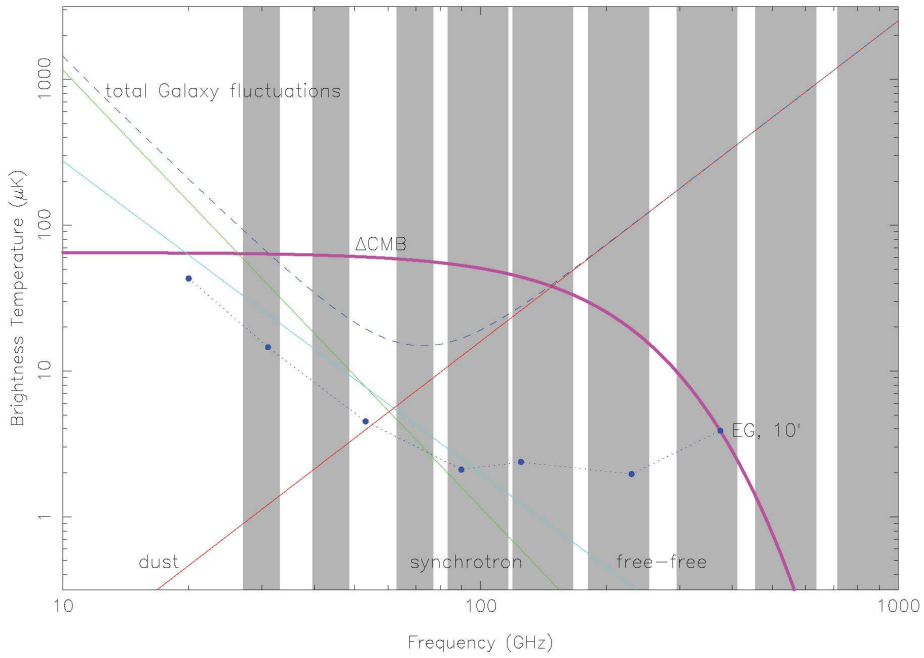

Figure 2.5: This figure, from the Planck Bluebook, shows the spectrum of CMB and the frequency coverage of the Planck channels (grey bands). It also shows the spectra of other known sources of fluctuations in the microwave sky. Dust, synchrotron and free-free temperature fluctuation levels referring to an angular resolution of $\sim 1^o$, correspond to the WMAP observed levels outside the Galactic plane.

of magnitude above the signal to recover is hard to believe. Fortunately, the foreground level is not constant across the sky and there are a few regions where it drops down considerably. One of these clean patches has been observed in the past (see Montroy et al., 2005, and references herein) and is the target of EBEX (Oxley et al., 2004). This patch is centered at RA=60°, DEC=-50° in Galactic coordinates and is more or less $\sim$ 1% of the sky.

Due to the theoretical expectations outlined in the previous Chapter, the detection of CMB B modes and the challenges posed by foregrounds are one of the main issues in modern CMB data analysis, and will be extensively discussed in the last Chapter. We conclude here by describing the expected foreground contamination to B modes in selected areas of the sky, focusing on the one just mentioned, sketched in Fig.2.8. It shows the spectrum of foregrounds B modes from Galactic dust and synchrotron expected outside P06 (dashed) and in the EBEX area (solid), as well as the CMB B modes (black, solid), at $\ell = 100$. The latter was obtained simply by assuming that the dust power, as estimated by WMAP, scales down as the ratio of the variance between the EBEX patch and the outside P06 area. Even if the contamination from the dust is still challenging, more than two orders of mag-

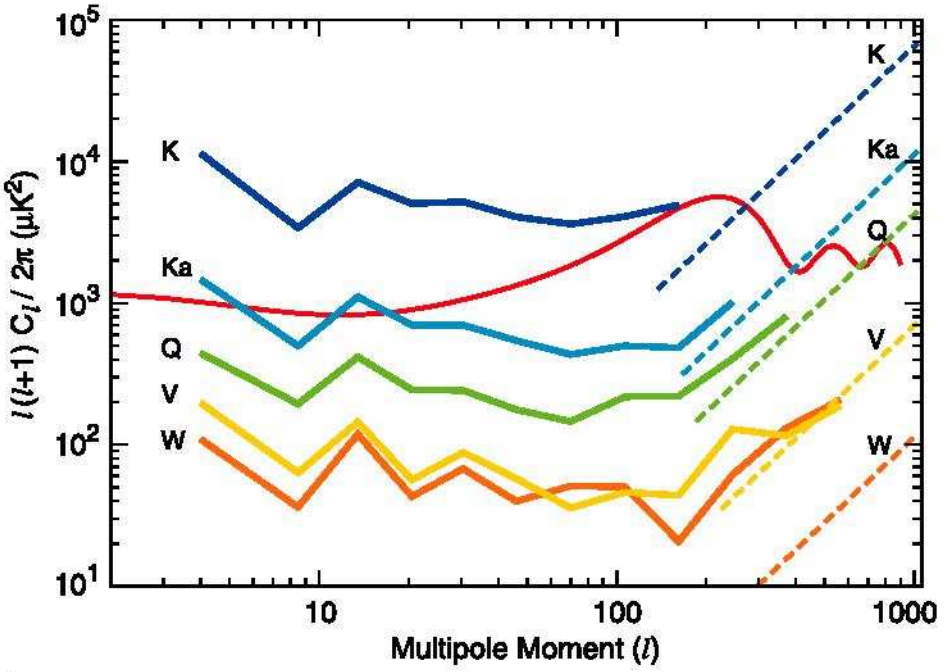

Figure 2.6: The measured foreground power spectra are shown for each WMAP band using the Kp2 mask. The smooth red line is the CMB from the best fit CDM model. Contamination from unmasked point source (expected to be a l^2 term) are shown in the dashed lines on the right. From Bennett et al. (2003).

nitude could be gained, simply by restricting the observations to this very clean patch. In the last Chapter we will show how component separation codes can help in finishing the job.

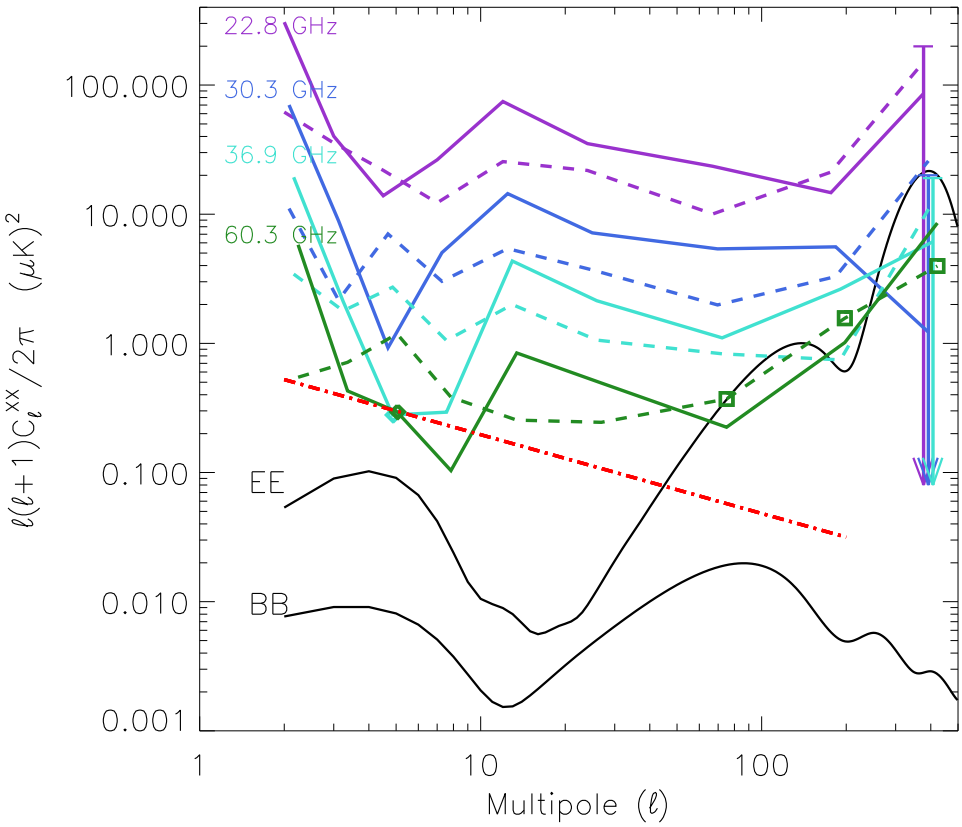

Figure 2.7: The absolute value of the EE (solid, violet through green) and BB (dashed, violet through green) polarization spectra for the WMAP data outside the Galactic plane. Black lines are CMB polarization power spectra from the best fit CDM model with a tensor contribution of r = 0.3. The dashed red line represents Eq.(2.4) evaluated at the minumum (60 GHz), for B modes. From Page et al. (2007).

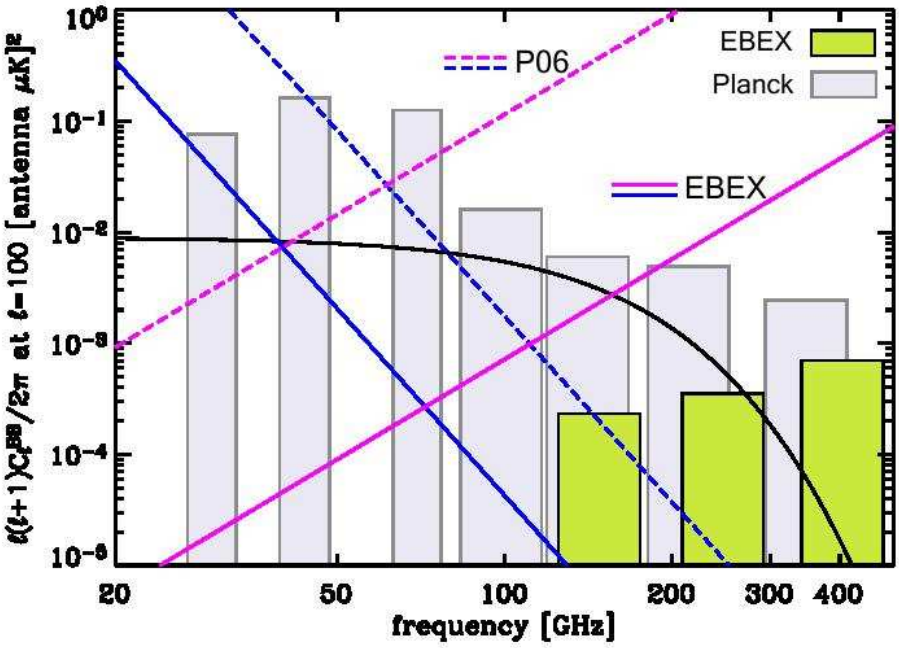

Figure 2.8: Comparison of power spectrum of foregrounds B modes, outside P06 (dashed) and in the EBEX area (solid). The black solid line is the level of CMB B modes for $r = 0.03$. All the spectra are considered at the angular scale corresponding to $\ell = 100$. The bands show the EBEX (yellow) and Planck (grey) frequency channels and instrumental noise level. From the successful EBEX proposal to NASA for funding support up to the Antartica flight.

Chapter 3

Component separation techniques

3.1 Introduction

In the last Chapter, our current knowledge of foregrounds has been summarized. An important lesson that the CMB community has learnt in the past years is that the Galactic foreground emission dangerously contaminates the cosmological signal (see for example Page et al., 2007)) and that this issue has to be addressed as a data analysis problem. Cosmologists are now aware of the fact that a new piece of pipeline is required for cleaning and/or separating foregrounds in the data and an entire line of research is being dedicated to the development and testing of algorithms for the so called *component separation*. The latter is a very general expression to mean any algorithm or data processing aimed at discriminating between different physical sources of emission.

The component separation problem can be stated as follows. Let's model the data **m** as

$$\mathbf{m}_u = \mathbf{s}_u + \mathbf{n}_u, \tag{3.0}$$

where **s** is the total signal and **n** is the instrumental noise. The subscript u indicates whatever data item these quantities depend on (time t, pixel p, multipoles ℓ, etc...). Quantities in the above equation are all vectors that have as many elements as the number of available measures n_m. In general the signal **s** denotes a linear combination of n_c physical components **c**:

$$\mathbf{s}_u = \mathbf{A}_u \mathbf{c}_u, \tag{3.0}$$

where **A** is called mixing matrix. The final purpose of component separation is to invert the System (3.1) in presence of noise and recover the components **c**. To cast the problem in a cosmological context, the components **c** play the role of the astrophysical sources (both Galactic foregrounds and CMB), while the mixtures **m** are the outputs of a multi-frequency experiment. The mixing matrix **A** takes into account the relative frequency scaling of each component in the data.

Quite a number of techniques have been recently proposed. These methods differ in many aspects and can be classified in various ways, according to anyone's tastes. For example, a typical distinction is the one between "blind" and "non-blind" techniques, depending on how strong are the assumptions made by the algorithm, i.e. what kind of *a priori* knowledges are imposed to infer the mixing matrix A.

I will use a different classification here dividing the different methods in three main categories, depending on what part of the whole data set they use in a single separation process. More specifically, the rationale behind the following scheme is the number of data items (pixels, multipoles, etc...) that are considered in the analysis at the same time:

- All the items,
- One item,
- Some of them.

Providing an extensive presentation of all the available methods is certainly beyond the purpose of this work and maybe found elsewhere (see Leach et al., 2008, and references therein). I'll mostly focus on three methods that I've been working on during the PhD: the fast independent component analysis (FASTICA, Maino et al., 2002), an approach based on the parameterization and fitting of the foreground and background unknowns (MIRAMARE, Stompor et al., 2008b), and the correlated component analysis (CCA, Bonaldi et al., 2006). Each of them can be considered a good example of the three categories introduced above. Indeed, as we will see later, FASTICA exploits the overall statistics of the signals to be separated, analyzing all the pixels together; MIRAMARE is designed to maximize the data likelihood on a single pixel; finally CCA uses a second order statistics (spatial correlation) of subsets of the data to infer the spectral behavior of the mixed components. Let's now have a closer look at these three techniques.

3.2 FASTICA

It is an application of the Independent Component Analysis (ICA), a general technique broadly used in signal processing. It was originally introduced to deal with problems closely related to the so called *cocktail party problem*, where a linear mixture of components has to be separated into the original sources. As it will be clear soon, ICA is a totally blind approach.

To introduce the technique, let's take into account Eq. (3.1) and work for simplicity in a space where the number of the components and number of frequency channels are both equal to 2. Moreover let's drop the noise term for a while. Suppose then to measure a linear combination m_1 of two components c_1 and c_2, $m_1 = a_{11}c_1 + a_{12}c_2$. It is quite obvious that it is not possible to find the original components, if the coefficient a_{1j} are unknown. Thus, it may be surprising that,

under some assumptions, a solution can be found when another, independent measure m_2 is added to the problem (here independent means that the weights have to be different). Then the system:

$$\begin{aligned} m_1 &= a_{11}c_1 + a_{12}c_2 \\ m_2 &= a_{21}c_1 + a_{22}c_2 \end{aligned} \tag{3.1}$$

can be solved without knowing neither the weights a_{ij} nor the components c_j. The ICA approach allows to solve this problem using the statistical properties of the components. Indeed, it turns out that it is enough to assume that the quantities c_j are statistically independent and that at most one of them is Gaussian. These assumptions are not unrealistic in many practical cases and mostly important for our purposes it can be safely assumed that the CMB is statistically independent from foregrounds and that, as we know from observations, the latter are highly non-Gaussian.

3.2.1 The basic idea

Non-Gaussianity is the key point to understand ICA (Hyvärinen, 1999). The Central Limit Theorem states that, given some assumptions, the distribution of a sum of independent random variables tends to be Gaussian. Thus, a sum of two (or more) independent random variables has a distribution that is closer to Gaussian than any of the original random variables.

Let's consider now a linear combination of the mixtures $\mathbf{m}$, through a generic vector $\mathbf{k}$

$$y = \mathbf{k^T m}. \tag{3.0}$$

According to the central limit theorem this variable y has to be more Gaussian than any m_i. But using Eq. (3.1), y is a linear combination of the components too:

$$y = \mathbf{k^T A c} = \mathbf{z^T c}, \tag{3.0}$$

where $\mathbf{z}$ simply represents a change of variables, $\mathbf{z} = \mathbf{A^T k}$. The quantity y is then more Gaussian than any c_j. Now, since there is full freedom on the choice of the coefficients $\mathbf{k^T}$, one can pick up that combination that minimizes the Gaussianity of y. Such a choice will set one and only one of the z_j different from zero, because any other combination of z_j^T would sum more than one component c_j causing the Gaussianity of y to depart from the minimum. In other words, by minimizing the Gaussianity of y, one selects out one of the components, $y = c_j$. In a case where the dimension of the space is larger than 2, once the first component has been found it is sufficient to go to the space orthogonal to that component and perform the same procedure again.

It is now clear why I put ICA in the first group of the component separation techniques, as introduced at the beginning of this Chapter: ICA infers the statistical properties of the signals by using all the available pixels in the map at the same time (of course this do not prevent us to perform the separation on a subset of data).

3.2.2 Measure of non-Gaussianity

To apply the ICA principle, we need a quantitative measure of the non-Gaussianity of a random variable, say y.

Kurtosis

Historically, the classical measure of non-Gaussianity is the kurtosis, also known as the (normalized) fourth order momentum of a distribution:

$$kurt(y) = E\{y^4\}. \tag{3.0}$$

Such a quantity equals to zero for a standard Gaussian variable. Typically non-Gaussianity is measured by the absolute value of the kurtosis and has the advantage of being very fast to compute. There are of course non-Gaussian random variables that have zero kurtosis, but they are very few and not relevant in the present context (Hyvärinen, 1999).

Unfortunately, kurtosis has some practical drawbacks, when its value has to be computed from a measured sample. The main problem is that the kurtosis is very sensitive to outliers (Huber, 1985), making it a rather poor estimator of non-Gaussianity in many cases.

Neg-entropy

Entropy is an important concept in information theory. It quantifies the degree of information that an observation of a variable can provide. The more a variable is random and unpredictable, the larger its entropy is. For a random vector y with density $f(y)$, the entropy H is defined as:

$$H(y) = -\int f(y)\, log(f(y))\, dy. \tag{3.0}$$

It can be demonstrated that a Gaussian variable has the largest entropy among all the random variables with equal variance (see e.g. Cover and Thomas , 2006). This means that entropy could be used as a measure of non-Gaussianity. It is indeed small for distributions that are concentrated on certain values, i.e., when the variable is clearly clustered, or has a probability distribution function that is very ”spiky”. To obtain a measure of non-Gaussianity that is zero for a Gaussian variable and always non-negative, it is useful to introduce a slightly modified version of the definition of differential entropy, called neg-entropy. Neg-Entropy J is defined as:

$$J(y) = H(y_{gauss}) - H(y) \tag{3.0}$$

where y_{gauss} is a Gaussian random variable with the same covariance matrix as y.

The advantage of using the entropy, or equivalently the neg-entropy, for estimating non-Gaussianity is that this choice is fully supported by information theory, being the entropy the optimal estimator of the non-Gaussianity. On the other hand, it results computationally demanding to compute the entropy of a variable in practical situations, since it requires a non-parametric estimation of the probability distribution function. Therefore it is necessary to introduce approximations of the neg-entropy, that maintains the robustness of the estimator while being faster to compute. These approximations are based on the maximum-entropy principle. In general we obtain the following approximation:

$$J(y) \approx \{E[G(y)] - E[G(\nu)]\}^2 \tag{3.0}$$

where G is some non-quadratic function and ν is a Gaussian variable. The point here is to choose wisely the function G. In particular, very robust estimator are obtained when G does not grow too fast. These two choices proved to be very useful for our purposes:

$$G_1(u) = \frac{1}{a_1} log\ cosh\ a_1 u, \tag{3.0}$$

$$G_2(u) = -exp(u^2/2), \tag{3.0}$$

where $1 \leq a_1 \leq 2$ is a suitable constant. These functions are used in the FASTICA algorithm as contrast functions, as I will show in the next sections.

3.2.3 The algorithm

The Independent Component Analysis was initially introduced in astrophysics by Baccigalupi et al. (2000), exploiting an adaptive technique for learning progressively the independent components in a given multi-frequency mixture by considering one realization (pixel) after the other. However, most applications nowadays in the context of CMB data analysis are based on the FASTICA algorithm (Hyvärinen & Oja, 2000) which provides an efficient and computationally cheap method for the cleaning of the CMB maps, using the whole statistics at once.

At the core of the FASTICA algorithm lies an assumption that the statistical independence of the components can be measured via their non-Gaussianity [1]. The component estimates are therefore sought for as linear combinations of the input data, which maximize the selected measure of the non-Gaussianity. In the FASTICA approach, the neg-entropy is used for this purpose, and a number of computationally efficient approximations have been proposed (Hyvärinen, 1999; Hyvärinen & Oja, 2000); later, they were found to perform very well in the context of the microwave sky maps, in total intensity (Maino et al., 2002) and polarization (Baccigalupi et al., 2004).

[1] We anticipate here that this assumption is not fully respected in CMB analysis since the foregrounds are expected to be correlated within each other. We will be back on this point later in this Chapter.

The FASTICA separation process is done in three steps, which we describe briefly below.

I - Pre-processing

Our assumed input data are a set of multi-frequency maps of the sky in a microwave band. Each map is discretized and is composed of n_{pix} pixels for which the relevant amplitudes, corresponding to some or all Stokes parameters, I, Q & U are recorded. For each sky pixel, p, and Stokes parameter, we model the data as,

$$\mathbf{m_p} = \mathbf{A c_p} + \mathbf{n_p}, \tag{3.0}$$

where all the vectors, $\mathbf{m_p}$, $\mathbf{c_p}$ and $\mathbf{n_p}$ (i.e., measured maps, sky components, and noise, respectively) have as many elements as the number of available maps, ($\equiv n_m$) and which either correspond to different frequency channels or Stokes parameters, and have to be analyzed separately. Hereafter, we always assume that a single mixing matrix exists for all considered pixels. Whenever this is not the case, it may be possible to limit the analysis to subsets of pixel selected in a way that the assumption of identical mixing matrix holds for each subset separately and proceed with the analysis as described below, accepting the loss of precision incurred due to a smaller number of the data points used in each of the ICA analysis.

From the data, $\mathbf{m}$, we first estimate the data correlation matrix

$$\mathbf{\Sigma} \equiv \left\langle \mathbf{m}\, \mathbf{m^t} \right\rangle, \tag{3.0}$$

where $\langle \dots \rangle$ denotes the average over all the map pixels. Then, given the known (by assumption) noise properties, $\mathbf{N}$, we get the signal correlations for the input maps as,

$$\mathbf{C} \equiv \mathbf{\Sigma} - \mathbf{N}. \tag{3.0}$$

Given the derived estimates of the signal correlations, the FASTICA algorithm subsequently pre-whitens the signals in the input data. That is done by computing a "square root" of the signals correlation matrix, $\mathbf{C^{1/2}}$, such as,

$$\mathbf{C} \equiv \mathbf{C^{1/2}}\, \mathbf{C^{1/2}}. \tag{3.0}$$

The input data, $\mathbf{m}$, are then replaced by their *signal* prewhitened and decorrelated version, $\mathbf{w}$,

$$\mathbf{w} \equiv \mathbf{C^{-1/2}}\, \mathbf{m}. \tag{3.0}$$

Indeed, we have,

$$\left\langle \mathbf{w}\, \mathbf{w^t} \right\rangle = \mathbf{1} + \mathbf{C^{-1/2}}\, \mathbf{N} \left(\mathbf{C^{-1/2}} \right)^{\mathbf{t}}. \tag{3.0}$$

Note that to perform this step, the signal correlation matrix, $\mathbf{C}$, has to be non-singular and positive definite. In the standard FASTICA implementation that requires, for example, that the number of the input maps, n_m, is equal to a number of the independent sky components, n_c, which are to be recovered. If that is not the case and that is not due to the noisiness of the signal correlation estimator we need to find another way to proceed. We describe a relevant approach later.

II - Mixing matrix determination

Assuming the pre-whitened input data, the FASTICA algorithm proceeds to estimating the mixing matrix, **B**, of the prewhitened data,

$$\mathbf{w} = \mathbf{B}\,\mathbf{c}, \tag{3.0}$$

and

$$\langle \mathbf{w}\,\mathbf{w}^{\mathbf{t}} \rangle = \mathbf{B}\,\langle \mathbf{c}\,\mathbf{c}^{\mathbf{t}} \rangle\,\mathbf{B}^{\mathbf{t}} = \mathbf{B}\,\mathbf{B}^{\mathbf{t}} = \mathbf{1}. \tag{3.0}$$

And therefore, the matrix **B** has to be orthonormal, if only we fix the arbitrary normalization of the sought-after components to unity.

The algorithm determines the matrix **B**, column-by-column, and is referred to as one-unit, fixed-point algorithm (Hyvärinen & Oja, 2000). Each column is found, as a vector maximizing the selected neg-entropy definitions and orthogonal to all the columns already determined earlier. The maximization of the neg-entropy is reached iteratively, as sketched in the following table:

Algorithm 1 BASIC FASTICA ALGORITHM

for *every pre-whitened data set j*
do
– *pick a random unit vector* $\mathbf{b_j}$*;*
while *the last correction to* $\mathbf{b_j}$ *not sufficiently small* ***do***
– *orthogonalize it with respect to all previously processed vectors* $\mathbf{b_i}$ *where* $i < j$*;*
– *find a correction to* $\mathbf{b_j}$ *so the new* $\mathbf{b_j}$ *is more non-Gaussian;*
end while
end for

The mixing matrix, **B**, determined in this manner, is by construction orthonormal.

III - Post-processing

Given the recovered mixing matrix the components are found applying the transpose of the matrix to the pre-whitened data, i.e.,

$$\mathbf{c} = \mathbf{B}^{\mathbf{t}}\,\mathbf{w}. \tag{3.0}$$

In general the amplitudes of the components recovered using the FASTICA approach are arbitrary and set to unity in Eq. (3.2.3). This freedom is due to the degeneracy between the values of the mixing matrix parameters and the component normalizations. Maino et al. (2002) resolves that problem by noticing that the freedom in the normalization essentially corresponds to the choice of the physical units for the component signal. For example, setting to the unity all the mixing matrix elements corresponding to one of the input frequency maps is equivalent to choosing the component map units as those of the input map of the selected frequency. Given the definition of the mixing matrix, Eq. (3.1), Maino et al. (2002)

normalization procedure requires that every element of the first row of **A** has to be equal to unity, i.e.,

$$\mathbf{A_{0,i}} = \mathbf{1}, \quad \text{for every } i = 0, \ldots, n_s - 1, \tag{3.0}$$

where given Eqs. (3.2.3) & (3.2.3) we have,

$$\mathbf{A} = \mathbf{C^{-1/2}}\,\mathbf{B}. \tag{3.0}$$

3.2.4 Achievements

Since its introduction in astrophysics, FASTICA has been successfully applied both on simulated and real CMB data sets.

Tests on simulated observations of the microwave sky were performed with nominal specifications of the Planck satellite both on total intensity (Maino et al., 2002) and polarization (Baccigalupi et al., 2004). They considered several observation channels containing the most important known diffuse signals: CMB, synchrotron, dust and free-free emissions (the latter was actually absent in the polarized simulated sky). The algorithm proved to be reliable and efficient in those conditions. In particular, the CMB angular power spectra was recovered at the percent level of accuracy up to $\ell \simeq 2000$ for total intensity, up to $\ell \simeq 1000$ for the E modes and up to $\ell \simeq 1200$ for the cross correlation TE. Simulations for the recovery of the B modes were performed by Stivoli et al. (2006) and will be treated in the next Chapter.

Beside the good results on simulations, FASTICA was able to deal with real data also. The first application was on the *COBE-DMR* 4yr data by Maino et al. (2003). Despite the low signal to noise ratio ($S/N \sim 1$) and low number of pixels available, the approach was able to extract the CMB signal with high confidence and to detect the foreground emission. The results were in excellent agreement with previous ones in terms of frequency scaling, foreground morphology and CMB power spectrum.

Another application of the code was performed by Donzelli et al. (2006) on the Background Emission Anisotropy Scanning Telescope (*BEAST*) data. After a detailed calibration procedure, they extracted the CMB power spectrum and found very good agreement with the one found by the *BEAST* collaboration.

Finally, Maino et al. (2007) performed the analysis of the WMAP 3rd year data. In that work we achieved two important results. First the CMB power spectrum was recovered up to degree scale. In Fig.3.1 (left panel) the black solid line is the WMAP 3rd year best fit while the different colored represent the different channel combinations used in the analysis. Here the Kp2 mask has been applied before to run FASTICA . The agreement is evident both on the low multipoles as well as to higher ones up to $\ell \simeq 150$. Large spread in the results is present for the highest bins.

Adopting the consistency with the WMAP power spectrum as a figure of merit, we can judge which is the optimal combination: this is the KQVW, indicating that the best tracer of low frequency foreground contamination at high frequencies is represented by the K band data. On the right panel of Fig.3.2 I show similar results from the full sky analysis (the Kp2 mask is here applied only before computing the power spectrum): the agreement with WMAP results are even more evident for both low and high ℓ. This is an indication of the fact that even in presence of strong and possibly correlated foreground on the Galactic plane, FASTICA is not only still able to properly recover the CMB pattern at high Galactic latitudes but it performs better than in the case of a pure Kp2 analysis. This means that the level of signal correlation along the Galactic plane, that violates one of the ICA assumptions, does not compromise the reconstruction at high Galactic latitudes and that the regions near the plane included in the full sky analysis are useful for better distinguish different signal statistics.

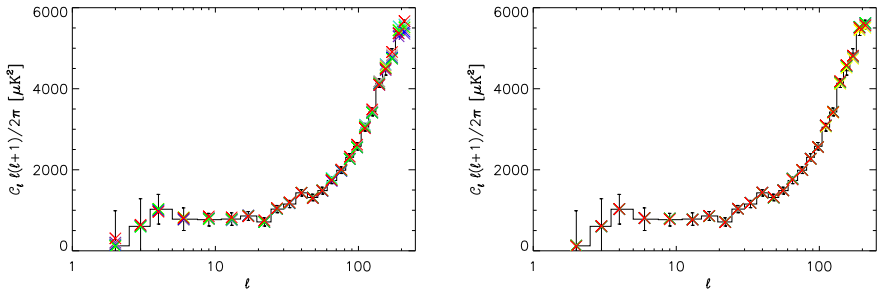

Figure 3.1: The binned power spectra from all WMAP channel combinations compared with the WMAP 3yr power spectrum for a complete Kp2 analysis (upper panel) and full sky analysis with Kp2 mask applied when computing the power spectrum (lower panel). Channel combinations are color coded: red for KKaQVW, green for QVW, blue for KQVW, yellow for KaQVW, brown for KKaQV, indigo for KaQV, turquoise for KQV and grey for KKaVW. From Maino et al. (2007)

The second main result of Maino et al. (2007), already noticed independently by several authors (see for example Hansen et al., 2006, and references therein) is the fact that FASTICA did confirm the north-south sky asymmetry. Analyzing the CMB power spectrum in the two hemisphere, we found the asymmetry shown in Fig.3.2 , left panel. In a pure FASTICA analysis, one might be tempted to try an explanation in terms of a difference in the overall foreground spectral indeces in the two hemispheres, since FASTICA assumes an uniform frequency scaling across the whole sky; but the fact that we obtained the same result with totally independent procedures, and most importantly the following test, make this explanation unlikely. Indeed, we performed the component separation on the northern and southern hemisphere separately and derived the CMB power spectrum separately for each of them. The spectra are reported in the right panel of Fig.3.2. Also

in this case the northern spectrum is systematically lower than the southern one. This result strongly disfavors an explanation based on foregrounds for the asymmetry found in the FASTICA CMB maps. In addition, we point out the remarkable agreement between our results and those of Hansen et al. (2006) (see their Fig.8), obtained with a completely independent technique.
On the FASTICA side, this confirms the reliability of the algorithm when exploited to reconstruct the finest structure in the CMB pattern out of a given dataset; on a purely scientific side, we confirm the existence of a marked asymmetry in the CMB anisotropy power between north and south in Galactic coordinates, which at the moment escapes explanations in terms of difference in the foreground properties on the corresponding two hemispheres.

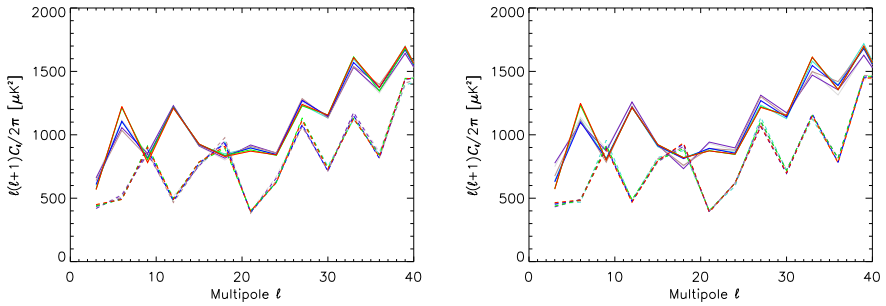

Figure 3.2: The power spectra derived on the northern (dashed line) and southern (solid line) hemisphere in the reference frame with north pole at $(\theta, \phi) = (80°, 57°)$. The left panel shows the spectra derived from the ICA CMB map out of Kp2. In the right panel the CMB maps are obtained applying FASTICA separately on the two hemispheres. Colors mark channel combinations as in previous figure. From Maino et al. (2007)

3.2.5 Extensions of the algorithm

The achieved results have been shown to be at least *on par* with, and often better than, those obtained with the other methods. That has been motivating a further exploration of the potential of the algorithm and its applicability in a progressively more realistic and demanding circumstances . This task has been, however, made significantly more difficult by a number of limitations characterizing the existing implementations of the algorithm and which had to be resolved by *ad hoc* means devised on a case-by-case basis (Maino et al., 2007).

Stompor et al. (2008a) addressed two issues that have been outstanding in that respect. The first one relates to a proper treatment of the redundant noisy data provided as inputs. In fact the original FASTICA algorithm was developed with mostly noiseless data in mind and required the user to reduce the redundant input data to an independent set, i.e. such that the number of input data sets is equal to the number of the anticipated independent components. In other words in the

original version of the code, the mixing matrix A was supposed to be square and extra measurement had to be discarded. In realistic cases with the noisy data, such an approach may unnecessarily lower the quality of the final results. The second issue is related to a lack of a framework for incorporating priors in the method. Though the priors may be seen as a departure from the spirit of the method, i.e. its "blindness", however, striving for the best possible estimator, it is interesting to investigate if some priors can be used consistently in the algorithm. In this context, the inclusion of the CMB frequency scaling as a prior seems to be of a particular interest, and the one we focus on hereafter. In relation to this last aspect, we point out that in the process of the preparation of the manuscript of the Stompor et al. (2008a) paper, a preprint by Vio & Andreani (2008) has been published, addressing this issue in a way which is similar to ours, and obtaining consistent results.

Lossless input data compression

As introduced in Stompor et al. (2008a), I'll describe now the generalization of the pre-whitening procedure, which properly treats the noisy, redundant data sets, preserving all the available information. The proposed procedure combines the standard data compression and Principle Component Analysis techniques. Our goal is to find a linear combination of the input maps which preserves all the information concerning the sky signal, **s**, but which consists of the independent data.

We use the eigenvalue decomposition to define the null space of the signal correlation matrix, $\mathbf{C} = \langle \mathbf{s}\,\mathbf{s^t} \rangle$. Let $\mathbf{U}$ be a matrix rows of which, denoted $\mathbf{u_i}$, correspond to eigenvectors of $\mathbf{C}$ and $\mathrm{diag}\,(\epsilon_i)$ – a diagonal matrix of the corresponding eigenvalues, i.e.,

$$\mathbf{C} = \mathbf{U}\,\mathrm{diag}\,(\epsilon_\mathbf{i})\;\mathbf{U^t}. \tag{3.0}$$

The null space is spanned by a subset of the $\mathbf{u_i}$ vectors, which we will call $\mathbf{V}$.

We also use $\mathbf{W}$ to denote the remaining vectors; so we have,

$$\mathbf{U^t} \equiv \left[\mathbf{W^t},\ \mathbf{V^t}\right]. \tag{3.0}$$

Any signal maps, **s**, in the n_m-dimensional space can be found only up to an unknown arbitrary vector from the null space. Therefore, to make our solution unique we can impose a set of additional constraints. For example, we can look for solutions which have no component in the null space, i.e.,

$$\mathbf{V}\,\mathbf{s} = \mathbf{0}. \tag{3.0}$$

We then have

$$\mathbf{s} = \widehat{\mathbf{W}}^\mathbf{t}\,\mathbf{y}, \tag{3.0}$$

where **y** is a vector of the length, n_{nz}, which defines the number of the independent components $n_c = n_{nz}$ and we have introduced $\widehat{\mathbf{W}} = \mathrm{diag}\left(\epsilon_\mathbf{i}^{-\mathbf{1/2}}\right)\mathbf{W}$ (the only non-zero ϵ value are included here), making use of the arbitrariness in the data normalization. Then we can plug that into Eq. (3.2.3) to obtain,

$$\mathbf{m} = \widehat{\mathbf{W}}^\mathbf{t}\,\mathbf{y} + \mathbf{n}. \tag{3.0}$$

A minimum variance (and maximum likelihood) estimate, $\hat{\mathbf{y}}$, of $\mathbf{y}$ is then given by,

$$\hat{\mathbf{y}} = \left(\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\widehat{\mathbf{W}}^{\mathbf{t}}\right)^{-1}\,\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\mathbf{m}. \tag{3.0}$$

$\hat{\mathbf{y}}$ is the new, by construction, non-degenerate data set to be processed by the current FASTICA algorithm with the noise correlations respectively given by,

$$\mathbf{N}_{\hat{\mathbf{y}}} = \left(\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\widehat{\mathbf{W}}^{\mathbf{t}}\right)^{-1}. \tag{3.0}$$

Note that the correlations of the noise between different maps already on the input are also introduced in the standard FASTICA algorithm due to the pre-whitening procedure and therefore are not a new feature of the proposed here approach.

The signal correlation matrix for the new data, $\widehat{\mathbf{y}}$, is given by,

$$\begin{aligned}\mathbf{C}_{\widehat{\mathbf{y}}} = \langle\widehat{\mathbf{y}}\,\widehat{\mathbf{y}}^{\mathbf{t}}\rangle \;&=\; \left(\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\widehat{\mathbf{W}}^{\mathbf{t}}\right)^{-1}\,\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\langle\mathbf{m}\,\mathbf{m}^{\mathbf{t}}\rangle\,\times \\ &\times\;\; \mathbf{N}^{-1}\,\widehat{\mathbf{W}}^{\mathbf{t}}\left(\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\widehat{\mathbf{W}}^{\mathbf{t}}\right)^{-1} = \mathbf{1}.\end{aligned} \tag{3.0}$$

We thus see that the new data set, thanks to the rescaling applied in Eq. (3.2.5), does not require any more any pre-whitening stage.

Now, in the case with the input data compression, the normalization and frequency scalings are determined in a way analogous to the one used in the standard FASTICA technique. This time, however, one needs to account for the fact that the maps on the input of the ICA procedure will, in general, not correspond now to any particular frequency but will be a linear combination of the single frequency maps as defined in Eq. (3.2.5), i.e.,

$$\hat{\mathbf{y}} = \mathbf{B}\,\mathbf{c}, \tag{3.0}$$

where the mixing matrix, $\mathbf{B}$, is estimated on the ICA routine output.

From Eqs. (3.2.5) & (3.2.3), we can obtain,

$$\hat{\mathbf{y}} = \mathbf{y} + \left(\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\widehat{\mathbf{W}}^{\mathbf{t}}\right)^{-1}\,\widehat{\mathbf{W}}\,\mathbf{N}^{-1}\,\mathbf{n}, \tag{3.0}$$

just showing that our independent data estimator in Eq. (3.2.5) is unbiased, i.e., in the absence of the noise, as assumed hereafter, we have $\hat{\mathbf{y}} = \mathbf{y}$. Given this and with help of Eq. (3.2.5), we further have,

$$\mathbf{s} = \widehat{\mathbf{W}}^{\mathbf{t}}\,\hat{\mathbf{y}} = \widehat{\mathbf{W}}^{\mathbf{t}}\,\mathbf{B}\,\mathbf{c}, \tag{3.0}$$

and hence

$$\mathbf{A} = \widehat{\mathbf{W}}^{\mathbf{t}}\,\mathbf{B}. \tag{3.0}$$

Our normalization procedure requires then,

$$\mathbf{c'}_{\mathbf{i}} = \mathbf{A}_{\mathbf{0i}}\,\mathbf{c}_{\mathbf{i}} = \left[\sum_{\mathbf{j}}\,\widehat{\mathbf{W}}_{\mathbf{j0}}\,\mathbf{B}_{\mathbf{ji}}\right]\,\mathbf{c}_{\mathbf{i}}, \tag{3.0}$$

where prime marks a component normalized to the same units as the input map corresponding to the 0 -th frequency. This generalizes the approach of Maino et al. (2002), which corresponds to the choice $\widehat{\mathbf{W}} = \mathbf{P^{-1}}$.

Similarly, the frequency scalings of the recovered components is now given by the generalized expression for the mixing matrix, **A**, Eq. (3.2.5). As before this reduces to the standard form whenever $\widehat{\mathbf{W}}$ is a square unit matrix.

CMB prior

The assumption is that the CMB is an "independent" component in the ICA sense. By applying the CMB prior we want to ensure that the scaling of the derived CMB component follows the *a priori* known law. Consequently, one of the columns of the mixing matrix, **A**, which is to be computed via the ICA algorithm needs to comply with that scaling law up to, at least, an overall, and irrelevant, constant factor.

The columns of the matrix, **A**, are uniquely related to the columns of the "compressed" mixing matrix, **B**. The relevant relation reads as follows,

$$\mathbf{b} = \widehat{\mathbf{W}}\,\mathbf{a}. \tag{3.0}$$

Here **a** and **b** denote the corresponding columns of the matrices, **A**, and **B**, respectively. Imposing the CMB prior, i.e., setting one of the columns of the matrix, **A**, to reflect the CMB frequency scaling, $\mathbf{a} = \mathbf{a^{CMB}}$, can be then formulated as fixing the relevant column of the **B** matrix to $\mathbf{b^{CMB}}$, where $\mathbf{b^{CMB}}$ and $\mathbf{a^{CMB}}$ are related as in Eq. (3.2.5).

Subsequently, we need to construct the full mixing matrix, **B**, one column of which is known ahead of the time. The construction is based on the fact that the sought-after matrix is orthonormal and therefore all the remaining columns have to span the subspace orthogonal to the direction defined by the assumed prior. Our amended ICA algorithm needs therefore to look for the matrix columns which simultaneously maximize our measure of non-Gaussianity and fulfill the latter requirement. Once all but one columns of **B** are determined, the last one is bound to coincide with the imposed prior.

The blue print of a simple numerical implementation is outlined below. Here we just note that if only the CMB map is required, we in fact do not need to proceed any further as the CMB map of the sky will be readily given by,

$$\mathbf{c^{CMB}} = \mathbf{b^{CMB^t}}\,\hat{\mathbf{y}}, \tag{3.0}$$

with the normalization expressed by Eq (3.2.5). This can become a basis to an extremely simple[2], non-iterative and therefore fast CMB cleaning algorithm. Note

[2]Note that if the input data set is free of any redundancies, then the Eq. (3.2.5) is equivalent to multiplying the prior, $\mathbf{a^{CMB}}$ by the input data pixel-by-pixel, i.e., $\mathbf{c^{CMB}} = \mathbf{a^{CMB}}\,\mathbf{m}$. Generally, this, however, is not true as the redundant noise of the input data contained in the null space of the matrix, **C**, has to be projected out before performing the dot product.

that in spite of its striking simplicity this algorithm retains all the pros and cons of the full FASTICA algorithm as far as the determination of the CMB component is concerned. However, as fewer sources of the uncertainty are involved in the simplified algorithm one may expect that a simple error estimation of the final cleaned CMB map based on the linear error propagation can provide its sufficiently precise estimates in this case. We discuss the relevant issues later in this Chapter.

Note that if indeed the compression procedure is done properly the scaling of the recovered component will be as imposed by the prior. That can be seen as follows. First, note that the frequency scaling of the recovered component is given by the relevant column, **a**, of the matrix **A** and thus can be expressed as,

$$\mathbf{a} = \widehat{\mathbf{W}}^{\mathbf{t}}\,\mathbf{b}^{\mathbf{CMB}} + \mathbf{B}\,\mathbf{V} = \widehat{\mathbf{W}}^{\mathbf{t}}\,\widehat{\mathbf{W}}\,\mathbf{a}^{\mathbf{CMB}} + \mathbf{V}\,\mathbf{d}, \tag{3.0}$$

where the second term results from the zero kernel of the "compression" matrix, **U**, and **d** is a vector storing arbitrary coefficients. However,

$$\mathbf{A}\,\mathbf{V} = \mathbf{a}^{\mathbf{CMB}}\,\mathbf{V} = \mathbf{0}, \tag{3.0}$$

because the singular vectors are determined for $\mathbf{A}^{\mathbf{t}}\,\mathbf{A}$ matrix and thus must be orthogonal to all columns of the matrix **A**, including the one corresponding to the CMB components. Given that, the coefficient vector, **d**, has to be zero and,

$$\mathbf{a} = \widehat{\mathbf{W}}^{\mathbf{t}}\,\widehat{\mathbf{W}}\,\mathbf{a}^{\mathbf{CMB}} = \mathbf{U}^{\mathbf{t}}\,\mathbf{U}\,\mathbf{a}^{\mathbf{CMB}} = \mathbf{a}_{\mathbf{CMB}}, \tag{3.0}$$

as **U** is a full-rank orthonormal matrix as defined in Eq. (3.2.5).

Note that in practice the rejected eigen-modes may only be approximately singular and thus some minor deviations of the posterior scaling from the prior can be expected.

In principle that could work with any component for which the scaling is assumed to be known. However, the single out component has to be known, on physical grounds, to be indeed independent. In particular, whenever we have multiple components with the scaling determined a priori, they all need to be mutually independent.

3.2.6 Pre-launch Planck simulations

One of the first application of the extended FASTICA algorithm has been done on a Planck total intensity simulated data set. As introduced in the last Chapter, Planck is an experiment designed to surpass previous CMB experiments in almost every aspect. Therefore, a complete and timely exploitation of the data will require methods that improve upon foreground removal via template subtraction and masking. The development and assessment of such methods is coordinated within the Planck 'Component Separation Working Group' (WG2).

Leach et al. (2008) reported the results of the WG2 activity in the framework of a *component separation challenge* using a common set of simulated Planck data.

The objective of such a challenge was to assess the readiness of the Planck collaboration to tackle component separation, based on the analysis of realistically complex simulations. It offered an opportunity for comparing the results from different methods and groups, as well as to develop the expertise, codes, organization and infrastructure necessary for this task.

Sky simulations were based on an early development version of the Planck Sky Model, as explained in the last Chapter. This data sets were complemented by a set of ancillary data including hit maps and noise levels, IRAS, 408 MHz, and Hα templates, as well as catalogues of known clusters and point sources (see Leach et al., 2008, for more details).

A bunch of methods have been applied to this data set to test the diffuse component separation. Table 3.1 reports the complete list of the methods with their main characteristics. Although we report here some of the results from all the exploited algorithms, we do not describe them. For details, see the references in the table. In the remaining part of this Section we outline the main FASTICA results in this

	Channels used	Components modeled	CPU time/runtime
CCA Bedini et al. (2005)	Planck, Haslam 408 MHz	CMB, dust, sync, FF	70 hr, 1.5 day
COMMANDER Eriksen et al. (2008)	WMAP, Planck 30–353 GHz,	CMB, dust, sync, FF, monopoles, dipoles	1000 hr, 2 day
FastICA Maino et al. (2002)	143–353 GHz	none	21 min, 20 sec
FastMEM Hobson et al. (1998)	Planck	CMB, SZ, dust, sync, FF	256 hr, 8 hr
GMCA Bobin et al. (2007)	Planck	CMB, SZ, sync., FF	1200 hr, 6 day
SEVEM Martínez et al. (2007)	Planck	CMB	30 hr, 30 hr
SMICA Delabrouille et al. (2003)	Planck, WMAP	CMB, SZ, dust, total galaxy	8 hr, 4 hr
WI-FIT Hansen et al. (2006)	70–217 GHz	CMB	400 hr, 8 hr

Table 3.1: A summary of methods, dataset, components and computational resources used in the Planck simulations. From Leach et al. (2008)

particular application.

In Fig.3.3, a quantitative measure of the raw residual CMB map is shown, provided by its RMS, calculated for 18 zonal bands, each of them 10 degrees wide in Galactic latitude. It is clear the FASTICA is performing worse than the other methods close to the Galactic center (see next Section for a discussion about this issue). Nevertheless, the algorithm is able to provide the lowest residual among all the codes at high Galactic latitudes.

FASTICA was applied to a limited range of frequencies where the diffuse emission is mostly represented by CMB and dust. The algorithm was particularly effective in recovering the dust emission at high frequencies and outside the Galactic plane. In Fig.3.4 I report a dust patch at 143 GHz as recovered by FASTICA (Leach et al., 2008). Although a proper characterization of the faithfulness of the

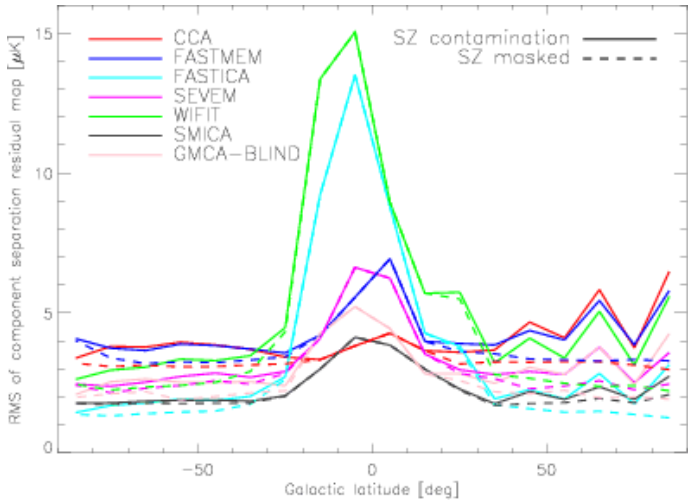

Figure 3.3: RMS of the residual error of the CMB map, calculated at 45'. For comparison $\sigma_{CMB} = 69.8\mu K$ and $\sigma_{noise} = 0.7\mu K$ at 143 GHz. From Leach et al. (2008)

reconstruction was not done, the visual inspection reveals a good recover.

3.2.7 Remarks

Up to date the study of the capabilities of FASTICA is not finished and the algorithm still have some evident limitations that I'm going to describe in this Section.

The first issue is represented by the fact that in a typical application, the algorithm is asked to break one of its fundamental assumptions. Indeed, the ICA approach assumes that the components to be recovered are independent, while Galactic foregrounds are expected to be correlated with each other, as it may be understood from what we discussed in the previous Chapter. All the tests made so far seem to show that FASTICA fails in separating the foregrounds among them, while it is still able to recover the CMB component that is independent on the Galactic emission (Maino et al., 2002). Then, in practical situations in which more than one foreground is known to be in the data, FASTICA can be seen as a foreground cleaning algorithm, rather than a component separation one.

Another intrinsic limitation is represented by the assumption of a single mixing matrix in the data. This is clearly an approximation, since we expect the foreground properties to be different across the sky. For example, we saw in the last Chapter that synchrotron and thermal dust emissions have different spectral indeces in different position in the sky. This is not a big issue for small patch experiments, where the spatial variation are expected to be small and possibly negligible, but it becomes important when one has to deal with all-sky surveys. At first order it is possible to ignore this fact and force the code to look for a single mixing matrix, but it has been demonstrated (Baccigalupi et al., 2004) that in this case the separation is worse in

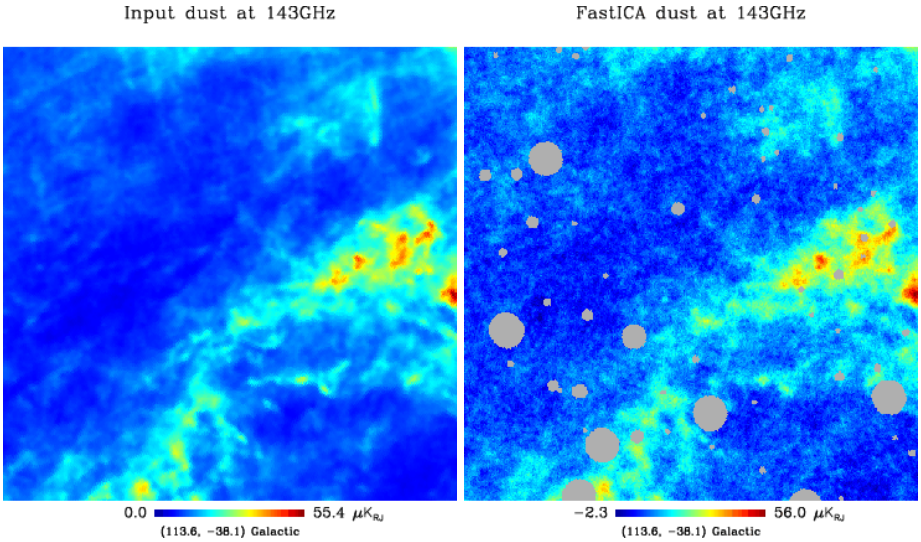

Figure 3.4: Comparison between input and output dust patch as recovered by FASTICA at 143 GHz, from Leach et al. (2008).

terms of foreground residuals in the CMB recovered map. One possible solution would be to cut the sky in different patches (given any possible driver to choose them) and perform separation separately on each of them. An extensive study on this possibility has not been done yet, but it is easy to foresee a couple of potential drawbacks here. First of all, given the smaller amount of available pixels in each separation, the statistics of the components would be harder to infer and the separation itself could be harder (this is probably not a big issue for the analysis on the Planck data, given the huge number of available pixels). Second, it is not trivial to put back together all the small separated maps, without incurring in sudden jumps at the borders.

Finally, there is a third issue that is becoming more and more important inside the CMB community: error assessment. The non-linear nature of the FASTICA algorithm makes it difficult to propagate fully consistently the noise and reconstruction errors from the input all the way to the maps of the separated components. In the past the problem have been addressed with help of numerical experiments and simulations (Stivoli et al., 2006; Maino et al., 2007). These can be considered as first attempts toward a better understanding of the error propagation within the FASTICA process. Of course, the final goal would be an analytic treatment of the problem, that would allow to properly propagate the errors to the next steps of CMB data analysis pipeline.

3.3 Parametric Approach

The defining premise of a parametric approach to the component separation is the assumption that the functional form of the frequency scaling for all involve components is known, and our ignorance can be quantified by means of relatively few, though non-linear, spectral parameters (see for example Stompor et al., 2008b, and reference therein). The approach consists in modeling the behavior of all the components in the mixing matrix and trying to recover the parameters of such a model via a fitting procedure, that in principle can be performed on a pixel-by-pixel basis. This aspect puts the parametric approach in the second group ("one item") of our category. The attractiveness of this approach lies in its simplicity, the flexibility of the parameterization schemes and possibility of phrasing the separation problem in a coherent maximum likelihood form (Jaffe et al., 2004). Its strength lies in exploiting nearly optimally the prior knowledge of the frequency scaling of the components, while ignoring the information which is less known and more debatable. Its weakness is related to the difficulties in performing the non-linear and high-dimensional parameter fits especially in a low signal-to-noise ratio regime on the pixel, and numerical efficiency, given the large number of pixels, $O(10^7)$, anticipated from the next generation of the CMB experiments.

In a recent paper, Eriksen et al. (2006) reconsidered a parametric approach to the component separation task originally introduced by Brandt et al. (1994) and proposed a method which was shown to be both numerically efficient and stable in the application to the nearly full-sky real (WMAP) and simulated (Planck) data. In that method, the numerical efficacy and stability is achieved by splitting the parameter and component map estimations into two separate steps. The spectral parameter fitting is performed pixel-by-pixel on the input data which are first smoothed and underpixelized. This preprocessing stage is found to be necessary because of the high noise in the full resolution pixels. It also makes the procedure more numerically tractable as fewer fits have to be done. The fitting criterion, employed by Eriksen et al. (2006), uses a likelihood function (but not a maximum likelihood as we will see later), and involves at the same time the (non-linear) spectral parameters as well as (linear) component amplitudes of the smoothed input maps. They use a Monte-Carlo Markov Chain sampling technique as a way to determine the best-fit parameters. Once the non-linear parameters have been estimated at low resolution, by marginalizing over the component amplitudes, the latter are recovered at full resolution simply by solving the System (3.1).More details on the method can be found in Eriksen et al. (2006).

In the work we describe now (Stompor et al., 2008b), we modified the Eriksen's method, deriving analytically its maximum likelihood form, improving some of the features outlined above, while preserving the efficency of the two steps approach.

3.3.1 The maximum likelihood parametric approach

As introduced with Eq.(3.1), we can model the multi-frequency sky signal as

$$\mathbf{d_p} = \mathbf{A_p}\,\mathbf{s_p} + \mathbf{n_p}. \tag{3.1}$$

Here, $\mathbf{d_p}$ is a data vector containing the measured signals for all n_f frequencies (but also for the n_s Stokes parameters, if one wants to explicitely deal with polarized maps), which are to be analyzed simultaneously; $\mathbf{s_p}$ is a vector of the underlying true values of the n_c components, to be estimated from the data; $\mathbf{A_p} \equiv \mathbf{A_p}(\beta)$ is the mixing matrix, which hereafter will be assumed to be parameterized by a set of unknown parameters $\{\beta_i\}$. With these definitions and models in hand we can write a single pixel log-likelihood which up to an irrelevant constant, K, is given by,

$$-2\,\ln\,\mathcal{L}\left(\mathbf{s_p},\beta_\mathbf{i}\right) = \mathrm{K} + \left(\mathbf{d_p} - \mathbf{A_p}\,\mathbf{s_p}\right)^t\,\mathbf{N_p^{-1}}\left(\mathbf{d_p} - \mathbf{A_p}\,\mathbf{s_p}\right). \tag{3.2}$$

Here, $\mathbf{N_p}$ is a square, symmetric noise matrix of the frequency maps for a pixel p. This likelihood is the basis of the Eriksen et al. (2006) approach. It is clearly straightforward to introduce a multi-pixel version of this likelihood:

$$-2\,\ln\,\mathcal{L}_{data}\,(\mathbf{s},\beta) = \mathrm{K} + (\mathbf{d} - \mathbf{A}\,\mathbf{s})^t\,\mathbf{N^{-1}}\,(\mathbf{d} - \mathbf{A}\,\mathbf{s}), \tag{3.3}$$

where all the matrices and vectors now span over many pixels. Hereafter, we will refer to this likelihood as the full data likelihood. In the simple case of the matrix $\mathbf{N}$ being block-diagonal (thus allowing for correlations between different Stokes parameters at each pixel), Eq. (3.3) becomes:

$$-2\,\ln\,\mathcal{L}_{data}\,(\mathbf{s},\beta) = \mathrm{K} + \sum_p \left(\mathbf{d_p} - \mathbf{A_p}\,\mathbf{s_p}\right)^t\,\mathbf{N_p^{-1}}\left(\mathbf{d_p} - \mathbf{A_p}\,\mathbf{s_p}\right). \tag{3.4}$$

This likelihood reaches its maximum for the values of $\mathbf{s}$ and $\beta_\mathbf{i}$ that satisfy the relations:

$$-\left(\mathbf{A}_{,\beta}\,\mathbf{s}\right)^t\,\mathbf{N^{-1}}\,(\mathbf{d} - \mathbf{A}\,\mathbf{s}) = \mathbf{0} \tag{3.5}$$

$$\mathbf{s} = \left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}\right)^{-1}\,\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{d}, \tag{3.6}$$

where $_{,\beta_\mathbf{i}}$ denotes a partial derivative with respect to $\beta_\mathbf{i}$. Solving the generally non-linear first equations above can be potentially problematic in the case of multiple spectral parameters, due to their generally non-linear character (Brandt et al., 1994). However, once the values of the spectral parameters are found the second set of linear equations provides straightforward estimates of the pixel amplitudes.

3.3.2 Non-linear parameters

The new idea in the work by Stompor et al. (2008b) is to marginalize the likelihood in Eq. (3.3) over component amplitudes **s**:

$$\begin{aligned} -2\,\ln \mathcal{L}_{marg}\,(\beta) &= -2\,\ln \int d\,\mathbf{s}\,\exp\left[-\frac{\mathbf{1}}{\mathbf{2}}\,(\mathbf{d}-\mathbf{A}\,\mathbf{s})^{\mathbf{t}}\,\mathbf{N^{-1}}\,(\mathbf{d}-\mathbf{A}\,\mathbf{s})\right] \\ &= \mathrm{K} - \left(\mathbf{A^t\,N^{-1}\,d}\right)^{t}\,\left(\mathbf{A^t\,N^{-1}\,A}\right)^{-1}\left(\mathbf{A^t\,N^{-1}\,d}\right) \\ &+ \ln\left|\left(\mathbf{A^t\,N^{-1}\,A}\right)^{-1}\right|, \end{aligned} \tag{3.5}$$

where |...| denotes the matrix determinant. This expression defines the likelihood function for the parameters, $\beta_{\mathbf{i}}$, given the data, **d**, and under the assumption of the flat priors for the component amplitudes, **s**. Thanks to this relation, the spectral parameters for any subset of pixels can be estimated via any of the standard methods of gridding, maximization or sampling. This approach is fully consistent with the one proposed by Eriksen et al. (2006), apart from the fact that it does not recover the components amplitudes at low resolution, something that usually is not useful in the analysis anyway. On the other hand, this analytically derived likelihood function has the advantage that it is really straightforward to incorporate in the formalism some of the essential features of the CMB data sets, as we will see later.

Marginalizing over the component amplitudes is not the only improvement made by Stompor et al. (2008b). Indeed, the spectral parameter estimate based on the maximum value of the marginalized likelihood in Eq. (3.5) turns out to be biased with respect to an ensemble of noise realizations. This can be seen by calculating the first derivative of Eq. (3.5):

$$\begin{aligned} -2\,\frac{\partial\,\ln \mathcal{L}_{marg}}{\partial \beta_{\mathbf{i}}} &= + \;2\left(\mathbf{A^t\,N^{-1}\,d}\right)^{t}\,\hat{\mathbf{N}}\,\mathbf{A^t\,N^{-1}\,A}_{,\beta_{\mathbf{i}}}\,\hat{\mathbf{N}}\left(\mathbf{A^t\,N^{-1}\,d}\right) \\ &- \;2\left(\mathbf{A}^{\mathbf{t}}_{,\beta_{\mathbf{i}}}\,\mathbf{N^{-1}\,d}\right)^{t}\,\hat{\mathbf{N}}\left(\mathbf{A^t\,N^{-1}\,d}\right) \\ &- \;2\,\mathrm{tr}\left[\hat{\mathbf{N}}\left(\mathbf{A^t\,N^{-1}\,A}_{,\beta_{\mathbf{i}}}\right)\right], \end{aligned} \tag{3.4}$$

and taking its average over the noise realization, denoted as $\langle ... \rangle$, with the spectral parameters set to their true values,

$$-2\,\langle\,\frac{\partial\,\ln \mathcal{L}_{marg}}{\partial \beta_{\mathbf{i}}}\bigg|_{\beta=\beta_{\mathbf{true}}}\,\rangle = -2\,\mathrm{tr}\,[\hat{\mathbf{N}}\,\mathbf{A^t\,N^{-1}\,A}_{,\beta_{\mathbf{i}}}], \tag{3.5}$$

which in general does not vanish. Here, we have introduced

$$\hat{\mathbf{N}} \equiv \left(\mathbf{A^t\,N^{-1}\,A}\right)^{-1}, \tag{3.6}$$

and exploited the fact that $\hat{\mathbf{N}}$ is symmetric. The bias is quantified by the logarithmic term in Eq. (3.5). To avoid it we can introduce an unbiased estimator based on the

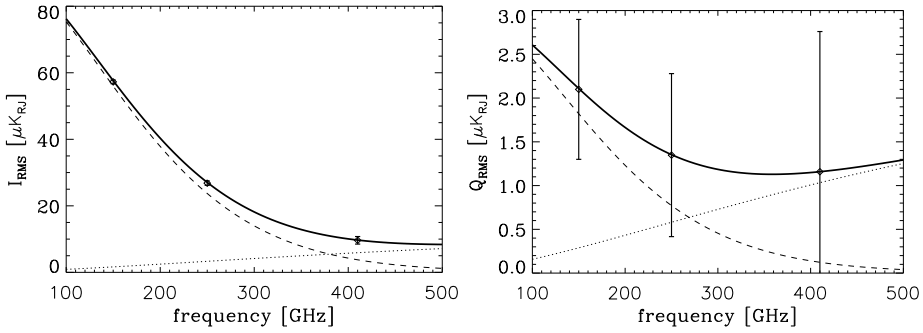

Figure 3.5: The RMS of the signal amplitudes for the CMB (dot-dashed) and thermal dust (dashed line) smoothed with the 8′ beam and for, left panel, total intensity, I, and, right panel, Stokes Q parameter, is shown as a function of frequency. The solid line indicates the RMS of the total coadded signal. The noise RMS per pixel ($n_{side} = 1024$) is also shown for the three frequency channels considered in the text. From Stompor et al. (2008b).

maximization of the marginalized likelihood as before but with the logarithmic term dropped:

$$-2\ \ln\ \mathcal{L}_{spec}\,(\beta) \equiv \mathrm{K} - \left(\mathbf{A^t\,N^{-1}\,d}\right)^t\ \left(\mathbf{A^t\,N^{-1}\,A}\right)^{-1}\left(\mathbf{A^t\,N^{-1}\,d}\right). \tag{3.7}$$

The equation above is one of the main results of the work by Stompor et al. (2008b). The additional attractive feature is that its solution coincides, case-by-case, with the maximum likelihood solution of the initial, non-marginalized likelihood as in Eq. (3.3). This can be seen by comparing the first derivative of both likelihoods, which in the case of the full data likelihood reads (see Eqs. (3.3) and (3.5)),

$$\begin{aligned} -2\ \frac{\partial\ \ln \mathcal{L}_{data}}{\partial \beta_{\mathbf{i}}} = \quad & - \quad 2\left(\mathbf{A^t\,N^{-1}\,d}\right)^t\ \hat{\mathbf{N}}\,\mathbf{A^t_{,\beta_i}}\,\mathbf{N^{-1}\,d} \\ & + \quad 2\left(\mathbf{A^t\,N^{-1}\,d}\right)^t\ \hat{\mathbf{N}}\,\mathbf{A^t_{,\beta_i}}\,\mathbf{N^{-1}\,A}\,\hat{\mathbf{N}}\,\mathbf{A^t\,N^{-1}\,d} \end{aligned} \tag{3.7}$$

and which thus agrees with the expression in Eq. (3.4), but only when the logarithmic term derivative is neglected.

Numerical examples

A few examples of the applications of the formulae above will be presented now. To perform such a study, we simulated the CMB and thermal dust polarized emission on a patch of sky with an area of 350 square degrees centered at $RA = 60^\circ$ and $DEC = -45^\circ$, adopting the HEALPix convention (Górski et al., 2005) , with a pixel resolution of 3.4′ ($n_{\rm side}$=1024). Simulated sky are made at three channels (150, 250 410 GHz) at 8′ resolution, with a white noise RMS level of 0.56, 0.66, 1.13 μK respectively, in antenna (Rayleigh-Jeans) units. We stress that this setup corresponds to a typical one for modern suborbital probes aiming at CMB polarization measurements (Oxley et al., 2004). The components have been simulated

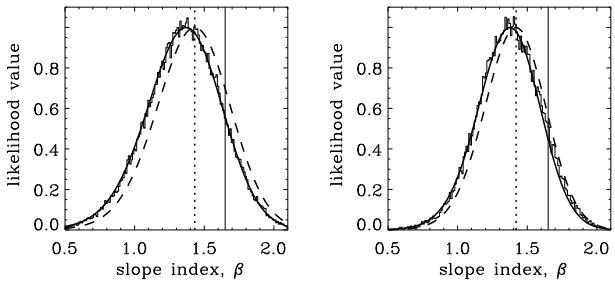

Figure 3.6: The one dimensional likelihoods for the slope index, β, computed for the total intensity and the Stokes Q parameter (left and right panels). The histograms show the result of the MCMC sampling of the full data likelihood, Eq. (3.3). The thick solid and dashed lines show the result for the biased and unbiased likelihood expressions, Eqs. (3.5) and (3.7) respectively. The dotted vertical lines show the β value corresponding to a peak found through a direct grid search. From Stompor et al. (2008b).

following the recipes introduced in the last Chapter with the only difference that in order to extrapolate the dust from 65 GHz back up to our frequency range we used 'Model 3' of Finkbeiner et al. (1999),

$$s_{\rm d}(\nu) = A_{\rm d}\frac{\nu}{\exp\frac{h\nu}{kT_{\rm d}}-1}\frac{\exp\frac{h\nu_0}{kT_{\rm d}}-1}{\nu_0}\left(\frac{\nu}{\nu_0}\right)^{\beta}, \tag{3.8}$$

with $T_{\rm d} = 18.1$K and $\beta = 1.65$, assumed constant on the patch. In Fig.3.5 we plot the frequency scaling of the CMB and dust RMS for total intensity and polarization, along with the error bars from simulated noise. The dust spectral index β is the non-linear parameter involved in the problem (the temperature $T_{\rm d}$ is assumed to be known) and the following analysis will be centered on it.

Three major observations made so far are visualized in Fig. 3.6. First of all, there is a perfect agreement between the numerical marginalization (the histograms, computed via MCMC) of the full data likelihood in Eq.(3.3) and the analytic formula derived, Eq.(3.5) (solid lines). Second, it is evident the presence of the bias when these two quantities are compared to the peak of the full data likelihood found through a direct grid search (vertical dotted line). Finally, the lack of such a bias when the spectral likelihood, Eq.(3.7) (dashed lines), is adopted.

But let's push the comparison a bit further. In the search for the best estimation of β, there are two implementations of the formalism that can be adopted. One is the *single-pixel* approach, as proposed by Eriksen et al. (2006), that is to first smooth the available data and downgrade it to a lower resolution, prior to calculating the spectral likelihood for each low-resolution pixel separately. The spectral parameters determined in this way are then applied to all the high resolution pixels falling into the low resolution one. The second implementation is the one proposed by Stompor et al. (2008b) in which, without any preprocessing, a subsets of the available maps, for which the assumption of a single set of spectral parameters is

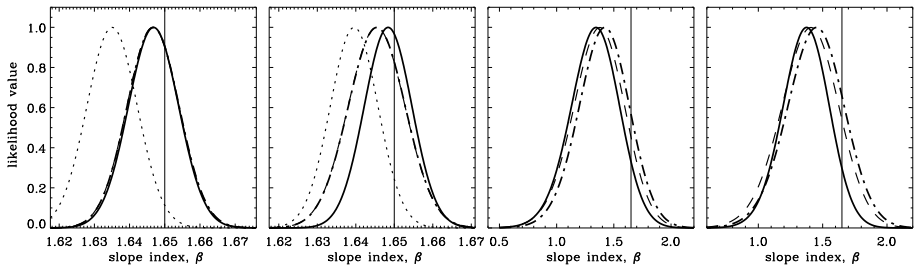

Figure 3.7: The one dimensional likelihoods for the slope index, β, computed for the total intensity and the Stokes Q parameter (two left and two right panels). The solid and dotted lines represent the unbiased and the biased cases, equations 3.7 and 3.5 respectively, computed using the *multi-pixel approach*. The dashed and dot-dashed lines (nearly perfectly overlapping) show the corresponding results derived for the *single-pixel approach*. The homogeneous and inhomogeneous noise cases are shown by the first and the second panel, respectively, of each pair of panels. From Stompor et al. (2008b).

physically justifiable, is analyzed. Let's call the latter *multi-pixel* approach.

The likelihood functions (both the biased and the unbiased) calculated in the two ways are presented in Fig. 3.7. The two left panels show the total intensity cases for homogeneous and inhomogeneous noise distribution (the latter was obtained by allowing for a $\sqrt{3}$ noise RMS variation across the patch, changing randomly across the patch). The unbiased and the biased likelihood are represented by the solid and dotted lines in the *multi-pixel* approach, while the nearly overlapping dashed and dot-dashed lines are for the *single-pixel* one. In the homogeneous case, the bias in the likelihood of Eq.(3.5) is evident for the *multi-pixel* approach, while the same likelihood still performs well in the *single-pixel* approach (I will give an intuitive explanation for that later on). Nevertheless, in the inhomogeneous noise case, the difference between the two approaches as well as the biased and unbiased likelihoods is much more evident and manifests itself as an overall shift and broadening, showing the macroscopic gain in adopting the expression in equation (3.7) throughout the analysis.

The two right panels of Fig. 3.7 show analogous cases but for the Stokes Q parameter. Here the biased *multi-pixel* likelihoods (dotted lines) are not shown as they do not fit the x-axis ranges. Note that unlike in the total intensity case the single and multi-pixel likelihoods differ even in the homogeneous noise case (middle panel, the solid and dashed lines). This reflects the fact that in the simulations used here there is more small-scale power in the dust polarization than in the total intensity, which leads to the loss of constraining power as a result of smoothing in the single pixel approach. The effect is additionally enhanced if the inhomogeneous noise is present as shown in the right panel. In both cases the bias and the loss of precision are small, as expected given a relatively smooth variation of the dust foreground as adopted here.

At this point one may wonder what happens when more pixels are added to

the analysis. As expected, by using more pixels, the constraints can be usually improved. However the rate at which that takes place is not trivial and will depend on the specific, and not known a priori, magnitude of the sky component in the newly added pixels, as well as the pixel instrumental noise levels, and therefore may not conform with the usual, $\sim 1/\sqrt{n_{pix}}$, expectation. Nevertheless, for the diffuse components, and in particular at high Galactic latitudes we find that once the number of included sky pixels is large enough the uncertainty of the slope determination falls roughly as expected, i.e., $\sim 1/\sqrt{n_{pix}}$. However, even in an extreme case when the newly added pixels happen to contain no information about a given component, the uncertainty, estimated using the multi-pixel approach, is guaranteed not to deteriorate. This may not be however the case with the single pixel approach, where including too many, for instance, dust-free pixels may suppress the noise power less than that of the dust and consequently increase the errors of the dust parameter determination. In a less extreme and more common case, the smoothing of the rapidly varying sky components whose amplitudes change across the low resolution pixel, will somewhat affect the precision of the spectral parameter determination, though will not bias the estimation result. The presence of the inhomogeneous noise on the scales smaller than the low resolution pixel will also have similar consequences.

In the examples considered so far, the smoothing generally leads to an improvement of the constraints on the spectral parameters because the noise is typically suppressed more significantly than the dust component. Anyway, in all the cases shown, the *multi-pixel* approach will produce nearly optimal constraints, independently of the actual sky distribution of the signals. It is also more flexible as it permits arbitrary pixel subsets for which identical spectral parameter values are assumed, and thus can better deal with the masked pixels and sky patch edge effects. Moreover, as discussed before, it also treats more optimally cases in which either the noise or the relative component content varies rapidly across the sky. For all these reasons the *multi-pixel* approach described in Stompor et al. (2008b) is therefore an approach of the choice here.

The evaluation of the spectral index is the first step toward the recovery of the actual maps of the components in the data, that will be described in the next Section.

3.3.3 Component amplitudes

We can calculate the maximum likelihood estimates of the component maps, given the maximum likelihood estimates of the spectral parameters derived in the previous Section and expression for the peak of the data likelihood, equation (3.6). I rewrite that equation here for simplicity:

$$\mathbf{s} = \left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}\right)^{-1}\,\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{d}. \tag{3.8}$$

The above equation also corresponds to a standard general least square solution for the components, if the spectral parameters are perfectly known ahead of the

time. In this case the solution error is quantified by the error correlation matrix $\hat{\mathbf{N}}$, equation (3.6):

$$\hat{\mathbf{N}} \equiv \left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}\right)^{-1}. \tag{3.8}$$

Whenever the spectral parameters are not known perfectly, and need to be constrained from the data, we can get some insight into the structure of the error correlation matrix from the curvature matrix computed at the peak of the likelihood. Stompor et al. (2008b) provided a complete analysis of the solution error (see Appendix A of the paper for the full computation). We can write explicitly this expression for the component amplitudes:

$$\begin{aligned}\tilde{\mathbf{N}}^{\mathbf{\Lambda}}_{\mathbf{ss}} = \hat{\mathbf{N}} + \left[\hat{\mathbf{N}}\left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}_{,\beta}\,\mathbf{s} - \mathbf{A}^{\mathbf{t}}_{,\beta}\,\mathbf{N^{-1}}\,(\mathbf{d}-\mathbf{A}\,\mathbf{s})\right)\right] \times \\ \tilde{\mathbf{N}}^{\mathbf{\Lambda}}_{\beta\beta}\left[\hat{\mathbf{N}}\left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}_{,\beta}\,\mathbf{s} - \mathbf{A}^{\mathbf{t}}_{,\beta}\,\mathbf{N^{-1}}\,(\mathbf{d}-\mathbf{A}\,\mathbf{s})\right)\right]^{\mathbf{t}}.\end{aligned} \tag{3.9}$$

This quantity is the $\mathbf{s}-\mathbf{s}$ diagonal block of the inverse curvature matrix. It describes the correlation pattern of the recovered sky components, while the $\beta-\beta$ block, $\mathbf{N^{\Lambda}}_{\beta\beta}$, provides estimates of the errors, and their correlations, of the recovery of the spectral parameters:

$$\begin{aligned}\tilde{\mathbf{N}}^{\mathbf{\Lambda}}_{\beta\beta} &= \left\{\left(\mathbf{A}_{,\beta_\mathbf{i}}\,\mathbf{s}\right)^{t}\,\mathbf{N^{-1}}\left(\mathbf{A}_{,\beta_\mathbf{j}}\,\mathbf{s}\right) - \left(\mathbf{A}_{,\beta_\mathbf{i}\beta_\mathbf{j}}\,\mathbf{s}\right)^{\mathbf{t}}\,\mathbf{N^{-1}}\,(\mathbf{d}-\mathbf{A}\,\mathbf{s})\right. \\ &- \left[\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}_{,\beta}\,\mathbf{s} - \mathbf{A}^{\mathbf{t}}_{,\beta}\,\mathbf{N^{-1}}\,(\mathbf{d}-\mathbf{A}\,\mathbf{s})\right]^{t} \\ &\times \left.\hat{\mathbf{N}}\left[\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}_{,\beta}\,\mathbf{s} - \mathbf{A}^{\mathbf{t}}_{,\beta}\,\mathbf{N^{-1}}\,(\mathbf{d}-\mathbf{A}\,\mathbf{s})\right]\right\}^{-1}.\end{aligned} \tag{3.7}$$

The two relations above are clearly an important tool available in the application of this method. One can analytically associate an error to the recovered component maps, a fundamental step that is still missing to many other component separation techniques.

3.4 Correlated Component Analysis

The last method we present in this work is the Correlated Component Analysis (CCA, Bedini et al., 2005; Bonaldi et al., 2006), which is a technique that exploits second order statistics of the data to estimate the spectral behaviour of the mixed components. Let's introduce the basic idea.

3.4.1 Mixing matrix estimation with CCA

As usual, we express the data vector $\mathbf{d}$ in each pixel as:

$$\mathbf{d} = \mathbf{A}\mathbf{s} + \mathbf{n}, \tag{3.7}$$

where $\mathbf{A}$ is a $M \times N$ mixing matrix, $\mathbf{s}$ is the N-vector of sources (components) and $\mathbf{n}$ the M-vector of instrumental noise (M is then the number of independent maps used in the analysis).

Given a generic signal **d**, defined in a two dimensional space with coordinates (ξ, η), the covariance matrix of this signal is:

$$\mathbf{C_d}(\tau, \psi) = \langle [\mathbf{D}(\xi, \eta) - \mu][\mathbf{D}(\xi + \tau, \eta + \psi) - \mu]^{\mathbf{T}}, \rangle, \tag{3.8}$$

where $\langle ... \rangle$ denotes expectation under the appropriate joint probability and μ is the mean vector. Every covariance matrix is characterized by the shift pair (τ, ψ) that links each pixel to a shifted one: τ and ψ are the increments in the ξ and η coordinates.

From Eq.(3.4.1) we can easily derive a relation between the data covariance matrix $\mathbf{C_d}$ at a certain lag, the source covariance matrix $\mathbf{C_s}$ at the same lag, the mixing matrix **A**, and the noise covariance matrix $\mathbf{C_n}$:

$$\mathbf{C_d}(\tau, \psi) = \mathbf{A}\mathbf{C_s}(\tau, \psi)\mathbf{A}^{\mathbf{T}} + \mathbf{C_n}(\tau, \psi). \tag{3.8}$$

The covariance matrix $\mathbf{C_d}$ can be estimated from the data as

$$\mathbf{C_d}(\tau, \psi) = \frac{1}{N_p} \sum_{\xi, \eta} [\mathbf{d}(\xi, \eta) - \mu_{\mathbf{d}}][\mathbf{d}(\xi + \tau, \eta + \psi) - \mu_{\mathbf{d}}]^{\mathbf{T}}, \tag{3.8}$$

where N_p is the number of pixels sampling the data. Given a noise process, we can model the noise correlation matrix $\mathbf{C_n}$: for example, if noise can be assumed to be signal-independent, white and zero-mean, then for a null lag $\mathbf{C_n}$ is a diagonal matrix whose elements are the noise variances in the measured channels, while for $(\tau, \psi) \neq (0, 0)$, $\mathbf{C_n}$ is the null $M \times M$ matrix. If the noise process deviates significantly from this ideal model, $\mathbf{C_n}$ can be computed using Monte Carlo simulations of noise maps.

Once $\mathbf{C_d}$ and $\mathbf{C_n}$ have been set, Eq. (3.4.1) can be used to identify the mixing operator **A**. The strategy of CCA is to parameterize the mixing matrix in order to reduce the number of unknowns and then to take into account enough nonzero shift pairs (τ, ψ) in order to estimate both **A** and $\mathbf{C_s}$. To solve the identification problem we perform the minimization:

$$(\Gamma, \Sigma(:,:)) = \text{argmin} \sum_{\tau, \psi} \| \mathbf{A}(\mathbf{\Gamma})\mathbf{C_s}[\mathbf{\Sigma}(\tau, \psi)]\mathbf{A}^{\mathbf{T}}(\mathbf{\Gamma}) + \\ -\hat{\mathbf{C}}_{\mathbf{d}}(\tau, \psi) - \mathbf{C_n}(\tau, \psi) \|, \tag{3.9}$$

where Γ is the vector of all parameters defining **A** and $\Sigma(:,:)$ is the vector containing all the unknown elements of matrices $\mathbf{C_s}$ for every shift pair.

The main product of CCA is an estimate of the mixing matrix **A**, locally on patches, whose dimensions can vary case by case, depending on how many pixels are needed to have the problem in Eq.(3.9) to converge (Bonaldi et al., 2006). Once the matrix has been estimated, it can be used to perform the source reconstruction with standard inversion methods.

3.4.2 The analysis on WMAP

In this Section, I will present the main results of the analysis we performed on the WMAP 3-year data using the CCA technique. Complete details can be found in Bonaldi et al. (2007). This work was mainly focused on the study of the foregrounds and the impact of the residuals of the separation on the CMB, as we will see now.

The data set

The basic data are the WMAP 3-year maps at all the frequency channels (K, Ka, Q, V and W bands). As requested by the algorithm, some preprocessing of the data was necessary. First of all, the maps were smoothed to the common resolution of 1°. Then, a proper mask had to be defined to put the code in the best work conditions. Point sources have been masked according to the point source catalogue provided by the WMAP team and, to mask out the most intense foreground region, we applied a Galactic cut of $\pm 3^\circ$. Finally, we masked a few highly contaminated regions, namely Cen A, the Large Magellanic Cloud, ρ Oph, Orion A, Orion B and Tau A.

The noise correlation matrix, $\mathbf{C_n}$, has been estimated with a Monte-Carlo on 10 simulated noise maps with the nominal characteristics of the instrument, preprocessed in the same way as the data.

Together with the WMAP channels, we found necessary to complement the data set with two templates, one for the thermal dust and one for the synchrotron. As explained in the previous Chapter, the former is obtained extrapolating to 850 GHz the FIRAS dust map using the best-fit model of Finkbeiner et al. (1999), while the synchrotron map is based on the Haslam et al. (1982) 408 MHz map. To avoid free-free contamination in the synchrotron template, the latter was obtained by subtracting from the Haslam map the free-free contribution estimated from the Hα map corrected for dust absorption (Dickinson et al., 2003).

Evidence for an extra component

The first step in the analysis was to define a conservative model for the data, containing only the standard mixture of CMB, synchrotron, free-free and dust. Since the thermal dust spectrum is only very poorly constrained by the WMAP data, the results are very weakly dependent on the assumed dust temperature and emissivity index. We then fixed these quantities to the commonly used values $T_{\rm dust} = 18\,{\rm K}$ and $\beta_d = 1.67$ in Eq.(3.8). Since also the free-free spectral index can be safely assumed as known, we are left with only one free parameter, the synchrotron spectral index β_s.

The distribution of β_s obtained with CCA is shown in Fig.3.8 (left panel). The two-peak feature is evident. One peak is at $\beta_s \simeq 2.7$ and has a dispersion $\sigma \simeq 0.2$, which is roughly what is expected for the synchrotron emission. The second peak, at $\beta_a = 2.3852$, is extremely narrow ($\sigma \simeq 0.0004$), hinting at a different

component. We explicitly checked that this distribution is unaffected by different choices of the thermal dust parameters. In Fig.3.8, right panel, we show the map of the recovered synchrotron spectral indices. The flattish component is mainly located at low Galactic latitudes, more than $\sim 40^\circ$ away from the Galactic center and does not correlate significantly with the synchrotron template.

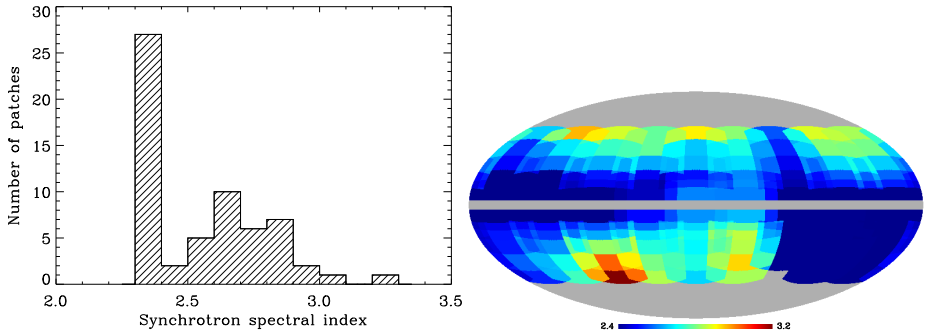

Figure 3.8: (*Left panel*) The distribution of the synchrotron spectral index recovered by CCA. It shows the presence of an unmodeled component with a very narrow distribution index. (*Right panel*) Map of the recovered spectral index. The flatter component do not show significant correlation with the synchrotron template.

To account for this extra component, it was necessary to put it in the model. According to de Oliveira-Costa et al. (2004) and Davies et al. (2006) the spectrum of the anomalous emission for $\nu > 20\,\mathrm{GHz}$ may be represented by a parabola in the $(\log\nu, \log S)$ plane:

$$\begin{aligned} \log T_{\mathrm{A,X}}(\nu) &= \mathrm{const} - \left(\frac{m_{60}\log\nu_{\max}}{\log(\nu_{\max}/60\,\mathrm{GHz})} + 2\right)\log\nu \\ &+ \frac{m_{60}}{2\log(\nu_{\max}/60\,\mathrm{GHz})}(\log\nu)^2, \end{aligned} \tag{3.8}$$

with ν in GHz and $\nu_{\max} = 20\,\mathrm{GHz}$. The free parameter, m_{60}, is the angular coefficient at 60 GHz in the $(\log\nu, \log S)$ plane. CCA found that the mean values of m_{60} over the sky patches are in the range $3.8 \leq m_{60} \leq 4.5$, and correlate with β_s. The linear best-fit relation is:

$$m_{60} = (2.1101 \pm 0.0005)\beta_s - (2.073 \pm 0.002). \tag{3.8}$$

The corresponding spectral shape is compatible with the anomalous emission detected by Davies et al. (2006).

Last clue on the nature of this extra component came from the study of its morphology, once the component maps were recovered. It turned out that it shows a high, but not perfect, correlation with the dust component. Davies et al. (2006) suggested that, if this component is due to spinning dust, it should be better correlated with the small grains dominating the mid-IR emission than with the big grains dominating at far-IR to sub-mm wavelengths. The map of this component yielded by our model could then constitute its first, albeit preliminary all-sky template.

Recovered CMB

Even if the work by Bonaldi et al. (2007) was mostly focused on the study of the foregrounds, the code was able to provide, as a side product, the CMB map showed in Fig.3.9. We did not go into a detailed analysis of this map, beside the fact that it shows a negligible spatial correlation with all the foreground templates used in the analysis when the extra component is considered into the model.

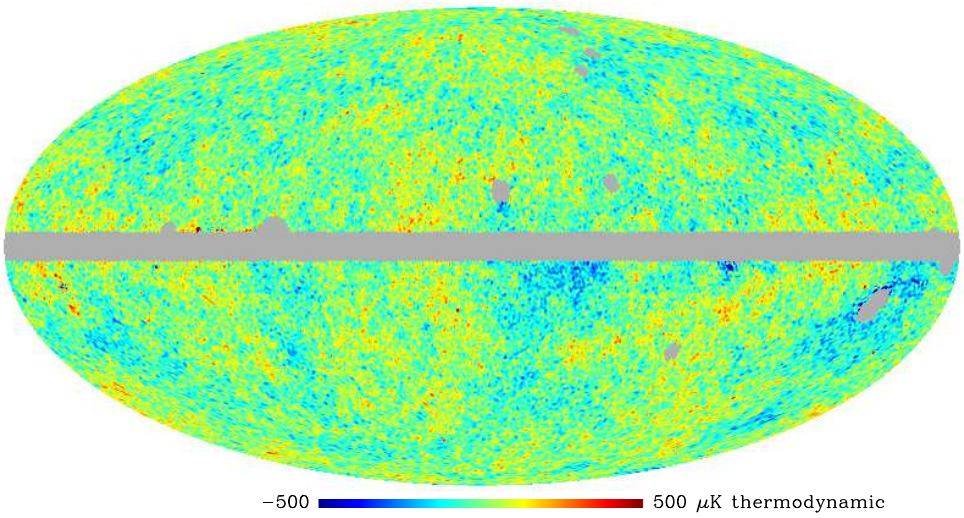

Figure 3.9: Recovered CMB maps, out of the WMAP 3 year data. This maps shows negligible corelation with the foreground templates used in the analysis.

The last question we tried to address was the following: to what extent do the different foreground modeling affect the estimates of the CMB power spectrum on large scales? To answer this question we performed the analysis with 5 different models for the foregrounds, including the one we introduced above. I'm not going into details here, but out of this 5 models, 3 provided reasonable results in terms of foreground recovery, all of them including a different model for the anomalous emission (see Bonaldi et al., 2007, for a complete discussion on this point). The recovered CMB power spectrum for each model is shown in Fig. 3.10. We generally find good agreement with the WMAP power spectrum, shown by the solid line, but there are significant differences among the various models at the lowest multipoles, especially evident for $l = 2$. At large multipoles the power spectrum seems to be unaffected by the choice of the model.

Regardless of which is the actual model we are exploiting, this result shows that the model of the Galactic components used to perform component separation can affect the CMB power spectrum on larger scales. Thus, large error bars are

required to take into account possible biases at these multipoles, something that the WMAP team never considered so far. In this analysis, as a rough estimate of uncertainties associated to foreground modeling, it makes sense to consider the spread of our three CMB power spectra. In Fig.3.11 we report the power spectrum as recovered by CCA, averaged on the three models, with the spread around the mean as error bars. It is compared to the best fit ΛCDM model (smooth solid line, shaded area is the cosmic variance).

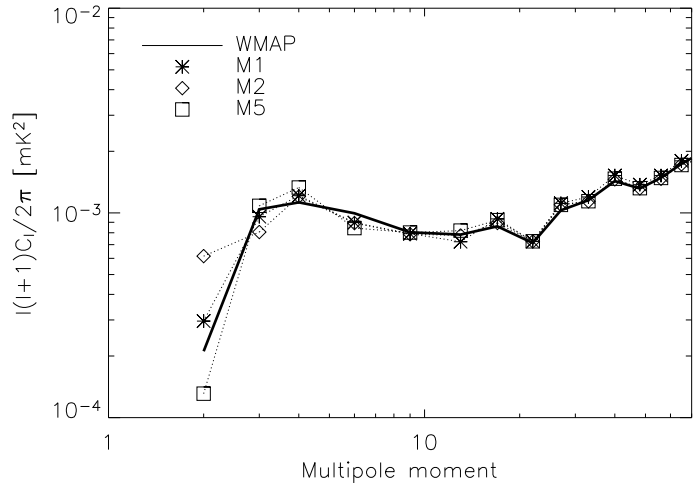

Figure 3.10: Binned CMB power spectra obtained from all models compared to the WMAP three-year power spectrum.

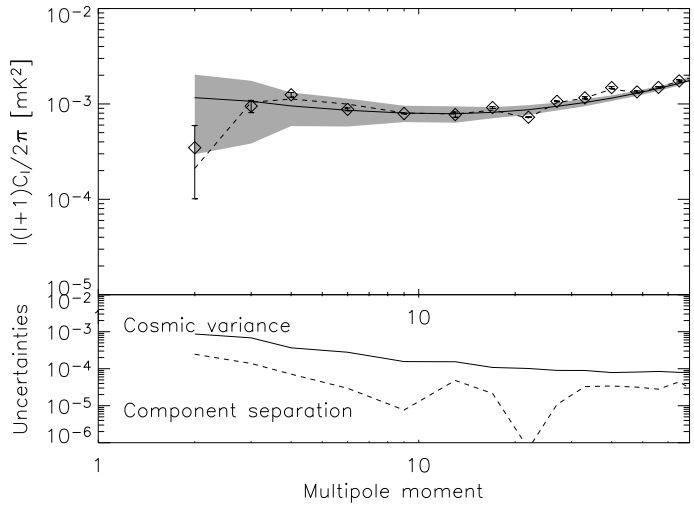

Figure 3.11: Upper panel: best fit ΛCDM model (smooth solid line) compared to the WMAP CMB power spectrum (dashed line) and to the mean final power spectrum resulting from our analysis (open diamonds; the error bars show the spread of results for different models). The shaded area shows the cosmic variance. Lower panel: uncertainties associated to foreground modeling (dashed line) compared to the cosmic variance (solid line).

Chapter 4

B mode recovery

4.1 Introduction

The main target of the planned probes for measuring the polarized component of the Cosmic Microwave Background (CMB) radiation is represented by the B modes, also known as 'curl' component (Zaldarriaga, Seljak, 1997; Kamionkowski et al., 1997). As introduced in the first Chapter, the CMB B mode signal is known to be generated by primordial gravitational waves and the weak lensing due to structures forming in the Universe (Zaldarriaga, Seljak, 1998), and thus contains unique information about the early Universe, and, potentially, physics of very high energies, beyond those reached in modern particle accelerators.
The B signal in the CMB polarization is more than one order of magnitude smaller than the 'gradient' mode (E) coming from all kinds of cosmological perturbations, and about two orders of magnitude lower if compared with the total intensity anisotropies (T). The CMB E mode and the TE cross-correlation, have been detected by the WMAP satellite as well as instruments operating on the ground (Kovac et al., 2002; Readhead et al., 2004) and from balloons (Montroy et al., 2005). No glimpse of the B has been seen so far and it is apparent that its detection will represent an experimental and data analysis challenge in terms of control and treatment of systematics and instrumental noise needed to attain the required precision. An additional important limiting factor for these experiments is related to foreground emissions. In the frequency range going from 70 to 150 GHz the diffuse Galactic emission is known to be sub-dominant with respect to the total intensity CMB signal at intermediate and high Galactic latitudes. However, the foreground, whose polarized pattern remains quite unknown expecially at high Galactic latitudes, is expected to highly contaminate the weak cosmological B signal everywhere in the sky and at any frequency (Baccigalupi, 2003).

In this context, it is crucial to develop reliable data analysis techniques and tools which are capable of cleaning the CMB polarized emission from the foreground contamination. In this Chapter, I will describe two applications of component separation aimed to B modes recovery.

4.2 FASTICA and the B modes

In Stivoli et al. (2006) we tested FASTICA performance in the reconstruction of the CMB polarized emission on a limited region of the sky, focusing on the recovery of B modes. Our work represented an exploratory study, not specialized to describe any particular operating or planned experiment, and not aiming at quantifying what is the minimum level of B modes detectable in presence of foregrounds. Our aim was rather to determine if the blind component separation techniques have the capability to recover the B modes corresponding to a tensor to scalar ratio of about 10%, observed on a limited patch of the sky in the presence of a substantial foreground contamination as estimated on the basis of the models of the Galactic emission.

4.2.1 Simulated maps

As said before, for this simulation we target a small patch in the sky where the foreground emission is known to be low. Such a region is centered on the position $(60°, -50°)$ in right ascension and declination. It has been observed by BOOMERANG 2K (Montroy et al., 2005), and is the target of the EBEX (Oxley et al., 2004), QUAD (Bowden et al., 2004) and QUIET experiments.

Two sets of simulated polarized skies were generated, one at low frequencies (40 GHz and 90 GHz) and one at high frequencies (150 GHz and 350 GHz). A combination of CMB plus synchrotron and CMB plus dust are simulated in the former and the latter pair, respectively. The sky models were based on the available data at the time in which the work was carried out.

The CMB emission is simulated accordingly to the cosmological concordance model (Spergel et al., 2003). The Hubble constant is H_0=72 km/s/Mpc, the overall geometry is flat, with a critical density made of baryons (4.4%), Cold Dark Matter (CDM) (22.6%), and the cosmological constant (73%). The radiation component consists of photons and three massless neutrino species. The optical depth to the last scattering surface is fixed at $\tau = 0.11$. The perturbations are Gaussian, with a primordial power spectrum characterized by a spectral index of scalar perturbations $n_s = 0.96$. Unless otherwise specified, the primordial gravity wave contribution was set to 10% of the scalar perturbation amplitude, with a spectral index fixed accordingly to the single field inflationary model, $n_t = -n_s/6.8$. When explicitly specified, we also considered the case in which no gravity waves are present. We included the contribution due to lensing in the power spectrum, which is responsible for substantial part of the power in the B modes of the CMB polarization anisotropies.

The polarized synchrotron emission was simulated with the following recipe. A template for the polarization angle θ was derived by exploiting the observations in the radio band: these measures indicate a rather high fluctuation level interpreted as the effect of the small scale structure of the Galactic magnetic field (Uyaniker et al., 1999; Duncan et al., 1999), scaling as $C_l^\theta \sim l^{-2}$ on degree and sub-degree

angular scale, up to the arc-minute (Tucci et al., 2002), consistently also with observations at medium Galactic latitudes (Carretti et al., 2005). It is worth noting here that WMAP three year analysis showed an evidence for a shallower slope in the polarization angle pattern, at least on large angular scales and intermediate Galactic latitudes (Page et al., 2007). In the last few years, the latter became the standard way to model the slope. In the Stivoli et al. (2006) paper the template for the polarization angle was obtained by adopting the form above for C_l^{θ}, and assuming Gaussian distribution. The polarized intensity was derived using the model by Giardino et al. (2002) who exploited the all sky template of synchrotron in total intensity at 408 MHz (Haslam et al., 1982), assuming a theoretically synchrotron polarization fraction of about 75%; since that template has a resolution of about one degree or less, they extrapolated the power to the smaller scales by exploiting the total intensity observations in the radio band (Uyaniker et al., 1999; Duncan et al., 1999).

The polarized emission from the diffuse thermal dust has been detected for the first time in the Archeops data (Benoit et al., 2004), indicating a 5% polarization fraction with respect to the total intensity emission, which is very well known at 100 μm and can be extrapolated at microwave frequencies fitting for the emissivity and temperature of two thermal components (Finkbeiner et al., 1999); in Stivoli et al. (2006), we adopted the model 8 of that work, where dust emissivity and temperatures do not vary across the sky. The dust polarization fraction reported by WMAP three years is also consistent with a few percent. The pattern of the polarization angle is much more uncertain, and due to the magnetized dust grains which get locally aligned along the Galactic magnetic field (Prunet et al., 1998; Jones et al., 1995). Since the geometry and composition of the dust grains is still very uncertain, the simplest assumption is that the Galactic magnetic field is 100% efficient in imprinting the polarization angle pattern to the synchrotron and dust emission (Baccigalupi, 2003).

In Figure 4.1 we show the contamination to the all sky CMB E and B spectra from the foreground emission corresponding to the synchrotron and dust diffuse Galactic signal after cutting out the Galactic plane up to $|b| = 50^{\circ}$, roughly corresponding to the latitudes considered in Stivoli et al. (2006). The foreground power has been evaluated by fitting the actual sky signal with a power law, $C_l = \alpha l^{\beta}$. As it is evident, the models of the foreground emission indicate that the contamination to the B modes of the CMB is relevant in all cases. The lines raising as l^2 represent the levels of instrumental noise which we considered here. We point out again that these simulations reflect the status of our knowledges and foreground modeling at the time the work was carried out.

For simplicity, we considered a circular sky patch with a radius $\theta_C = 10^{\circ}$ and 20°, corresponding to about 0.76% and 3.04% of the entire sky, respectively. The center in Galactic coordinates is at $l = 260^{\circ}$, $b = -62^{\circ}$, within the region considered by different experiments as said before. The sky emission at the various frequencies, corresponding to the Q Stokes parameter, is shown in Figures 4.2 &

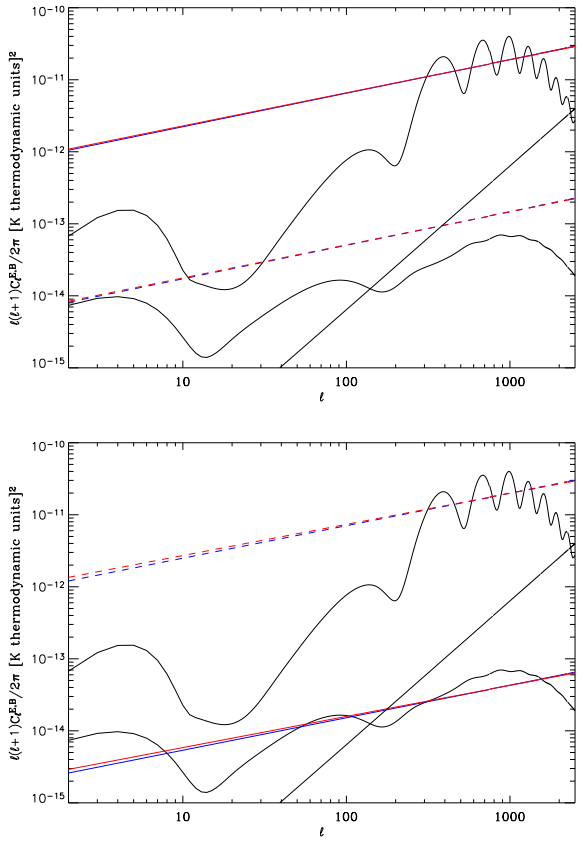

Figure 4.1: Power spectra of the different polarized sky signals relevant to the microwave observations. The almost flat straight lines represent the foreground contamination obtained by cutting out the Galactic plane up to $|b| = 50°$, and fitting those with a power law; the steep straight lines raising as l^2 represent the instrumental noise assumed in this work. The left and right panels show the predictions for the 40, 90 GHz and 150, 350 GHz frequency bands, respectively. The solid lines represent a lower frequency, while the dashed ones a higher one. These signals are plotted against the full sky CMB power spectra of anisotropies in thermodynamical units for E and B as assumed in Stivoli et al. (2006). Foregrounds models changed substantially in the last years.

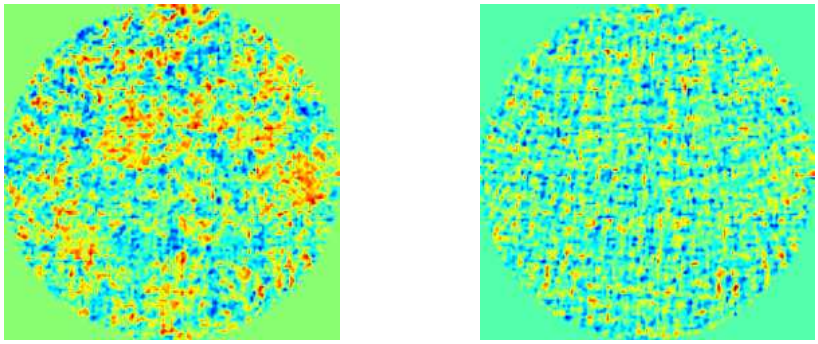

Figure 4.2: The total (CMB plus foregrounds) Q Stokes parameter emission in the sky area considered in the text, at 40 (left) and 90 GHz (right). At 90 GHz the signal appears dominated by the CMB signal, while the synchrotron contamination is evident at 40 GHz.

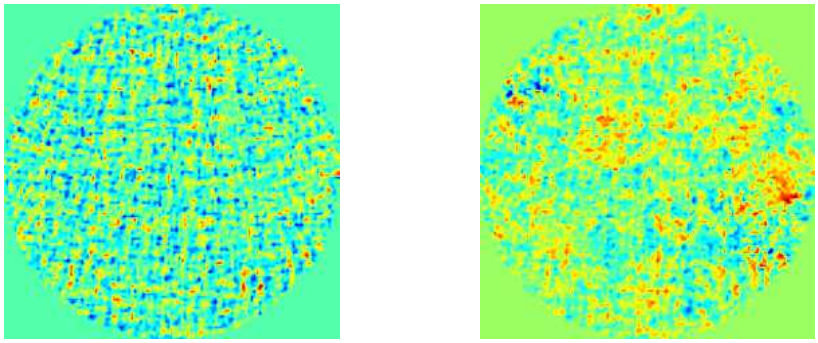

Figure 4.3: The total Q Stokes parameter emission in the sky area considered in the text, at 150 (left) and 350 GHz (right). At 150 GHz the signal appears dominated by the CMB emission, while the dust emission dominates at 350 GHz.

4.3. At 90 and 150 GHz the CMB signal appears relatively free of foreground contamination, while at 40 and 350 GHz the foregrounds dominate.

4.2.2 Polarization pseudo-power spectra

Since we performed component separation on a portion of the sky, we quantify the quality of the reconstruction with help of the angular pseudo-power spectra, which are relevant and straightforwardly calculable for the limited sky observations as described here. In a computation of the polarized, E and B pseudo-power spectra on a finite portion of the sky a transfer of power between the E and B modes occurs (see Chon et al., 2004, and references therein). Since the B modes are sub-dominant, the leakage of the E-mode power alters their spectrum more substantially and consequently needs to be explicitly considered in the presented analysis. We denote the pseudo-power spectra of E and B as $\tilde{C}_l^E$ and $\tilde{C}_l^B$, respectively, while the symbols without a tilde will correspond to their full-sky versions. Hereafter we compute the power spectra using a recipe adopted from Hansen et al. (2002). Consequently, we

introduce a window function, $G(\theta,\phi)$, (Gabor, 1946) which is applied to the data prior to a computation of the spherical harmonic transforms on a portion of the sphere and a calculation of the pseudo-power spectra. The leakage between the polarization modes may be written as

$$\tilde{C}_l^E = \sum_{l'} C_{l'}^E K_2(l,l') + \sum_{l'} C_{l'}^B K_{-2}(l,l') \,, \tag{4.0}$$

$$\tilde{C}_l^B = \sum_{l'} C_{l'}^B K_2(l,l') + \sum_{l'} C_{l'}^E K_{-2}(l,l') \,, \tag{4.0}$$

where C_l^E and C_l^B are the polarization full-sky power spectra (Zaldarriaga, Seljak, 1997), while the kernels $K_2(l,l')$ and $K_{-2}(l,l')$ depend on the form and the size of the cut, described by a generic function $G(\theta,\phi)$ which is zero in the sky regions which are not considered. The explicit expressions for the kernels are:

$$K_{\pm 2}(l,l') = \sum_{l''} g_{l''}^2 \frac{(2l'+1)(2l''+1)}{32\pi^2} W^2(l,l',l'')(1\pm(-1)^{l+l'+l''}) \,. \tag{4.0}$$

Here g_l are found by the inverse Legendre transform of the Gabor window $G(\theta,\phi)$ and the Wigner symbols W are defined as:

$$W(l,l',l'') = \begin{pmatrix} l & l' & l'' \\ -2 & 2 & 0 \end{pmatrix} . \tag{4.0}$$

We exploited these formulae for circular cut sky area of different size with top hat shape:

$$G(\theta) = \begin{cases} 1, & \theta \le \theta_C \,, \\ 0, & \theta > \theta_C \,. \end{cases} \tag{4.0}$$

As one can see from Equations (4.2.2) and (4.2.2), the sky cut mixes the polarization E and B modes, as quantified by the $K_{-2}(l,l')$ kernel. Obviously, the mixing gets reduced as the size of the window is increased. Since the cosmological fluctuations are dominated by the scalar contribution in the cosmological concordance model (Spergel et al., 2003), even if the diagonal of the kernel $K_2(l,l')$ is one order of magnitude larger than the diagonal of $K_{-2}(l,l')$, we expect the E mode to contaminate substantially the B signal even for large regions of the sky, while on the other hand $\tilde{C}_l^E \simeq C_l^E$.

In Fig.4.4 we show the pseudo-power spectrum of the B mode, $\tilde{C}_l^B$, as defined in (4.2.2) for a top hat window with $\theta_C = 10^\circ$ and 20°. In the latter case, the leakage from the E modes is slightly weaker. (Hereafter, we limit our analysis to a range of l-modes ≤ 1000 in order to speed up the calculation of the pseudo C_ls.) For comparison the dashed lines show the full sky B mode power spectra (with power normalized on the patch). As we see, the shapes of the two spectra are substantially different and the E mode contamination is relevant for the pseudo C_l^B.

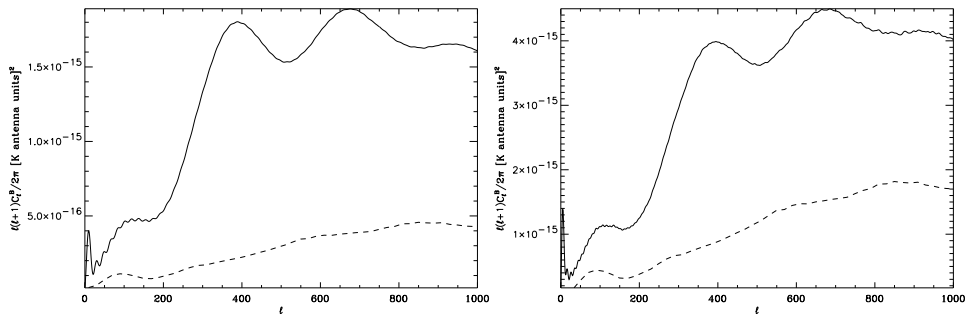

Figure 4.4: Pseudo-power spectra for the B modes, in the case of a circular top hat cut of $\theta_C = 10^\circ$ (left panel, solid line) and $\theta_C = 20^\circ$ (right panel, solid line). The dashed lines in both panels represent the full sky C_l^B normalized to the patch area fractions. The contamination due to the E mode is evident.

4.2.3 B modes reconstruction and error estimation

The sky signals in the patch considered are processed by the FASTICA code, and the outputs, in E and B, are shown in Figures 4.5, 4.6, and 4.7, 4.8, respectively. Those are plotted at 40 and 150 GHz in antenna units, as the code outputs are at the lowest frequency by default. In each panel, the two dotted curves correspond to the input pseudo $C_l^{E,B}$ of the CMB signal, $\pm 1\sigma$ where σ represents the cosmic variance on our patch of the sky: that is specified by a fraction f_{sky} and binned over $\Delta l = 50$ multipoles (Tegmark, 1997), and is given by

$$\Delta\tilde{C}_l^{E,B} = \sqrt{\frac{2}{(2l+1)\Delta l f_{sky}}}(\tilde{C}_l^{E,B} + \tilde{C}_{n,l}^{E,B}) \,, \tag{4.0}$$

where $\tilde{C}_{n,l}^{E,B}$ are the contribution of the noise. We assume a Gaussian and uniformly distributed noise over the analyzed region, with *rms* equal to a half of that of the CMB Stokes parameter Q or U on a single pixel, at each frequency. The noise amplitude is not related to any particular experiment, and was chosen as a starting point for the analysis performed in the next Section, where the noise amplitude is varied.

The symbols in the figures represent the signal recovered by the FASTICA separation process averaged over the 100 MC simulations of the CMB and noise, while the error bars show, a 1σ uncertainty derived from the simulations. Thus, they represent the extra uncertainty due to the separation process, given the foreground templates assumed here. At the bottom of each figure, we also plot the average and standard deviation of the residuals, obtained by subtracting the input from the output pseudo-power spectra for each realization. The averages provide a measure of biases of the reconstruction on each realization, while the error bars estimate the extra dispersion introduced due to the separation process.

The first feature to be noted is that the separation is clearly successful, for E and B

as well. Note that the B-mode pseudo-power spectra are generally comparable or lower than the foreground and noise contamination. As we stressed before, the ICA technique looks for the independent components into the data, assuming rigid and different frequency scaling and a different statistics for all of them, with no other prior; the fact that this procedure is able to extract with such a precision a signal which is comparable or lower than the foreground contamination in presence of noise is remarkable. Once again, the observed performance is made possible by the large number of pixels in the map, as well as the high level of statistical independence between background and foreground emission. These two facts bring the algorithm close to an ideal environment, ensuring the convergence very close to the correct answer, with a precision represented by the errors shown in the figures. A second, most interesting aspect to be noted is that we detect the error due to the separation process; that is clearly visible in all the figures as the excess in the error bars with respect to what predicted by cosmic variance and noise. The error from component separation is comparable or smaller, to the sample variance of the simulated templates. The error of the separation is either due to the randomness of the noise realizations on one hand and the fact that, for a single realization, background and foreground may not be completely independent. The latter factor can be a source of the extra randomness in the ICA performance thus contributing to the total error.

Although we were mainly interested in the extraction of the CMB B modes from the data, it is interesting also to look at the foreground recovery. In Fig.4.9 we plot the reconstructed pseudo-power spectra of the separated synchrotron compared with the original ones, reported with dotted lines. In this case the FASTICA is able to properly reconstruct the polarized signals of the foreground with good precision. On the other hand, the dust reconstruction fails, as it comes out heavily contaminated by the CMB, and with wrong normalization. This manifests that the separation with dust is more problematic, as it may be also noted by looking at Fig.4.8, which shows excess power in the recovered spectra and residuals with respect to the input ones, and which is mostly concentrated at low multipoles where the dust spectrum is highest. This occurrence should not be interpreted in terms of the different pattern of the foreground emission for dust and synchrotron, but in terms of the relative weight of it with respect to the background emission, as already noticed in earlier works (Maino et al., 2002; Baccigalupi et al., 2004). Due to the difference in the frequency scalings in the bands considered, in the 40, 90 GHz case the foreground and background signals are closer in amplitude with respect to the higher frequency combination; thus, at 150 GHz the CMB dominates over the dust while at 350 GHz the CMB emission is negligible. Indeed, this bias disappears if the foreground amplitude is raised by a factor of a few as we see next, and consequently the dust template can be better reconstructed.

Note that our Monte Carlo analysis does not include varying the foreground template, a factor that should be accounted for in order to quantify the error in the separation process in a comprehensive way. However, our modest knowledge of the foreground emissions does not allow to estimate their statistics to a level high

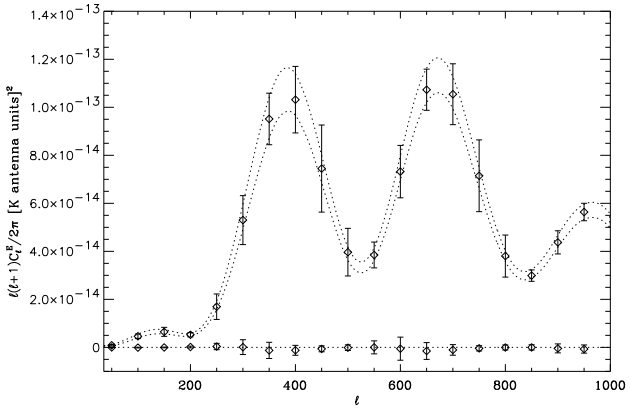

Figure 4.5: Pseudo-power spectra of the reconstructed $\tilde{C}_l^E$ modes of the CMB in the 40, 90 GHz frequency combination, in the $S/N = 2$ case. The region between the dotted lines is the input CMB signal $\pm\sigma$ cosmic and noise variance at 40 GHz on the sky area considered. At the bottom we show the average and standard deviation of the residuals on each realization. From Stivoli et al. (2006).

enough to vary their template in the Monte Carlo. Finally, we stress that this work represented a first attempt to evaluate the reconstruction errors introduced by the non-linear nature of the algorithm. As pointed out in the previous Chapter, an analytic treatment of the component separation through FASTICA is still missing.

4.2.4 Varying noise, foreground amplitudes and analyzed area

We performed a first study of a dependence of the results on some of the key simulation parameters. Specifically, we varied noise amplitude, foreground fluctuation amplitude, and extension of the sky area considered. We explored the corresponding parameter space by moving along the multiple 'directions' within its volume and use the results to set constraints on the applicability of the FASTICA approach given the assumed foreground pattern, however, still in systematic free cases.
In order to quantify the error introduced by the algorithm with respect to the one due to the cosmic variance and noise, and focusing on the B mode reconstruction, we introduced the quantities

$$d_l = \frac{\Delta\tilde{C}_l^{rec.}}{\Delta\tilde{C}_l^B} \quad , \quad r_l = \langle\frac{\tilde{C}_l^{rec.} - \tilde{C}_l^B}{\tilde{C}_l^B}\rangle_{\rm ICA} \quad , \quad a_l = \langle\frac{|\tilde{C}_l^{rec.} - \tilde{C}_l^B|}{\tilde{C}_l^B}\rangle_{\rm ICA}, \qquad (4.0)$$

meaning of which we explain now. d_l is the ratio between the dispersion of the recovered spectra over 100 realizations and the quantity defined in (4.2.3). Generally we expect this quantity to be larger than 1, accounting for the error introduced by the separation itself: a number close to 1 means that the separation procedure introduces an error which is negligible with respect to the input one; on the contrary,

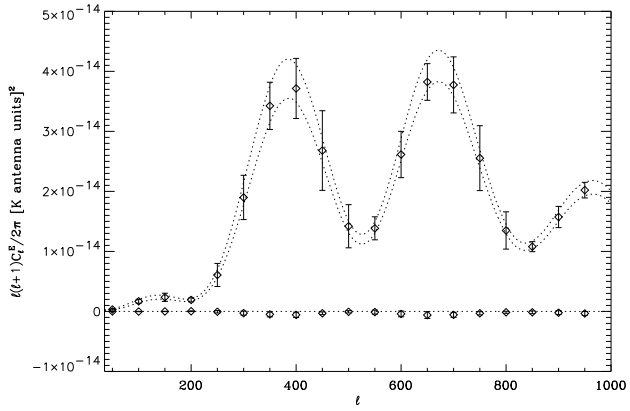

Figure 4.6: Pseudo-power spectra of the reconstructed $\tilde{C}_l^E$ modes of the CMB in the 150, 350 GHz frequency combination, in the $S/N = 2$ case. The region between the dotted lines is the input CMB signal $\pm\sigma$ cosmic and noise variance at 150 GHz on the sky area considered. At the bottom we show the average and standard deviation of the residuals on each realization. From Stivoli et al. (2006).

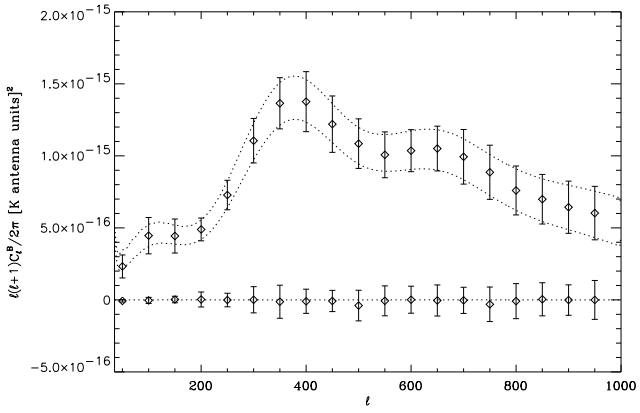

Figure 4.7: Pseudo-power spectra of the reconstructed $\tilde{C}_l^B$ modes of the CMB in the 40, 90 GHz frequency combination, in the $S/N = 2$ case. The region between the dotted lines is the input CMB signal $\pm\sigma$ cosmic and noise variance at 40 GHz on the sky area considered. At the bottom we show the average and standard deviation of the residuals on each realization. From Stivoli et al. (2006).

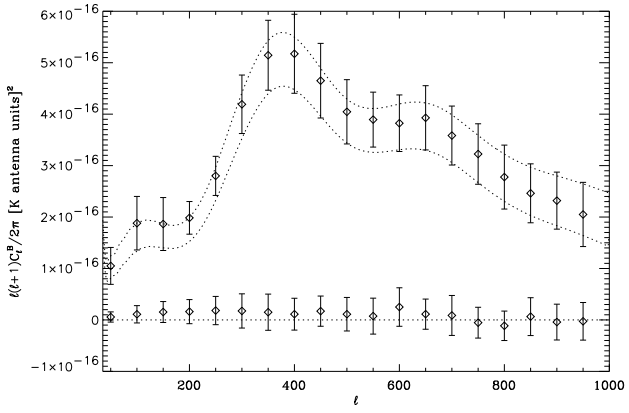

Figure 4.8: Pseudo-power spectra of the reconstructed $\tilde{C}_l^B$ modes of the CMB in the 150, 350 GHz frequency combination, in the $S/N = 2$ case. The region between the dotted lines is the input CMB signal $\pm\sigma$ cosmic and noise variance at 150 GHz on the sky area considered. At the bottom we show the average and standard deviation of the residuals on each realization. From Stivoli et al. (2006).

a value larger than 2 means that the separation error is dominating. d_l is a measure of the extra uncertainty introduced by the algorithm.

On the other hand, as we see in a moment, d_l was getting closer to 1 when the noise is increased. This is due to the fact that in some cases, the separation quality deterioration, caused by the increase of the noise level, proceeds at the slower rate than the noise level increase itself leading to a decrease in d_l. The r_l and a_l quantities are respectively the residual and the absolute residual of the recovered pseudo-power spectra in the single separation, averaged over 100 realizations. These quantities gave us a measure of a bias of the reconstruction, and thus were expected to be close to zero. Note that part of the differences in the numerators of r_l and a_l comes from the instrumental noise; therefore, in a highly noisy configuration, their value can become large not because the separation fails, but because of the high noise itself.

Especially if checked together for each case, these quantities allowed us to attempt to give a definition for a 'successful' separation, which is when FASTICA is able both to recover the CMB signal giving the value of d_l on one side and the values of r_l and a_l on the other close to, and less than unity, respectively. In tables 4.1, 4.2 and 4.3, we report the value of these quantities for some relevant multipoles as a function of the varying parameters.

We began varying the noise with respect to the simulated dataset considered in the previous Section. We found out the results to be quite stable up to $S/N = 1$. As it may be noted by looking at the first block of four rows in the tables, the algorithm performance, in terms of d_l, decreased mildly or remained constant, and decreases nearly linearly with the noise amplitude in terms of r_l and a_l. For values of the noise larger than the signal, the code started failing to reconstruct the sig-

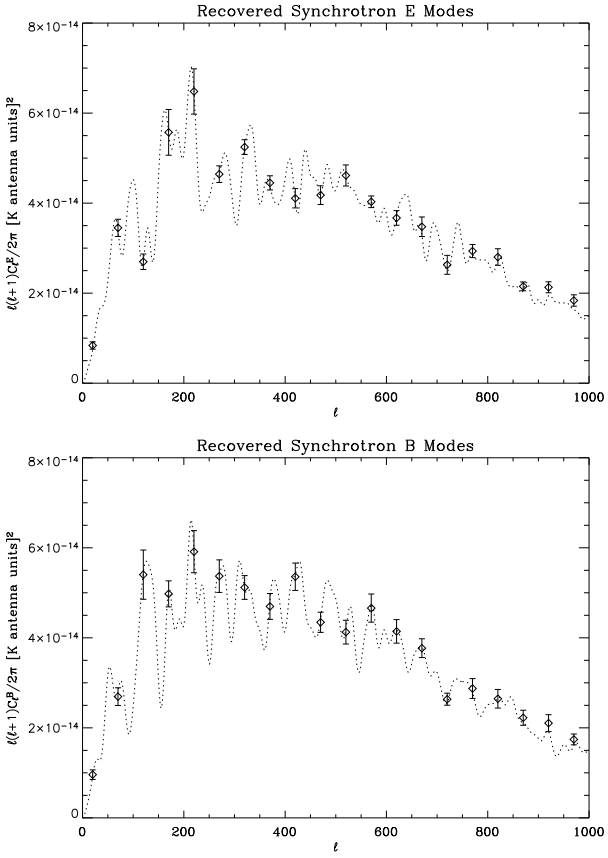

Figure 4.9: Pseudo-power spectra of the reconstructed $\tilde{C}_l^E$ modes (left) and $\tilde{C}_l^B$ (right) of synchrotron at 40 GHz in the $S/N = 2$ case. Dotted lines are the input spectra. From Stivoli et al. (2006).

nals, what at first showed as a residual foreground contamination persisting in the reconstructed B modes of the CMB, then as a failure to reach the convergence or to estimate a non-negative definite signal correlation matrix due to the large noise sample variance.
The foreground variation was realized by keeping its mean over the considered area unchanged and increasing solely its *rms* by a factor 2, 4, and 6 for synchrotron, and 2, 4, 6 and 10 for dust. In Figs. 4.10 and 4.11 we report the foreground B modes at 40, 90 GHz and 150, 350 GHz, respectively, for the *rms* considered. For reference, we also plotted the theoretical CMB pseudo-spectra. At 40 and 350 GHz, the contamination to the CMB is worse of course. Despite the high level of foreground fluctuations, the method exhibited again a remarkable stability or even improvement in the interval considered for this parameter, as it may be seen by looking at the second block of four rows in the tables; it starts failing only when the foreground *rms* is increased by a factor of about 6 for synchrotron, and by a factor of about 10 for dust. This can be interpreted as due to the fact that foreground recovery is indeed easier and more precise given a larger foreground amplitude. Indeed, for an ICA based component separation technique which utilizes the independence of the components to be recovered, the quality of the reconstruction of each of them depends on how well the other ones are extracted (Maino et al., 2002; Baccigalupi et al., 2004).
The last row in each table shows the effect of the variation of the sky area considered, while all the other parameters are kept fixed. As expected, things got generally better after doubling the radius of the cut, but since at a resolution of about 10 arc-minutes a patch with $\theta_C = 10^\circ$ has already a number of samples (pixels) large enough to faithfully represent the signal statistics, increasing θ_C doesn't improve the separation dramatically. However, a wider area represents a benefit concerning the possibility of detecting the B modes from primordial gravitational waves, as we discuss in Section 4.2.5.

As a final remark, we noticed that increasing the noise amplitude causes the reconstruction to be less accurate, in all cases when the dust is taken into account. This observation, anticipated in Section 4.2.3, is due to the fact that at 150 GHz, the dust emission is negligible with respect to CMB and noise. Indeed, as noticed in earlier works (Maino et al., 2002; Baccigalupi et al., 2004), the separation is more accurate when the signals are comparable in all frequency bands. This is supported by the fact that the performance improves or remains unaltered when the dust fluctuation amplitude is increased, while in the synchrotron case a clear degradation of the separation may be seen.
As a final test to evaluate the separation we studied the recovered frequency scaling index $\alpha = \log[s(\nu_2)/s(\nu_1))]/\log(\nu_2/\nu_1)$ of the different output components s, computed through the ratios between column elements in the inverse of the separation matrix (Maino et al., 2002). In all the cases we studied, this quantity resulted to be close to the theoretical one, with dispersions $\Delta\alpha$ increasing roughly linearly with foreground amplitude and noise; an exception was still represented by the

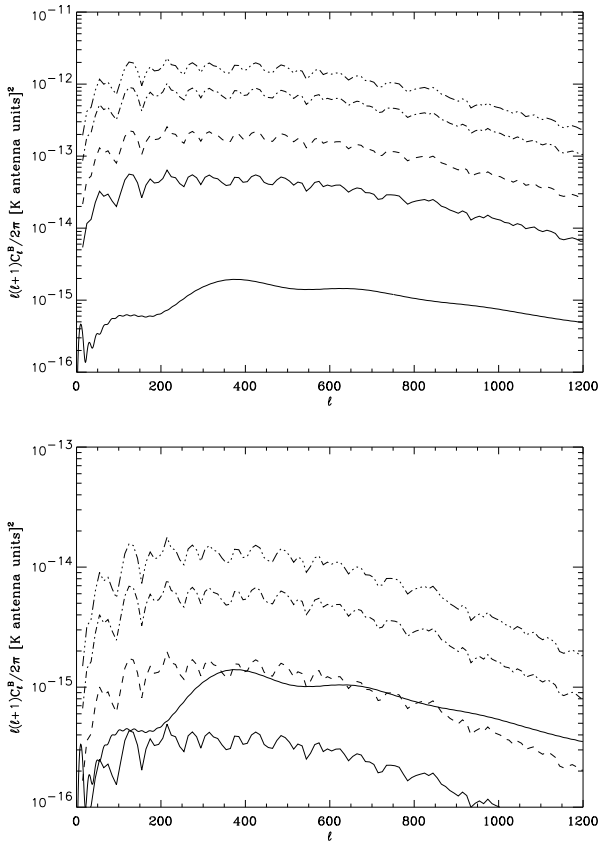

Figure 4.10: Pseudo-power spectra of the synchrotron B modes calculated for the sky region and with the amplitudes as considered in the text, at 40 (left) and 90 GHz (right). The different curves, with raising power, correspond to the foreground *rms* multiplied by 1, 2, 4, and 6, respectively. In each panel the solid smooth line represents the B modes of the CMB. From Stivoli et al. (2006).

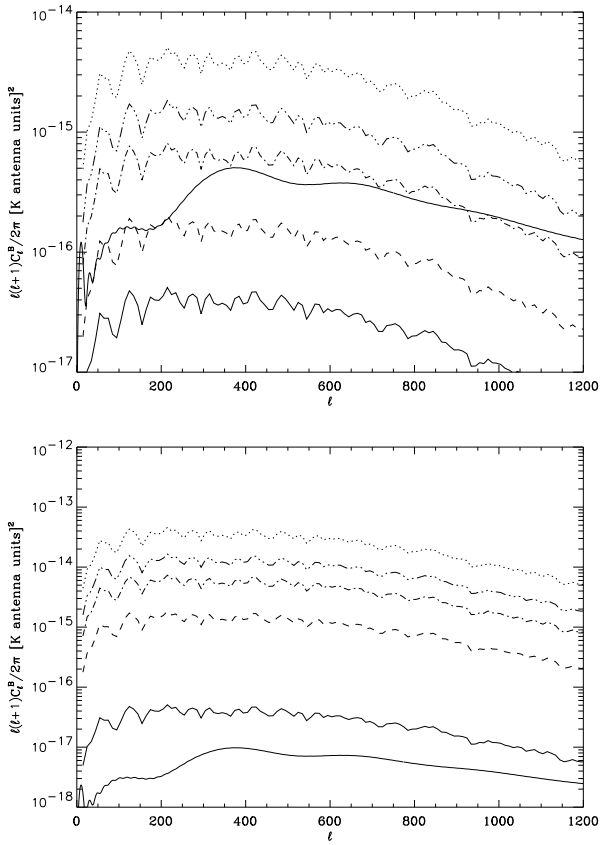

Figure 4.11: Pseudo-power spectra of the dust B modes calculated for the sky region and with the amplitudes as considered in the text, at 150 (left) and 350 GHz (right). The different curves, with raising power, correspond to the foreground *rms* multiplied by 1, 2, 4, 6 and 10, respectively. In each panel the solid smooth line represents the B modes of the CMB. From Stivoli et al. (2006).

S/N	Fore. Ampl.	Cut Radius	$d_{l=100}$ Sync.	$d_{l=400}$ Sync.	$d_{l=950}$ Sync.	$d_{l=100}$ Dust	$d_{l=400}$ Dust	$d_{l=950}$ Dust
∞	1.00	10	1.16	1.69	1.31	2.08	1.99	1.74
2.00	1.00	10	1.44	1.55	1.18	1.81	1.56	1.31
1.50	1.00	10	1.38	1.42	1.07	2.24	1.56	1.31
1.00	1.00	10	1.88	1.72	1.38	3.39	1.56	1.21
2.00	2.00	10	1.60	1.70	1.33	1.66	1.64	1.34
2.00	4.00	10	1.62	2.03	1.58	1.63	1.69	1.33
2.00	6.00	10	1.78	2.39	1.81	1.51	1.69	1.32
2.00	10.00	10	-	-	-	2.00	2.04	1.50
2.00	1.00	20	1.28	1.53	1.03	1.75	1.41	1.24

Table 4.1: Relative extra uncertainty, d_l, evaluated for the reconstructed B mode power spectrum of the CMB. The results are given for three different values of the multipole l and for multiple choices of the sky and noise parameters as listed in the table.

dust case, when the dust *rms* is increased: as explained above, increasing the dust *rms* induces an improvement in the reconstruction, which appears also in the recovery of α. The relative dispersion $\Delta\alpha/\alpha$, evaluated over the 100 Monte Carlo realizations both for CMB and foregrounds, is shown in table 4.4.

4.2.5 Measuring the primordial tensor amplitude

As we stressed already, one of the most important goals of the forthcoming CMB polarization experiments is the measure of the ratio r between the primordial amplitudes of tensor and scalar cosmological perturbations, i.e. gravity waves and density fluctuations. The most relevant question in this context is how small that ratio can be in order to be detected when foregrounds are taken into account, and in particular what this minimal detectable value is when the CMB background is separated from the foregrounds with the technique considered here. Unfortunately, the foreground simulations were (and are) still too uncertain to push the analysis toward a complete cosmological parameter estimation pipeline and address this question comprehensively. Nevertheless, given the importance of this topic, we present in this Section some general though preliminary remarks and we illustrate them with some examples.

For our purpose here the most relevant result of the previous Sections is the observation that the FASTICA-based separation yields errors which are comparable to those from cosmic variance and noise, for the model with $r = 0.1$. Therefore, in such a case, as far as these simulations are concerned, we should be able to detect the tensor contribution in the presence of foregrounds when the latter are treated with ICA, with a confidence close to the one achievable without foregrounds.

To illustrate this issue, we compared the pseudo B mode recovery in our fidu-

S/N	Fore. Ampl.	Cut Radius	$r_{l=100}$ Sync.	$r_{l=400}$ Sync.	$r_{l=950}$ Sync.	$r_{l=100}$ Dust	$r_{l=400}$ Dust	$r_{l=950}$ Dust
∞	1.00	10	0.02	0.05	0.05	0.10	0.05	0.05
2.00	1.00	10	0.01	0.06	0.08	0.11	0.06	0.07
1.50	1.00	10	0.02	0.06	0.10	0.13	0.06	0.17
1.00	1.00	10	0.02	0.08	0.22	0.22	0.08	0.20
2.00	2.00	10	0.01	0.06	0.08	0.08	0.05	0.05
2.00	4.00	10	0.02	0.08	0.14	0.08	0.06	0.06
2.00	6.00	10	0.02	0.08	0.14	0.08	0.06	0.03
2.00	10.00	10	-	-	-	0.08	0.08	0.05
2.00	1.00	20	0.02	0.04	0.03	0.09	0.05	0.03

Table 4.2: Residuals of the CMB pseudo B modes recovered against synchrotron and dust, averaged over 100 realizations of noise and CMB.

S/N	Fore. Ampl.	Cut Radius	$a_{l=100}$ Sync.	$a_{l=400}$ Sync.	$a_{l=950}$ Sync.	$a_{l=100}$ Dust	$a_{l=400}$ Dust	$a_{l=950}$ Dust
∞	1.00	10	0.12	0.06	0.06	0.14	0.06	0.05
2.00	1.00	10	0.12	0.07	0.11	0.15	0.07	0.14
1.50	1.00	10	0.12	0.08	0.17	0.17	0.07	0.17
1.00	1.00	10	0.12	0.11	0.46	0.22	0.10	0.29
2.00	2.00	10	0.09	0.07	0.14	0.13	0.06	0.13
2.00	4.00	10	0.10	0.11	0.18	0.13	0.07	0.14
2.00	6.00	10	0.10	0.11	0.19	0.13	0.08	0.14
2.00	10.00	10	-	-	-	0.13	0.09	0.13
2.00	1.00	20	0.09	0.07	0.09	0.09	0.05	0.08

Table 4.3: Absolute value of residuals of the CMB pseudo B modes recovered against synchrotron and dust, averaged over 100 realizations of noise and CMB.

S/N	Fore. Ampl.	Cut Radius	CMB vs. Sync.	Synchrotron	CMB vs. Dust	Dust
∞	1.00	10	0.01	0.03	0.72	0.81
2.00	1.00	10	0.05	0.07	0.60	0.89
1.50	1.00	10	0.14	0.15	0.75	1.09
1.00	1.00	10	0.28	0.35	1.17	1.97
2.00	2.00	10	0.14	0.14	0.63	0.69
2.00	4.00	10	0.22	0.23	0.85	0.56
2.00	6.00	10	0.35	0.36	1.49	0.53
2.00	10.00	10	-	-	3.04	0.46
2.00	1.00	20	< 0.01	0.02	0.45	0.73
1.00	1.00	20	0.18	0.20	-	-

Table 4.4: Relative dispersions $\Delta\alpha/\alpha$ around the expected values of the frequency spectral indices, for both CMB and foregrounds.

cial model with $r = 0.1$ with one in which the tensors are absent, $r = 0$. We focused on the spectral region where primordial tensors are most relevant, e.g. $l \simeq 100$. Of course, as a result of the leakage of E modes into B due to the limited sky coverage, also in the latter case the amplitude of E modes on these scales matters. We addressed this issue to some extent by considering the results for each value of r obtained for sky areas with different aperture. In Figure4.12 we plot the recovered pseudo B modes in these two cases, zooming on the relevant range of multipoles. Two different sizes for the sky cut were considered; the left and right panel refer to $\theta_C = 10^\circ$ and $\theta_C = 20^\circ$, respectively. In both panels the higher amplitude spectrum represents the model with $r = 0.1$. As we quoted above, we saw that in the entire interval, the separation error is comparable to the cosmic variance, in particular in the l-range, where the B modes from tensors have their main impact, i.e. $l \simeq 100$. The central role of the leakage from E modes is also evident. When the area gets smaller and the pollution consequently larger, the detection of B modes becomes harder. Points on the plot are the pseudo B modes recovered by the code against the synchrotron template, in the case with $S/N = 2$, for both the models (asterisks for $r = 0$ and diamonds $r = 0.1$).

Looking at the left panel of the figure, it is clear that, even if the algorithm did not introduce any extra uncertainty, we could not make any claim on primordial B modes detection with sufficient confidence, because the leakage is already too high for the 10 deg cut. The situation gets better for the $\theta_C = 20^\circ$ case (the right panel), where the two models seem to be distinguishable statistically by our method. For the latter case, more quantitative results are showed in table 4.5. To define whether or not the algorithm is able to distinguish between the two models, we compared the statistics of the recovered power spectra. In the second column, we report the percentage of recovered power spectra (for $r = 0$) that fall inside the

	$\tilde{C}^{out}_{r=0}$ vs $\tilde{C}^{out}_{r=0.1}$	$\tilde{C}^{in}_{r=0}$ vs $\tilde{C}^{in}_{r=0.1}$	$\tilde{C}^{out}_{r=0.1}$ vs $\tilde{C}^{out}_{r=0}$	$\tilde{C}^{in}_{r=0.1}$ vs $\tilde{C}^{in}_{r=0}$
$\ell = 70$	0%	0%	0%	0%
$\ell = 120$	0%	0%	5%	2%
$\ell = 170$	50%	43%	53%	46%

Table 4.5: This table reports the capability of the algorithm to distinguish between the two models with and without the gravity wave content for the $\theta_C = 20^\circ$ case. The second column shows the percentage (see text for more details) of spurious detection of the tensor contribution, while the fourth column shows the percentage of a false non-detection. The other two columns report the same quantities but derived in the ideal, CMB-only, foreground-free case.

95% confidence region for $r = 0.1$ as calculated directly from the statistics of the recovered power spectra them-self. These are the power spectra that provide a false detection of tensor contribution. Vice-versa, the fourth column shows how many of them miss a true presence of primordial B modes, since it reports the percentage of recovered power spectra for $r = 0.1$ that fall inside the 95% confidence region for $r = 0$. Third and fifth column give analogous numbers but computed for simulated, CMB-only spectra and thus not requiring any further processing. These two columns provide an idea of the best achievable levels.

We thus conclude that, given the available foreground simulations, FASTICA eliminates the foregrounds in the two cases of r we consider, with a precision sufficient to make them distinguishable even with the suboptimal pseudo BB estimator. As already stressed above, this is due to the fact that the separation process induces an error comparable to those coming from cosmic variance and noise.

Further analysis on the recovering of the true B modes and, mostly important, on the minimum value of r that can be detected with this technique, could be performed (for problems relevant to this issue and what can be expected for the considered cases, see e.g. Lewis (2002, 2003) and reference therein). However they could be misleading given the substantial uncertainties that still affect the foreground simulations and because of the absence of systematics in this simulation. We actually plan to address these issues in a future work, probably exploiting the new *pure* power spectrum estimator from Smith (2006), in a new, more realistic simulated environment.

4.3 A parametric approach for EBEX

I report here an application of the parametric approach to the foreground cleaning originally introduced by Brandt et al. (1994) and further extended in total intensity by Eriksen et al. (2006) in the context of CMB observations aiming at the measurement of the cosmological B modes. It has to be noted that this work was done before the full analytic treatment by Stompor et al. (2008a) and it can be consid-

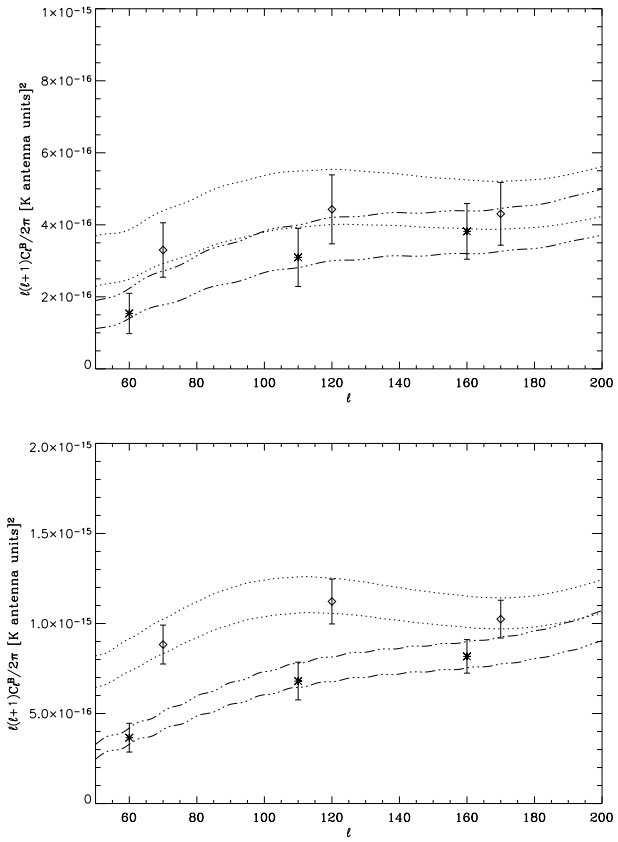

Figure 4.12: Comparison of the recovered pseudo B modes in two different cosmological models with $r = 0$ (dashed-dotted lines and asterisks) and $r = 0.1$ (dotted lines and diamonds shifted by $\Delta l = 10$ for clarity), in the low frequency combination with $S/N = 2$. Left panel refers to $\theta_C = 10^\circ$, while right panel to $\theta_C = 20^\circ$. Regions between the lines are the dispersion coming from cosmic variance and noise. From Stivoli et al. (2006).

ered a path finder for it. It has been also used in the successful proposal submitted by the EBEX team in 2007. An exhaustive and updated report about this work is in preparation (Stivoli et al., 2008).

As introduced in the previous Chapter, the parametric approach gives up completely on the information stored in the spatial correlation or statistics of the different signals to recover, fitting for the unknowns on a pixel by pixel basis. On the other hand, the built-in parameterization procedure allows to cast the separation problem in a coherent maximum likelihood form, quantifying uncertainties, achieving simplifications or even semi-analytic solutions after marginalization over sub-set of parameters, as well as taking into account a number of realistic effects, like the directionality dependence of the frequency scaling of foregrounds and the presence of offsets and calibration errors.

4.3.1 Simulated Sky

In this Section we describe how we simulate the sky emission in which we apply our foreground cleaning technique. Since this is a more recent work, it should be noted that the foreground models used here are an updated version with respect to the ones used in Sec.4.2.1.
Similarly to the previous Section, our analysis concerns the patch centered at RA=60°, DEC=-50°, known to be characterized by low foreground emission (Montroy et al., 2005) and which is the target of EBEX (Oxley et al., 2004); for simplicity, we take it circular with 350 squared degrees area, and we refer to it in the following as the EBEX patch.
The only source of data on polarized foregrounds at high Galactic latitudes in the microwave band is represented by WMAP (Page et al., 2007), covering the frequency range between 22 and 94 GHz. Outside the P06 mask, cutting out about 25% of the sky across the Galactic plane, the diffuse foreground emission is observed on angular scales corresponding to the degree or larger. The data are consistent with the physical expectations. The component of the polarized Galactic emission which is rising in frequency is given by the thermal vibration of dust grains, inducing a black body characterized by a temperature which is roughly between 10 and 20 K across the sky, corrected because of the emission of stars, which cause a steeper rise in frequency with respect to a pure black body. In total intensity, the Galactic thermal dust emission has been observed with high accuracy and angular resolution at about 3000 GHz Finkbeiner et al. (1999); due to their magnetic moment, the grains get locally aligned with the Galactic magnetic field (Jones et al., 1995), causing a linear polarization in the emitted thermal radiation. The polarized intensity is found to be about 10% of the total one, consistently with the findings at low Galactic latitudes by Archeops (Ponthieu, 2005). The frequency range in which the synchrotron and dust emission become equal is between 60 and 94 GHz (Page et al., 2007).

The polarized dust emission is the only foreground of interest in this study, since we focus on high frequencies, larger than 150 GHz, as this is the range cov-

ered by the EBEX channels. The dust is not observed in polarization in the area we consider, and we adopt the following simulation strategy. First, we assume that the dust polarized intensity is a constant fraction p of the total one. Second, in order to construct a model for the Stokes parameters $Q = P\cos\theta$ and $U = P\sin\theta$ where θ is the polarization angle, we simulate an all sky Gaussian template for the polarization angle using the `anafast` code within HEALPix, following a strategy already adopted by Giardino et al. (2002): we generate two realizations r_1 and r_2 of a power spectrum given by $C_l = \propto l^{\beta}$, where β is constant, and pose $\cos\theta = r_1/\sqrt{(r_1^2 + r_2^2)}$, $\sin\theta = r_2/\sqrt{(r_1^2 + r_2^2)}$. Finally, as for the Galaxy the power spectra for Q, U, E and B are expected (Zaldarriaga, 2001) and found (Page et al., 2007) to be comparable, we determine the constants p and β matching their power spectra with the ones estimated from WMAP outside the P06 mask, Eq.(25) of Page et al. (2007). The match is found at 65 GHz, scaling the dust total intensity down from 3000 GHz according to the two component modified black body corresponding to the model 8 of Finkbeiner et al. (1999), substantially consistent with the WMAP data (Bennett et al., 2003) and is

$$\text{dust flux} \propto \frac{f_1(q_1/q_2)(\nu/3000\ \text{GHz})^{\alpha_1+3}}{e^{h\nu/kT_1} - 1} + \frac{f_2(\nu/3000\ \text{GHz})^{\alpha_2+3}}{e^{h\nu/kT_1} - 1}\,, \qquad (4.0)$$

where $f_1 = 0.0363$, $f_2 = 1 - f_1$, $\alpha_1 = 1.67$, $\alpha_2 = 2.70$, $q_1/q_2 = 13$, $T_1 = 9.4$ K, $T_2 = 16.2$ K, and h and k are the Planck and Boltzmann constants, respectively. For the polarization fraction, we find good agreement for $p = 11\%$, consistently with the Archeops findings on the plane, which are slightly lower, 5%, because of cancellation effects along the line of sight (Ponthieu, 2005). For what concerns the slope of the polarization angle power spectrum, we obtain only an upper limit, $\beta \simeq -3.2$, due to the fact that for higher values the fluctuations of the polarization angle, plus the one in total intensity of the thermal dust, cause an excess of fluctuating power in the shape of the power spectrum with respect to WMAP. Note that this corresponds to a rather severe damping of the polarization angle fluctuations on small angular scales with respect to the models adopted before, as in Stivoli et al. (2006); this implies that on the scales probed by patches with an extension comparable to the EBEX one, the polarization angle is essentially uniform compared with the fluctuations of the total intensity. In other words, in the EBEX patch, according to our modeling, the Q and U templates are essentially given by the total intensity of the dust in that region, times a polarization fraction $p = 11\%$, times the large scale modulation given by the almost constant value of $\cos\theta$ and $\sin\theta$ in that region. Since, as we know, this component separation algorithm does not depend on global properties like the average of signals on the area considered, that means that our results are rather independent on the direction of the polarization angle in the EBEX patch. Finally, we checked that the variation of the dust frequency scalings across the sky, interpreted mainly at fluctuations in the temperatures of its components (Finkbeiner et al., 1999), does not lead to a significant effect on the EBEX patch; the present knowledge of this feature, from the correlation of data

at 100 and 240 μm, indicates a relative fluctuation of the frequency scaling across the EBEX patch corresponding to the percent or less, which is subdominant with respect to the noise level considered in this work.

The CMB emission is assumed to be Gaussian, simulated using the `synfast` routine of HEALPix, out of a power spectrum generated with the public Code for Anisotropies in the Microwave Background (CAMB[1]) with Hubble expansion rate $H_0 = 70.8$ km/sec/Mpc, optical depth at re-ionization $\tau = 0.07$, cosmological abundances given by $\Omega_b = 4.4\%$, $\Omega_{cdm} = 22\%$, $\Omega_\Lambda = 1 - \Omega_b - \Omega_{cdm}$, representing the energy densities of baryons plus leptons, cold dark matter, cosmological constant, respectively, divided by the critical one; the primordial power spectrum is assumed to possess a scalar spectral index $n_s = -0.94$, a tensor one $n_t = -0.015$, corresponding to the simplest single field inflationary models. with a tensor amplitude corresponding to 10% of the scalar one, $T/S = 0.1$. The overall normalization power has been determined using the WMAP data in total intensity and correlation with the E mode of polarization. This set of cosmological parameters represents that of the cosmological concordance model apart from tensors, which are at present limited from above, $T/S \leq 30\%$, (Spergel et al., 2007).

In Fig.4.13, top panels, we show the pattern of the Q Stokes parameter for the dust (left) and the CMB (right). With respect to the CMB, fluctuating mostly on the degree scale, the dust has considerable power on larger angular scales, and follows closely the total intensity pattern. In the lower panel we show the power spectrum of the dust, where for simplicity the oscillations given by the underlying realization have been averaged out, compared with the E and B modes of the CMB at various frequencies, using again the Finkbeiner et al. (1999) recipe for scaling the dust. As it is evident, the CMB is dominating up to the highest frequency through the E modes. On the other hand, the curl component of CMB anisotropies, i.e. the B modes, is always sub-dominating or at least highly contaminated.

4.3.2 Data modeling

In the following application we applied the Eriksen et al. (2006) approach described in the last Chapter. This work was done before the improvements by Stompor et al. (2008a).
The signal in the EBEX patch is pixelized with HEALPix, at $n_{side} = 1024$, corresponding to a resolution of about 3.4 arc-minutes. We take the frequencies and nominal noise from the EBEX experiment as reference (Oxley et al., 2004). The signal is smoothed with a circular beam corresponding to a full width half maximum (FWHM) of 8 arc-minutes. The noise rms is 1.43, 3.96, 43.1 thermodynamic μK at 150, 250, 410 GHz, respectively. The sky signal is M_ν a mixture of CMB and dust emission, where the dust is normalized to have the same rms at 150 GHz, and then scaled in frequency as (4.3.1). Therefore, we write

$$M_\nu = M_\nu^{cmb} + M_\nu^{dust} + n_\nu \tag{4.0}$$

[1] See `camb.info`.

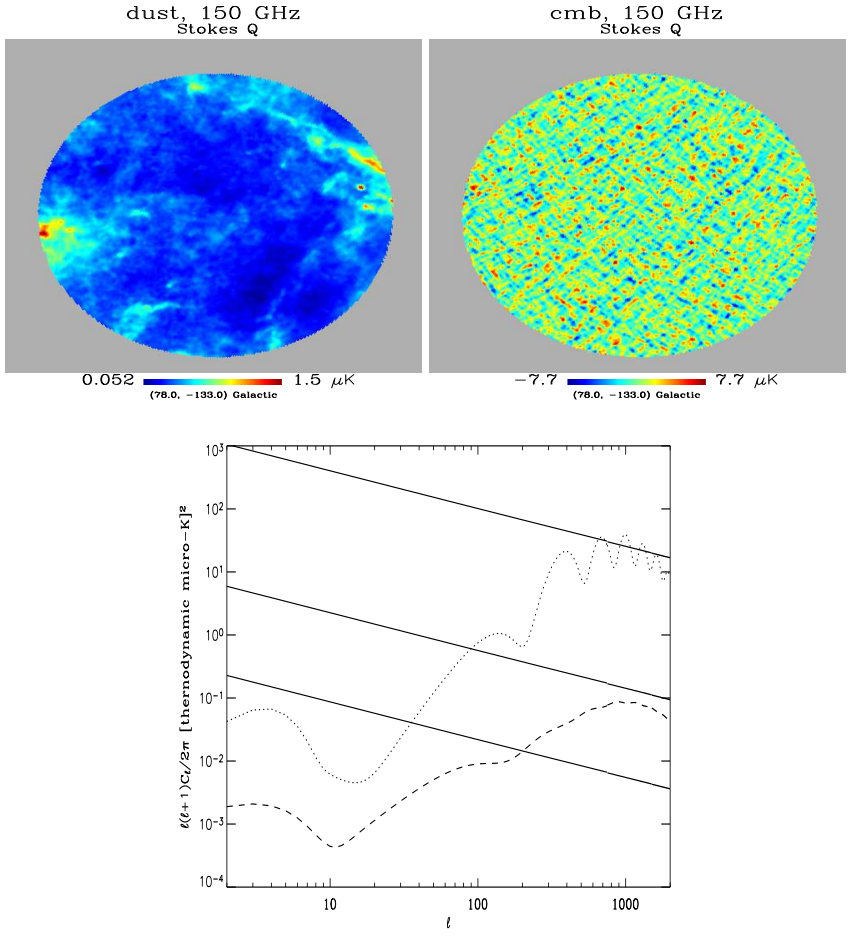

Figure 4.13: Top panels: a representation of the simulated dust (left) and CMB (right) Q Stokes parameter patterns, in antenna units at 150 GHz. Lower panel: the angular power spectra for the CMB B (dashed) and E (dotted) modes, in comparison with the level of foreground contamination in the EBEX patch, at 150 (lower straight curve), 250 (middle straight curve) and 410 (higher straight curve) GHz.

where

$$M_\nu^{cmb} = s_c \frac{\left(\frac{h\nu}{kT_o}\right)^2 e^{\frac{h\nu}{kT_o}}}{(e^{\frac{h\nu}{kT_o}} - 1)^2} \quad , \quad T_o = 2.725K \, , \tag{4.0}$$

and

$$M_\nu^{dust} = s_d F(\nu) \tag{4.0}$$

where F represents the parameterized dust frequency scaling, different from the one adopted to simulate the sky in all three cases as we discuss in a moment, and the two quantities s_c and s_d are the CMB and dust amplitudes on a given pixel, respectively. The component separation consists in estimating the amplitudes such that they are close to the real ones, $s_c \simeq \bar{s_c}$ and $s_d \simeq \bar{s_d}$, as well as getting the unknowns in the dust frequency scaling, along with all the corresponding uncertainties. As we stressed already, the latter has been simulated using model 8 of Finkbeiner et al. (1999), which corresponds to a two dust grain component, with relative abundances and different emissivities as in (4.3.1). This means more than three parameters, so that they cannot be estimated using the present technique. On the other hand, on our frequency range, the spectrum may be parameterized efficiently in several simplified ways. The ones adopted here are

$$F_1(\nu) = (\frac{\nu}{\nu_0})^\beta \quad , \tag{4.0}$$

$$F_2(\nu) = \frac{\nu}{e^{\frac{h\nu}{kT_2}} - 1} \frac{e^{\frac{h\nu_0}{kT_2}} - 1}{\nu_0} \left(\frac{\nu}{\nu_0}\right)^\beta \quad , \quad T_2 = 18.1\text{K} \, , \tag{4.0}$$

$$F_3(\nu) = \left(\frac{\nu}{\nu_0}\right)^{(\beta+1)} \left[\frac{1/(e^{h\nu/kT_3} - 1) + 6.7/(e^{h\nu/kT_4} - 1)}{1/(e^{h\nu_0/kT_3} - 1) + 6.7/(e^{h\nu_0/kT_4} - 1)}\right] ,$$

$$T_3 = 20.4 \text{ K} \, , \; T_4 = 4.77 \text{ K} \, . \tag{4.-1}$$

with $\nu_0 = 150GHz$. They correspond respectively to a simple power law and two versions of a modified black body, respectively model 3 by Finkbeiner et al. (1999) and Wright et al. (1991), containing one parameter only, β, which together with s_c and s_d represent the set of quantities to be recovered in each pixel.

Spectral index recovery

The model $M(\theta)$, where $\theta = (s_c, s_d, \beta)$, can be rewritten as

$$M_p(\theta) = A_p s_p \, , \tag{4.-1}$$

where we explicitly wrote the dependence on the pixel p. Here $A_p = A_p(\beta)$ is an estimation of the mixing matrix $\bar{A}$, modeled with a single unknown parameter β parameterizing the dust frequency scaling with one of the three recipes (4.3.2,

4.3.2, 4.1) and $s_p = (s_c, s_d)$ represents the signal amplitudes on the pixel. As showed in the last Chapter, Eq.(3.2), the logarithmic likelihood on the pixel can be written as:

$$-2 \ln \mathcal{L}\left(\mathbf{s_p}, \beta_\mathbf{i}\right) = \mathrm{K} + \left(\mathbf{d_p} - \mathbf{A_p}\,\mathbf{s_p}\right)^t \mathbf{N_p^{-1}} \left(\mathbf{d_p} - \mathbf{A_p}\,\mathbf{s_p}\right). \tag{4.0}$$

Integrating this relation over pixel amplitudes s_c and s_d, we get Eq.(3.5), that we report here for simplicity:

$$\begin{aligned} -2 \ln \mathcal{L}_{marg}(\beta) &= -2 \ln \int d\,\mathbf{s} \exp\left[-\frac{\mathbf{1}}{\mathbf{2}} \left(\mathbf{d} - \mathbf{A}\,\mathbf{s}\right)^\mathbf{t} \mathbf{N^{-1}} \left(\mathbf{d} - \mathbf{A}\,\mathbf{s}\right)\right] \\ &= \mathrm{K} - \left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{d}\right)^t \left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}\right)^{-1} \left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{d}\right) \\ &+ \ln\left|\left(\mathbf{A^t}\,\mathbf{N^{-1}}\,\mathbf{A}\right)^{-1}\right|, \end{aligned} \tag{4.-1}$$

The likelihood above, function of β only, provides an estimation for the one dimensional marginalized probability distribution of β. The minimization proceeds semi-analytically, yielding the solution for β. It is to be noted here that this Equation is the biased estimator that we corrected in Stompor et al. (2008a). However, in the following application, it provides an estimation of the spectral index that is good enough for our purposes.
We will now compare the solution which is achieved in this way with the one coming out of the MCMC procedure. Beside the bias, Equation (4.-1) has the advantage of being in principle numerically stable with respect to noisy data and faster than the MCMC. On the other hand it has the intrinsic limitation of being usable only when a solution can be computed, i.e. when a friendly enough likelihood is used, like Gaussian for example. For now we concentrate on the results obtained with the parameterization (4.3.2). Before concluding this Section we will show how the results are analogous using the other dust models.

Recovering the frequency scaling parameter β via MCMC requires decreasing the resolution in order to increase the signal to noise ratio (Brandt et al., 1994). In our case, a configuration which works consists in pixels with 4 squared degrees area. In the EBEX patch, this corresponds to 12 low resolution pixels in the map. This means that if the frequency scaling is varying in different sky directions on scales smaller than 4 squared degrees, the MCMC approach is unable to recover that information. We plot the marginalized two dimensional probability distributions of the recovered parameters on one of the large pixels in Fig.4.14 (top and bottom left panels). We stress that the CMB and dust amplitudes here represent the ones on the low resolution pixels. In the bottom right panel we marginalize on both amplitudes. The dashed solid line represents the value $\bar{\beta}$ for which the scaling with (4.3.2) corresponds exactly to the one actually used to simulate the sky, Equation (4.3.1), between 150 and 410 GHz. The right value is recovered within 1σ. The plots for the other large pixels in the EBEX patch are analogous. In Fig.4.15 we

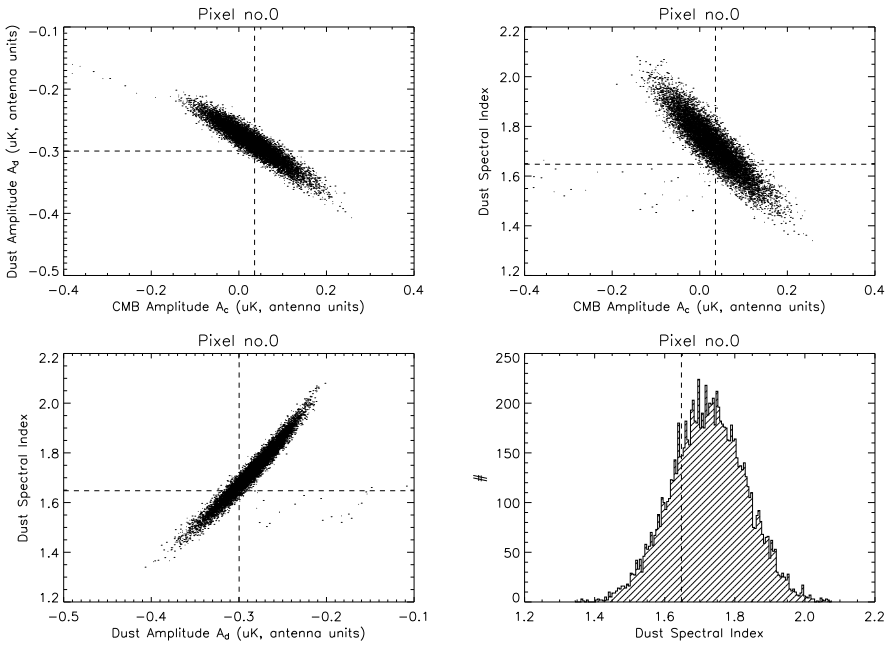

Figure 4.14: MCMC results for a representative pixel. Top panels and bottom left panel: marginalized two-dimensional probability distributions. Bottom right panel: spectral index probability distribution, marginalized over the amplitudes. The dashed straight lines are the input values of the parameters.

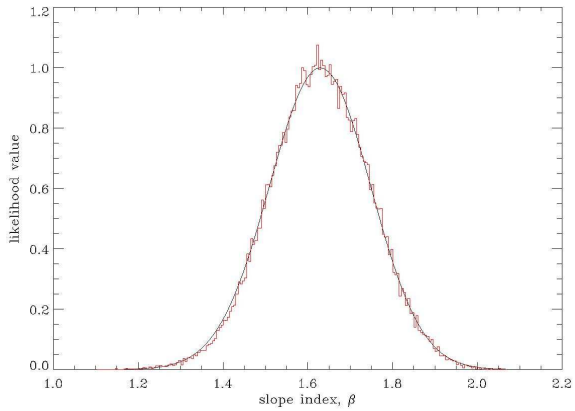

Figure 4.15: Comparison between the spectral index probability distribution obtained with the marginalization procedure (4.-1) and with the MCMCs described in the text.

show the comparison between the MCMC estimated probability distribution for β versus the one obtained from the semi-analytic procedure mentioned above (4.-1), showing substantial agreement, the bias being too small in this case. We stress that there is no need for decreasing the resolution if the unbiased marginalization procedure (3.7) is adopted.

Finally, we briefly discuss the results obtained using the other parameterizations, (4.3.2, 4.1). In Fig.4.16 we plot the dust rms for Q as a function of frequency according to the best fits on β on all parameterizations (4.3.2, 4.3.2, 4.1) versus the scaling (4.3.1), which was used in the actual sky simulation; the error bars, which have been artificially reduced by a factor of 10 and attached to the dust curve for display reasons, represent the error in each pixel. The decreasing line represents the CMB rms. The latter is dominating except at 410 GHz because the emission is dominated by the E modes, see Fig.4.13. This plot shows that the different parameterizations are essentially equivalent in the frequency range of interest with respect to the noise, as well as that the fitting procedure is effective in all cases. From now on, we only consider the parameterization given by the model 3 by Finkbeiner, Eq.(4.3.2).

Amplitudes

Once the distribution function for the frequency scaling has been obtained, we are left with the problem of solving a linear system for the two amplitudes, which in the present case is still analytic. This can be done now at full resolution, either if the frequency scaling recovery has been performed with an MCMC or semi-analytically. Given our best value for the spectral index β_0, that leads to $A \equiv A(\beta_0)$, we get for each component x, pixel by pixel:

$$\mathbf{s} = \left(\mathbf{A^t N^{-1} A}\right)^{-1} \mathbf{A^t N^{-1} d}. \qquad (4.-1)$$

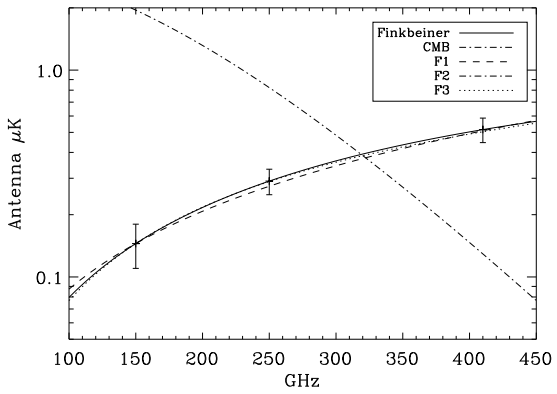

Figure 4.16: Comparison between the frequency behavior of the dust rms in Q obtained with the different parameterizations (4.3.2,4.3.2,4.1), versus the one actually used to simulate the sky. The CMB rms is also shown, together with the error bars in each pixels, divided by a factor of 10 for display reasons, and attached to the dust curve in order to make it evident how the difference in parameterization is irrelevant in the fitting procedure.

Note that due to our simplifying assumption the noise in the recovered component maps is Gaussian and uncorrelated.

The results are shown in Fig.4.17. The top panels show the CMB residuals, input minus output, for Q (left) and U (right). The middle panel show the dust residuals. For comparison and quantification of the achieved precision, the lower panels show the noise realization. The first and most remarkable feature is the absence of large scale structure in the residuals, which behave as an additional and effective noise component; as we see in a moment, this is crucial in order to remove the large scale contamination of the dust in the CMB B modes which is apparent in Fig.4.13. Second, the level of precision is comparable with the input noise, as it is expected since the latter is the real limiting factor in a pixel by pixel analysis.

A proper error estimation on the component amplitudes was not performed at the time of this analysis. As pointed out in the last Chapter it could be done analytically exploiting Eq.(3.9). Here we estimated the uncertainty in the component separation performing a small Monte-Carlo varying the CMB and the noise realizations. The uncertainty on the spectral index estimation was propagated to the second step of the analysis by drawing 10 random values for β out of its one dimensional marginalized posterior. As a figure of merit for the separation method, we used the B-mode power spectrum. Figure 4.18 shows the predicted EBEX errors on the determination of CMB B-mode polarization (stars) including errors from foreground subtraction. Errors on the CMB signal show less than $1/3$ increase over binned cosmic variance and instrument noise (shaded red) for $\ell < 900$. Error bars have been obtained on ten map-domain simulations varying CMB and instrument noise. This figure was used in the successful EBEX proposal in 2007.

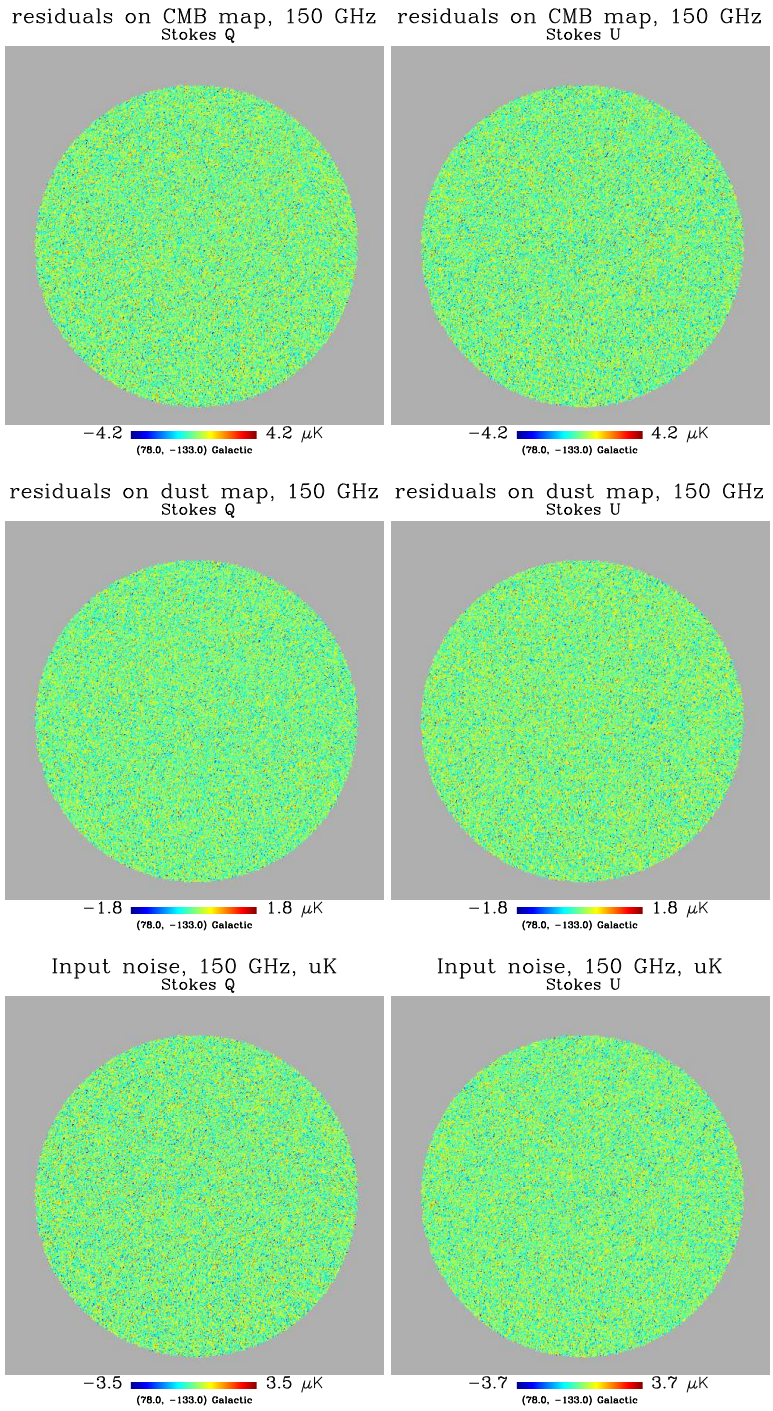

Figure 4.17: Separation residuals for Q (left) and U (right), CMB (top) and dust (middle). The noise realization is shown in the bottom panels. The separation residuals appear free of any large scale structure from any of the CMB or dust components, behaving as white noise.

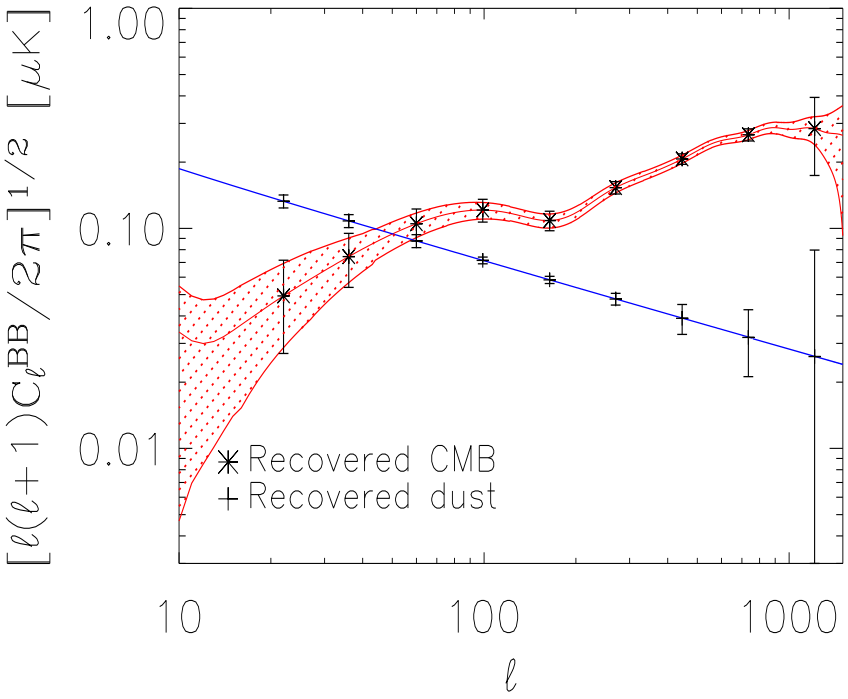

Figure 4.18: CMB and dust recovered B modes on the EBEX patch. Error bars are computed on 10 different realizations of CMB and noise.

Conclusions

One of the major challenges to the present and future observations of the Cosmic Microwave Background (CMB) anisotropies is represented by the foreground emission. The latest data on all sky by the Wilkinson Microwave Anisotropy Probe (WMAP) satellite confirmed that the diffuse emission from our own Galaxy must be controlled and possibly removed in total intensity and even more in polarization in order to access the cosmological information contained in the CMB anisotropies.

In this work, we developed and applied some of the most advanced existing techniques for the separation of background and foregrounds in multi-frequency CMB experiments to simulated and real data, with a special regard to the extraction of the curl (B) component of the CMB polarization, which might contain the signal from primordial gravitational waves in the anisotropies on degree angular scale.

We focused the discussion on three methods: Fast Independent Component Analysis (FASTICA), a blind algorithm that exploits the statistics of the signals in order to achieve separation, analyzing the signal in all resolution elements (pixels) at once, a parametric approach, implemented in a code called MIRAMARE, designed to cast and maximize the data likelihood on each pixel singularly by means of a suitable parameterization of foreground and background unknowns, and the Correlated Component Analysis (CCA) that uses spatial correlations of subsets of the data to constrain the spectral parameters of the mixed components.

The first of these methods, FASTICA, recently underwent major theoretical developments. In the work by Stompor et al. (2008a), we achieved two important results; the code is now able to optimally deal with redundant noisy data, via a lossless input data compression, as well as to include the CMB frequency scaling as a prior, showing consistent improvements in those cases in which the data set is not large enough to properly infer the statistics of the signals.

Several application of FASTICA have been described. First of all, the work we have done on the WMAP 3-year data (Maino et al., 2007). In that study, most of the analysis have been carried out on the reconstructed CMB power spectrum that we managed to recover up to degree angular scale. Taking advantage of K band, the lowest for WMAP and the best tracer so far for assessing the CMB contamination from low frequencies, among other results we confirmed the north-south sky asymmetry, already noticed independently by several authors, showing the reliability of the algorithm when exploited to reconstruct the finest structure up to date in the CMB pattern.

A second application of FASTICA has been done in the context of the Working Group 2 within the Planck collaboration, the group dedicated to the study and testing of the component separation techniques. In a work recently submitted by Leach et al. (2008), we reviewed the results of a challenge among all the main component separation techniques today available within the Planck collaboration. FASTICA was capable of providing the lowest residuals in the CMB map at high Galactic latitudes, as well as recovering a remarkably clean dust template and to extract the TT power spectrum up to the 6-th acoustic peak in the CMB anisotropies.

Finally, we applied FASTICA on a simulated data set, specifically tailored to describe a typical CMB sub-orbital experiment, aiming at the detection of the B modes (Stivoli et al., 2006). This was one of the first works specifically dedicated to the recovery of B modes in component separation. More specifically, the aim was to test the capabilities of the algorithm to recover the B modes observed on a limited patch of the sky in the presence of a large foreground contamination. The code proved to be quite stable for interesting values of B modes signal with respect to variations in the foreground modeling as well as instrumental noise level, showing a limited deterioration in the quality of the outputs with respect to an ideal separation. The quantification of the additional error introduced by the foreground cleaning stage was indeed one of the main results of that work.

For what concerns the parametric approach, starting from a pretty advanced formulation of the method, in the work by Stompor et al. (2008b) we implemented a new version of the algorithm, casting it in maximum likelihood framework and healing a few problems which were present in the previous, related approaches. In this context, our work has been mostly theoretical, exploiting the possibilities of the method in a semi-analytical approach. Parameterization in component separation is particularly suitable for an explicit treatment of the separation error and for modeling instrumental systematics, as calibration errors and offsets, directly in the likelihood.

We applied this technique to a simulated data set related to one particular sub-orbital probe in which we are directly involved, the E and B EXperiment (EBEX), in order to show the capabilities of the instrument to detect the B modes in the Antarctica flight scheduled for the end of 2009. In the simplified set up where no instrumental systematics were considered, the results were excellent, and were submitted as part of the successful proposal to NASA in 2007 for funding the experiment up to the final stages of the data analysis.

The last application we considered in this thesis is the one we performed in Bonaldi et al. (2007) on the WMAP 3-year data set. Given the particular nature of the algorithm, this analysis, differently by the previous ones, was more foreground-oriented. We achieved two important results. First, we provided a full sky map of the anomalous emission, a dust correlated component that still evades a full understanding. Second, we studied the impact of the foreground modeling in the estimation of the low multipoles in the TT power spectrum, showing that extra care has to be used in addressing the error bar to the quadrupole of the CMB.

We finally write down some concluding remarks here. First of all, we point out

that, as for most of the cases when data analysis is concerned, the works presented in this thesis are parts of a large and collaborative ongoing work. The results summarized above represent major steps forward in the understanding of the relative techniques, in view of their application to most sophisticate simulations and most importantly to real data, but much more has still to come. The next use of FASTICA and CCA will probably be the analysis on the WMAP polarized data together with the application on more and more realistic simulations on Planck, in preparation for the true data analysis of the Planck data themselves. For what concerns the parametric approach, now that it is fully implemented in the MIRAMARE code, it is likely to follow the same path of the other ones. It will be applied to the WMAP data set and on the sub-orbital and Planck simulations. Thanks to its potential and the promising results, it will probably be the baseline for the foreground cleaning in the EBEX analysis.

Efficient and robust foreground cleaning is a requirement for the prosecution of CMB related science. It is becoming an independent and articulated sector of CMB data analysis, where a fruitful input from signal processing science is exploited. Probably, its most important application of this area of CMB data analysis is the establishment of the minimum detectable level of primordial gravitational waves, encoded in the B modes of the CMB polarization anisotropy, given the level of Galactic foreground emission. Although a definitive answer to this issue will have to wait until theoretical investigations match real data, the results contained in thesis contain the first steps in this direction.

Bibliography

Acquavia V., et al., 2006, *Phys. Rev. D* 74, 510

Alpher R. A., Bethe H. A., Gamow G., 1948, Physical Review, 73 , 803

Ashdown M., et al., 2006, *Astron. & Astrophys.* 476, 741

Baccigalupi C., et al., 2000, *MNRAS* 318, 769

Baccigalupi C., 2003, *New Astron. Rev.* 47, 1127

Baccigalupi C., et al., 2004, *MNRAS* 355, 5570

Banday A.J., et al., 2003, *MNRAS* 345, 897

Bedini L., et al. Journal on Applied Signal Processing, 2005, 14, 2400

Bennett C. L., et al., 2003, *Astrophys. J. Supp.* , 148, 97

Benoit A., et al., 2004, *Astron. & Astrophys.* 424, 571

Bobin J., et al. 2007, IEEE Transaction on Image Process, in Press

Bonaldi A., Bedini L., Salerno E., Baccigalupi C., de Zotti G., 2006, *MNRAS* , 373, 263

Bonaldi A., Ricciardi S., Leach S., Stivoli F., Baccigalupi C., de Zotti G., 2007, *MNRAS* , 382, 1791B

Bowden M. et al., 2004, *MNRAS* 349, 321

Brandt W. N., et al., 1994, *Astrophys. J.* , 424, 1

Carretti E., Bernardi G., Sault R.J., Cortiglioni S., Poppi S., 2005, *MNRAS* , 358, 1

Chon G., Challinor A., Prunet S., Hivon E. Szapudi I., 2004, *MNRAS* 350, 914

Cover T. M., Thomas J. A., *Elements of Information Theory*. Wiley.

Delabrouille J., et al., 2003, *MNRAS* , 347, 1089

Davies R. J., et al., 2006, *MNRAS* , 370, 1125

de Bernardis P., et al., 2000, *Nature* , 404, 955

de Oliveira-Costa A., et al., 2004, *Astrophys. J.* , 606, 89

Dennison B., Simonetti J. H., Topasna G. H., 1998, Pub. Astron. Soc. Aust., 15, 147, 1998.

Dickinson C., Davies R. D., Davis R. J., 2003, *MNRAS* , 341, 369

Dodelson S., 2003, Modern Cosmology, Academic Pr.

Donzelli S. et al., 2006, *MNRAS* , 369, 441

Draine B. T., Lee H. M., 1984, *Astrophys. J.* 285, 89

Draine B. T., Lee H. M., 1998, *Astrophys. J.* 508, 157

Duncan A. R., Reich P., Reich W., Fürst E., 1999, *Astron. & Astrophys.* 350, 447

Dunkley J., et al., 2008, submitted to *Astrophys. J.*

Eriksen H. K., et al., 2006, *Astrophys. J.* , 641, 665

Eriksen H. K., et al., 2008, *Astrophys. J.* , 676, 10

Finkbeiner D. P., Davis M., Schlegel, D. J., 1999 *Astrophys. J.* , 524, 867

Finkbeiner D. P., 2003 *Astrophys. J.* , 146, 407

Gabor D., 1946, *J. Inst. Elect. Eng.* 93, 420

Gaustad J. E., et al., 2001, PASP, 113, 1326

Giardino G., et al., 2002, *Astron. & Astrophys.* 387, 82

Górski K. M., et al., 2005, *Astrophys. J.* , 622, 759

Guth A. H., 1981, *Phys. Rev. D* 23, 347

Hanany S., et al., 2000, *Astrophys. J.* , 545, L5

Hansen F. K., Górski K. M., Hivon E., 2002, *MNRAS* 336, 1304

Hansen F. K., Banday A. J., Erisken H. K., Górski K. M., Lilje P. B, 2006, *Astrophys. J.* , 648, 784

Haslam C. G. T., Stoffel H., Salter C. J., Wilson W. E., 1982, *Astron. & Astrophys. Suppl.* 47, 1

Hildebrand R. H., et al., 1999, *Astrophys. J.* 516, 834

Hinshaw G., et al., 2006, *Astrophys. J.* 170, 288

Hobson R., et al., 1998, *MNRAS* 300, 1

Hu W., White M., 1997, *Phys. Rev. D* 52, 2

Huber P., 1985, *The annal of statistics*, 13(2), 435

Hummel E., et al., 1991, *Astron. & Astrophys.* 246, 10

Hyvärinen A., 1999, *IEEE Signal Processing Lett.* 6, 145

Hyvärinen A. & Oja E., 2000, Neural Networks, 13 (4-5), 411,

Jaffe A. H., et al., 2004, *Astrophys. J.* 615, 55

Jones T. J., Klebe K., Dickey J. M., 1995, *Astrophys. J.* 389, 602

Kamionkowski M., Kosowsky A., Stebbins A., 1997, *Phys. Rev. D* 55, 7368

Kovac J., et al., 2002, *Nature* 420, 772

Lagache G., et al., 2003, *Astron. & Astrophys.* 405, 813

Leach S., et al., 2008, `astro-ph/0`805.0269

Lewis A., Challinor A., Turok N., 2002, *Phys. Rev. D* 65, 023505

Lewis A., 2003, *Phys. Rev. D* 68, 083509

Lewis A., Challinor A., 2005, Physics Reports, 429, 1.

Liddle A., Lyth D. H., Cosmological Inflation and Large Scale Structure, Cambridge University Press (2000)

Maino D., et al., 2002, *MNRAS* 334, 53

Maino D., Banday A. J., Baccigalupi C., Perrotta F., Górski K., 2003, *MNRAS* 344, 544

Maino D., Donzelli S., Banday A. J., Stivoli F., Baccigalupi C., 2007, *MNRAS* , 374, 1207

Martínez-González E., et al., 2003, *MNRAS* , 345, 1101

Masi S., et al., 2001, *Astrophys. J.* , 553, 93

Mather J. C., et al., 1992, AIP Conference Proceedings Vol. 245, 266

Montroy T., et al., 2005, *Astrophys. J.* , 647, 813

Nolta M. R., et al., 2008, submitted to *Astrophys. J.*

Oxley P., et al., 2004, Earth Observing Systems IX. Proceedings of the SPIE, 5543, 320

Page L., et al., 2007, *Astrophys. J. Supp.* 170, 335

Penzias A. A., Wilson R. 1965, *Astrophys. J.* 142, 419

Perlmutter S., et al., *Astrophys. J.* 517, 565-586

Platania P., et al., 1998, *Astrophys. J.* 505, 473

Ponthieu N., et al., *Astron. & Astrophys.* 444, 327, 2005

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in C, Cambridge University Press

Prunet S., Sethi S. K., Bouchet F. R., Miville-Deschenes M. A., 1998, *Astron. & Astrophys.* 339, 187

Readhead A. C. S., et al., 2004, *Science* 306, 836

Reich P. & Reich W., 1988, *Astron. & Astrophys.* 74, 7

Reich P., et al., 2003, in The Magnetised Interstellar Medium, eds. Uyaniker, Reich, Wielebinski

Reichardt C., et al. 2008, astro-ph/0801.1491

Riess A.G., 1998, Bulletin of the American Astronomical Society, 30, 843

Rybicki G. B., Lightman A. P., 1979, Radiative Processes in Astrophysics

Sachs R. K.., Wolfe A. M., 1967, *Astrophys. J.* 147, 73

Serkowski K., et al. 1975, *Astrophys. J.* 196, 261

Smith K. M., 2006, *Phys. Rev. D* 74, 3002

Smoot G. et al., *Astrophys. J.* 396, L1

Spergel D. N. et al., 2003, *Astrophys. J. Supp.* 148, 175

Spergel D. N., and the WMAP collaboration 2007, *Astrophys. J.* 170, 377

Steigman G., 2007, Annual Reviews of Nuclear and Particle Science, 57, 463

Stivoli, F., Baccigalupi, C., Maino, D., & Stompor, R., 2006, *MNRAS* , 372, 615

Stivoli F., Stompor R., Leach S., Baccigalupi C., 2008, in preparation

Stompor R., Leach S., Stivoli F., Baccigalupi C., 2008, b, submitted to *MNRAS*

Stompor R., Stivoli F., Leach S., Baccigalupi C., 2008, a, in preparation

Tegmark M., 1997 *Phys. Rev. D* 55, 5895

Tegmark M., et al., 2004, *Astrophys. J.* 606, 702

Tucci M., et al., 2002, *Astrophys. J.* 579, 607

Ulvestad J. S., 1982 *Astrophys. J.* , 259, 96

Uyaniker B., Fürst E., Reich W., Reich P., Wielebinski R., 1999, *Astron. & Astrophys. Suppl.* 138, 31

Vio R., Andreani P., astro-ph/0802.0400

Watson R., et al., 2005, *Astrophys. J.* 624, 89

Wolleben M., et al., 2006, *Astron. & Astrophys.* 448, 411

Wright E. L., et al., 1991, *Astrophys. J.* 375, 608

Zaldarriaga M., Seljak U., 1997, *Phys. Rev. D* 55, 1830

Zaldarriaga M., Seljak U., 1998, *Phys. Rev. D* 58, 023003

Zaldarriaga M., 2001, *Phys. Rev. D* 64, 103001